Gödels's First Incompleteness Theorem

by
Tanu Prasad

A dissertation submitted in partial fulfilment of the requirements for the degree of
B.Sc Honours in Mathematics
2020-2023

Supervised by: Divakaran D
Evaluated by: Divakaran D and Ramanujam R

Gödel's First Incompleteness Theorem

Tanu Prasad

Abstract

In mathematical logic, one strives to understand if there exists a finite procedure to identify true statements. This leads to two other questions which are as follows: Does this procedure identify only true statements? And does it identify all true statements? The primary aim of this exposition is to understand Gödel's First Incompleteness Theorem. In layman's terms, the theorem states that given a "sufficiently complex" system of axioms, there exists a statement that is true but cannot be proved. But, one might suggest, why not include this true statement in the set of axioms? The catch here is that the expanded system of axioms is also sufficiently complicated. Thus, the proof of Gödel's First Incompleteness Theorem allows us to construct another statement that is true but cannot be proved from the expanded set of axioms. One might then ask, why not add all true statements? But, then given a statement, we won't be able to determine if the statement is an axiom or not and thus defeats the purpose.

Dedication

To Maa, Papa, Tatha and Milli. For being there always!

Declaration

I hereby declare that the work in this thesis has been carried out by me, in the B.Sc. (Honours) Program, under the supervision of Dr. Divakaran D, and in the partial fulfillment of the requirements for the award of the degree of B.Sc. (Honours) at the Azim Premji University, Bangalore. I further declare that this work has not been the basis for the award of any degree, diploma or any other title elsewhere.

Acknowledgements

I would like to express by deepest gratitude to my supervisor Dr. Divakaran D and Prof. Ramanujam R for their continuous guidance and support. This project would not have been possible without my family and friends who have supported emotionally throughout my journey. I would also like to thank my cat, *Mili* who has repeated mentions in my thesis.

Contents

Chapter 1

Introduction

In the field of mathematics, we have various mathematical structure like group, fields, rings etc which are defined based on some properties. What makes these properties useful for mathematicians and other practitioners in this subject is the truth and falsehood of these properties (which are essentially mathematical statements) which can be used to arrive at various other mathematical results. That is, say of instance we have to prove that -1 is not natural number. So, we know that $\forall x \in \mathbb{N}, x > 0$ and we have $0 > -1$. Hence, we can conclude that -1 is not a natural number. So, we have used a property possessed by all natural numbers to conclude that a given element is not a natural number. More so, not just in the field of mathematics, all of the above can also be applied to arrive at conclusions in real-life decision making. That is, say for instance, we believe that if it windy then it will rain and if it rains then there would be a power cut. From the two beliefs, we can infer that if it is windy then it would result in a power cut. It is important to notice that in both the above mentioned examples, we have used a set of assumptions to arrive at what we have to deduce.

But, have you ever wondered if truth and proof are related? And if so, how? Or even before that what does it mean to say that something is true? What does it mean to prove something? The statement of the First Incompleteness Theorem makes strong claims about how the notions of truth and proof are associated to each other. In layman terms, the First Incompleteness Theorem states the following: **There exists statements that are true but cannot be**

proved.

It is natural to think about the idea of completeness, given the Incompleteness theorem and if it exists, it is important to understand how can two contradicting concepts co-exist in a given paradigm. Completeness, as you may have guessed, claims that if a given statement is true then one can produce a proof for it. Proof and truth as concepts seem separate but in reality are very much in agreement with each other. The primary cause for this follows from the fact that checking truth involves providing a proof.

In this process of writing a proof, there are certain assumptions that one makes and these assumptions are used to arrive at the desired result. These assumptions can either be predefined or can be defined specifically for the intended proof. These assumptions are termed as *axioms*. Axioms are mathematical statements whose truth value is taken to be true. In short, the First Incompleteness Theorem says that the latter category of axioms (that do not form a part of our deductive system) cannot be used to axiomatize any true statement. However, this problem does not arise if we were to consider the predefined axioms (contained in our deductive system), thereby resulting in the coexistence of Incompleteness and Completeness. What are these two category of axioms? What is a truth value? What is a deductive system? The answers to all these questions form the contents of the subsequent chapters. Formally, the First Incompleteness Theorem, written in **first order language** is:
$\mathfrak{N} \models \phi$ but $\mathcal{A} \nvdash \phi$. Not to worry looking at the not-so-familiar looking symbols arranged in the above mentioned manner! The contents of this paper will help understand, in detail, the different tools required to explain the statement of the First Incompleteness Theorem. And because the above statement is in first order language, we would first make sense this language. But before moving forward to understanding the same, it is a good time to pause a look at why do we even need a language and what makes up a language.

As mentioned previously, every mathematical structure is described using a bunch of properties. These properties can be seen as a string of various symbols which have a predefined meaning so as to make sure that these are interpreted in the required way only. Like for instance, consider the following example-Taking the commutativity axiom, defined in $\mathbb{R}^2$ and we interpret the $+$

sign to be as follows: Let us consider $a, b \in \mathbb{R}^2$ such that $a = (x_1, y_1)$ and $b = (x_2, y_2)$. Now, we define $(x_1, y_1) + (x_2, y_2) = (x_1y_2, y_1x_2)$. Taking LHS, we have, $(x_1, y_1) + (x_2, y_2) = (x_1y_2, y_1x_2)$ and taking RHS, we have, $(x_2, y_2) + (x_1, y_1) = (x_2y_1, y_2x_1)$. We observe that the commutativity property no longer holds true. So, we observe that **when we change the interpretations of a particular symbol in a statement, we see that the truth value of the statement also changes.** So, what is communicated, will not perceived the way it should be. Hence, it is important to define a language.

Now, given any language, it is necessary to define its vocabulary and grammar (which is the syntax) along with the meaning of the lexicon defined(which is the semantics). That is, simply put, a meaning given to a string of symbols arranged in a particular order - makes up a language. Because we are familiar with the English language, we would use this language to make sense of what is syntax and the semantics.

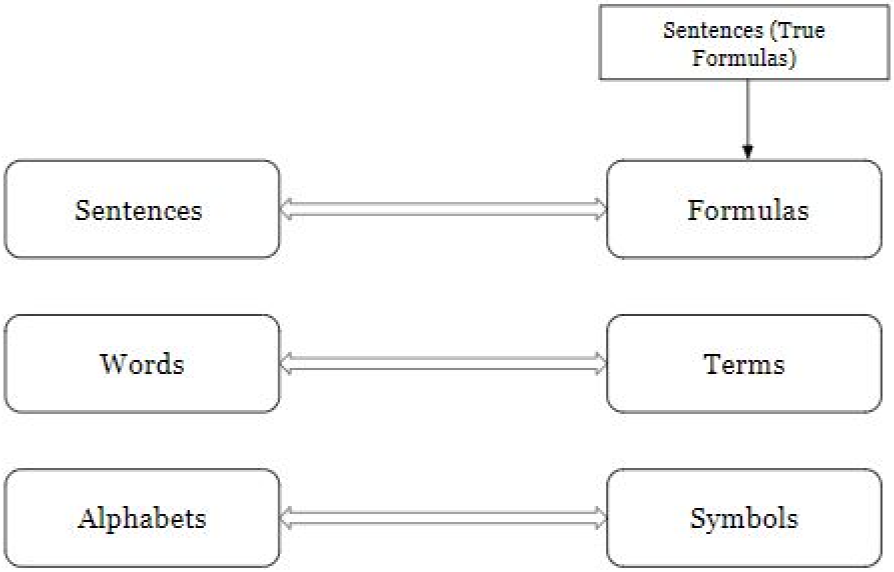
Consider the following sentence in English: *Meet my cat, Mili!.* The sentence is the result of effective/acceptable ordering to the words: {”*Meet*”,”*my*”,”*cat*”,”*Mili*”,”,”,”!”} where in, each of these words are the result of effective/acceptable structuring applied to the alphabets: {”*m*”,”*e*”,”*t*”,”*c*”,”*a*”,”*y*”,”*i*”,”*l*”,”*i*”,”,”,”!”}. By this, we mean that neither *cat Meet, my Mili!* nor *Mtee my cta!, Mili* are sentences. This is because the former is the result of unacceptable ordering of the words (that makes up a sentence), that is, the arrangement ”*cat Meet, my Mili!*” is grammatically wrong. While the latter is the result of unacceptable ordering of the alphabets (that make up the words), that is, ”*cta*” is not a word in English. So, the alphabets could be thought as the smallest units of the English language. It is syntax of the English language that accounts for this acceptable arrangement of the alphabets.

To explain the semantics of a language, consider this sentence, given by Noam Chomsky, an American Linguist, in the year 1957 in English, which goes as: *Colorless green ideas sleep furiously.* This sentence here is syntactically correct, but if looked at semantically, one cannot make any sense of it. It is important to notice that a sentence that does not make semantic sense is different from a sentence that is semantically not true. Consider the sentence: *Mili, my cat, recites poems to me every night!* If one goes to understand what it means, it is clear that the above sentence is not true in the real-world, but it might very well be the case that cats recite poems in the fairy tale world. So, syntactically, the sentence is correct, the sentence also makes semantic sense, but is untrue.

If we look at mathematical statements, say for instance: $1 + 1 = 2$. Similar to previous examples, syntactically, it is the symbols $1, +, 1, =, 2$ placed one after the other, with no meaning associated to these symbols. But, semantically, we are adding 1 to itself and calling this new number to be 2.

Now that we are well aware of what makes up a language, to able to decode what the statement First Incompleteness Theorem means, the line up of tasks that we need to accomplish is as follows: First understand the definition of the first order language, following which we would look at what truth and proof means in the paradigm of first order logic, which will in turn help produce the proof of the First Incompleteness Theorem.

Using our prior knowledge of the English Language, we have the following analogy with the first order language:



As mentioned previously, in the English Language, we use sentences that are made up of words and each of these words are made of alphabets. Similarly, the first order language consists of symbols(which are analogous to alphabets) make up *terms* (which are analogous to words in

English) and these terms form the *formulas* (that are analogous to sentences in English).

We have three new words - symbols, terms, formulas introduced above and understanding each of this in detail, in addition to getting a clear sense of what it means to conclude the truth or falsehood of a given statement, follows as the content of the next chapter. All the definitions and theorems (along with the proofs) have been studied and presented from the book Leary and Kristiansen 2015

Chapter 2

Syntax + Semantics = First Order Language

2.1 Syntax

2.1.1 First Order Language, $\mathcal{L}$

When we say 'write down in terms of a given language' - the first step is to identify the set of symbols that are used to construct the statements following which we need clearly indicate what each of these symbols means, when put together to form the statements. So, formally,

Definition 1. *The first order language, denoted by $\mathcal{L}$, is* ***an infinite collection of symbols such that none of the symbols are properly contained in another.*** *The symbols can be categorized as follows:*

- *Logical connectives (eg: $\vee$, $\wedge$, $\neg$, $\Rightarrow$, $\Leftrightarrow$)*
- *Parentheses*
- *Variables*
- *Quantifiers ($\forall, \exists$)*
- *Constant symbols*

- *Function symbols*
- *Relation symbols*
- *Equality symbol*

Besides the various classifications of symbols, so as to be considered as symbols in the first-order language, one important property possessed by these symbols is the notion of the symbols not being properly contained in one another. To explain what this means, using an example, let us consider the symbol $*$, which is used in various programming languages - This symbol is used to perform the multiplication operation between any two numerical inputs provided. We also have the $**$ symbol that is used to perform the 'raise to' operation between two integers. Notice that $*$ is properly contained in $**$, that is $**$ can also be seen as $*$ performed twice. This creates confusion as to how should $**$ be interpreted. That is, is it a new operation or is it performing multiplication two times. Hence, considering the way first-order language is defined, we cannot have both $**$ and $*$ present in the same set.

Here, the parentheses act as punctuation in a sentence or a statement. In most cases, parentheses, the equality symbol, logical connectives, and quantifiers are included by default in the set. Hence, they are not mentioned explicitly always. Also, do not confuse constant symbols with numbers! Given the notion of the mathematical language we are used to, it is natural to associate only numbers with constant symbols. Numbers can be a type of constant symbol but are not the only constant symbols that exist. This also means that not all numbers are constant symbols. It completely depends on how the language is defined. That is, say for instance, we can define a language where the alphabets in English are the constant symbols.

One of the crucial things to understand is the difference between a function and a function symbol itself. A function can be thought of as a machine (say by the name f) that makes modifications to what it is fed that is, the input(s), thereby producing a new product, which is the output. Now, the machine might require one or more raw materials depending on what it has to produce. Similarly, in mathematics, a function represented as f may require one or more inputs(or arguments in precise mathematical language). Based on the number of arguments required, we can categorize functions. Functions that require one argument are said to be arity

1 functions or Unary functions, functions with two arguments are said to be functions with arity 2 or Binary functions, and so on. Also, the arity of the function symbol is the same as that of the function. As you may have guessed, all the above explanations can be applied to relations and relation symbols, with the only difference being the output of the relation - a truth value (that is, true/false). Same as functions, a set with 2-tuples is said to be an arity 2 relation. So, essentially, function symbols and relation symbols represent functions and relations respectively.

Example 2.1.1. *The language set used to describe Number theory, $\mathcal{L}_{NT}$ is defined as follows:* $\{0, E, +, \cdot, S, >\}$, *where,*

- 0 *is a constant symbol*
- S *is unary function symbol that represents the successor function.*
- $+, \cdot, E$ *are binary function symbols that represent addition, multiplication, and exponentiation, respectively between two integers.*
- $>$ *is a binary relation symbol that represents the "is greater than" relation between two integers.*

2.1.2 Terms in $\mathcal{L}$

Broadly put, terms are a sequence of selected symbols (in $\mathcal{L}$) arranged in a particular way such that together they constitute a different meaning as opposed to when called individually.

Definition 2. *A term of $\mathcal{L}$ is a nonempty finite string t, consisting of symbols from $\mathcal{L}$, such that:*

- t *is a variable or*
- t *is a constant or*
- t *is* $ft_1t_2t_3....t_n$ *where* f *is a n-arity function symbol and* t_i *is a term in* $\mathcal{L}$ *where* $i = 1, 2, 3, 4....., n$

Here, the third classification of the term can also be stated using a meta-linguistic symbol. Meta-linguistic symbols are those symbols contained in a language that is used to describe another language, that is, $t :\equiv ft_1t_2t_3....t_n$ where f is an n-arity function symbol and t_i is a term in $\mathcal{L}$ where $i = 1, 2, 3, 4....., n$ and $\equiv$ is the meta-linguistic symbol. Also, we see that the third classification of the term has been defined recursively, where t is obtained by performing some operation on the term(s) of $\mathcal{L}$. From here on, in the subsequent chapters of this thesis, we would extensively use this idea of recursively defining certain concepts.

Example 2.1.2. *Considering the language of number theory, $\mathcal{L}_{NT}$ defined previously:*

- 0 *is a term*
- x *is a term*
- $Sx, S0, SS0$ *are terms*
- $E(SS0)(x)$ *is a term*

2.1.3 Formulas in $\mathcal{L}$

As mentioned previously, an extension to terms, formulas are analogous to sentences.

Definition 3. *A formula ϕ of $\mathcal{L}$ is a nonempty finite string of symbols in $\mathcal{L}$ such that:*

- $\phi :\equiv = t_1t_2$, *where t_1, t_2 are terms in $\mathcal{L}$*
- $\phi :\equiv Rt_1t_2t_3....t_n$ *where R is a n-arity relation symbol in $\mathcal{L}$ and t_i is a term in $\mathcal{L}$ where $i = 1, 2, 3, 4....., n$*
- $\phi :\equiv (\neg\alpha)$, *where α is a formula in $\mathcal{L}$*
- $\phi :\equiv (\alpha \vee \beta)$, *where α and β are formulas in $\mathcal{L}$*
- $\phi :\equiv (\forall v)(\alpha)$, *where v is a variable in $\mathcal{L}$ and α is formula in $\mathcal{L}$*

The formulas that are not recursively defined as termed atomic formulas. So, the first two classifications of formulas, in the above definition are the atomic formulas, and using these we can construct formulas that belong to one of the remaining three categorizations. One interesting distinction between terms and formulas of the first-order language is the particular use of function symbols in terms and relation symbols in formulas. It is reasonable to question why is it the way it is or if could it have been the other way around. The answer to this is the following - Terms do not have truth value. But Formulas do. The existence of relation symbols causes that particular entity to possess a truth value and one of the larger aims is to study the truth and falsehood of mathematical statements. Another way to look at it is as follows - Function symbols are already included in the formulas as the result of terms.

Example 2.1.3. *Considering the language of number theory, $\mathcal{L}_{NT}$ defined previously*

- $\phi :\equiv (\neg(x > 0))$ *is a formula in* $\mathcal{L}_{NT}$
- $\phi :\equiv = (E(SS0)(x))(SS0)$ *is a formula in* $\mathcal{L}_{NT}$
- $\phi :\equiv (\forall x)(x > 0)$ *is a formula in* $\mathcal{L}_{NT}$

A part of what makes the contents of this chapter difficult to comprehend, despite it sounding very obvious at various junctures, is the difference in the notations used to describe the "very obvious". In the everyday mathematics that we do, we are used to the *infix notation.* To explain with an example, we are habituated to writing the arithmetic operator (like $+, \times$) between its operands. Like, for instance, $3 + 2$. But here, the notation followed is $+32$, provided the symbols $3, 2, +$ are included in the language set. The former follows *infix notation* while the latter follows *polish notation.* So far, we have been using the Polish notation, in order to maintain the consistent behavior of the operator and to enable easy splitting up of terms in a given expression. Not to worry though, we would eventually get to actively use what we are familiar with, which is the infix notation.

2.1.4 Sentences in $\mathcal{L}$

A subset of the set of $\mathcal{L}$-formulas, sentences are those formulas that have a definite truth value, that is, formally, these are the formulas that do not have *free variables* contained in them.

The existence of free variables in a formula causes the truth value of the formula to be indeterministic. We have defined sentences to be those formulas that are either true or false. So, it is important that these formulas do not contain any free variables. So, all sentences are formulas but all formulas are not sentences.

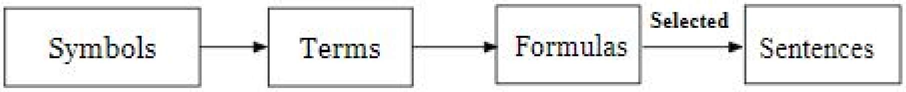
Free Variables

Formally, if we are given a variable v and a formula because we have two classifications of formulas - the first being the atomic formulas and the second being the recursively defined formulas, we will have to describe the idea of variables being free in both these kinds. Intuitively, this can be explained as follows - Consider the two $\mathcal{L}_{NT}$ formulas $> (S(S(0)))(0)$ and $> (E(S(S(0)))(x))(S(0)) \vee (= (S(0))(0))$. Intuitively, we can see that the first essentially checks $1 > 0$ while the second checks either $2^x > 1$ or $1 = 0$. We can easily determine the truth value of the first formula because there are no variables in it. In the case of the second formula, it is of the form $\phi :\equiv \alpha \vee \beta$, where $\alpha :\equiv (E(S(S(0)))(x))(S(0))$ and $\beta :\equiv = (S(0))(0)$. The truth value of ϕ depends on the truth values α and β. But here, the truth value of α is dependent on the value of x, thereby making its truth value indefinite. That is, α is true if x takes any value greater than or equal to 2 but when $x = 1$, α is false.

Definition 4. *Given a variable v and a $\mathcal{L}$-formula ϕ, v is free in ϕ if:*

- *ϕ is an atomic formula and v occurs in ϕ or*
- *ϕ is not atomic, $\phi :\equiv (\neg\alpha)$ and v is free in α or*
- *ϕ is not atomic, $\phi :\equiv \alpha \vee \beta$ and v is free in either α or β or*
- *ϕ is not atomic, $\phi :\equiv (\forall u)(\alpha)$ and v is not in u and v is free in α.*

So, diagrammatically, we have,



Example 2.1.4. *Consider a formula* $\phi :\equiv (\forall v_1)(\neg(\forall v_5)(= (v_2)(+v_1v_5)))$*, that is defined in a language that contains all the necessary symbols for the above formula to be defined. We need to check if any free variables are contained in* ϕ*, in order to conclude whether it is a sentence or not. We have* $\phi :\equiv (\forall v_1)(\alpha)$*, where* $\alpha :\equiv (\neg(\forall v_5)(\beta))$ *and* $\beta :\equiv = (v_2)(+v_1v_5)$*. Considering* β*, because it is an atomic formula, we can conclude that* v_2, v_1, v_5 *are all free in* β*. Now, considering* α*, by definition, because* v_1, v_2 *are not equal to* v_5 *and* v_1, v_2 *are free in* β*, we can infer that* v_1, v_2 *are free in* α*. Following from this, because* v_2 *is not equal to* v_1 *and* v_2 *is free in* α*, the only free variable in* ϕ *is* v_2*.*

Now that we have the syntax in place, the contents of the next section would get us to conclude that $\mathcal{L}_{NT}$-*formulas* of the kind $\phi :\equiv = (SSS0)(+(SS0)(S0))$ are indeed same as $3 = 2 + 1$. At the end of the next section, the meaning of each of these symbols, contained in the set $\mathcal{L}_{NT}$ would be well defined.

2.2 Semantics

As mentioned previously, our language set $\mathcal{L}_{NT}$ constitutes the function **symbols** and relation *symbols* but not the function and relation itself. That is, if we observe, nowhere have we mentioned that $SS0$ is equal to 2, yet. This is precise because 2 or any constants are not contained in our language set, $\mathcal{L}_{NT}$. Not to confuse constants with constants symbols here. The constant symbol 0 is contained in $\mathcal{L}_{NT}$, not the constant 0 itself. The same would apply to all the constants generated when any of the function symbols are translated to the function it represents, thereby producing the output. In order to be able to conclude the obvious, we first need to fix a set that contains all these constants, and then considering the arity of the function and relation symbols, we need to precisely define the function and relation on this set of constants.

Definition 5. *Formally, for a fixed language, say* $\mathcal{L}$*, we have* $\mathcal{L}-$ *structure denoted by* $\mathfrak{N}$*, is a nonempty set* A *(the set of constants) called the universe of* $\mathfrak{N}$ *along with:*

- *For every constant symbol* $c \in \mathcal{L}$*, we have* $c^{\mathfrak{N}} \in A$

- *For a n-arity function symbol, say* f *of* $\mathcal{L}$*, we have* $f^{\mathfrak{N}} : A^n \to A$
- *For a n-arity relation symbol, say* R *of* $\mathcal{L}$*, we have a relation,* $R^{\mathfrak{N}} \subset A^n$

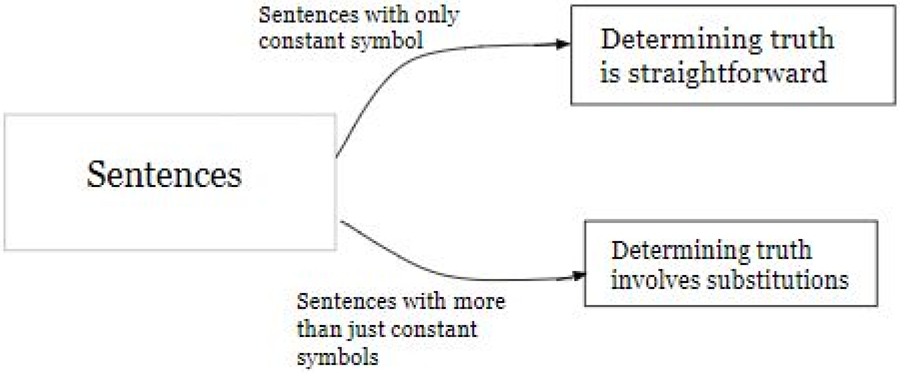
The representation of the structure is similar to that of groups or rings, that is, it is an ordered tuple with the universe as its first element, followed by the functions defined on the universe (that corresponds to the function symbol in the language set) and then the relations defined on the universe (that corresponds to the relation symbol in the language set).

Example 2.2.1. *Given* $\mathcal{L}_{NT}$*, we have* $\mathfrak{N}$ *to be a* $\mathcal{L}_{NT}$ *structure that is represented as follows:* $(\mathbb{N}, 0^{\mathfrak{N}}, S^{\mathfrak{N}}, +^{\mathfrak{N}}, \cdot^{\mathfrak{N}}, E^{\mathfrak{N}}, >^{\mathfrak{N}})$

- $\mathbb{N}$ *is the universe, where* $0^{\mathfrak{N}}$ *represented by* 0 *in* $\mathcal{L}_{NT}$*, will be contained in* $\mathbb{N}$
- *The Unary function* $S^{\mathfrak{N}} : \mathbb{N} \to \mathbb{N}, S^{\mathfrak{N}}(x) = x + 1$
- *The binary function* $+^{\mathfrak{N}} : \mathbb{N}^2 \to \mathbb{N}, +^{\mathfrak{N}}(x, y) = x + y$
- *The binary function* $\cdot^{\mathfrak{N}} : \mathbb{N}^2 \to \mathbb{N}, \cdot^{\mathfrak{N}}(x, y) = xy$
- *The binary function* $E^{\mathfrak{N}} : \mathbb{N}^2 \to \mathbb{N}, E^{\mathfrak{N}}(x, y) = x^y$
- *The binary relation* $>^{\mathfrak{N}}$ *that is defined as follows: Given two numbers,* $a, b \in \mathbb{N}$*, we say* $(a, b) \in >^{\mathfrak{N}} \subset \mathbb{N}^2$ *iff* $a > b$*.*

2.3 Truth and Falsehood

Given a $\mathcal{L}_{NT}$-sentence, we now know what the sentence means. The next step is to determine if this sentence is true. The sentences in first-order logic can be broadly categorized into two types and we need to check the truth in both these kinds. Pictorially, the intuition can be represented as follows:



Identifying the truth or falsehood of sentences with variables is tricky and it requires substitution, to arrive at a conclusion. That is, for sentences that have variables, the only way to learn the truth of the sentence, is by assigning some constant value to the variable and then checking whether the sentence holds true for this value. In the paradigm of first-order logic, this assignment can be made precise, which is as follows:

Definition 6. *Consider $\mathfrak{N}$ to be the structure of some language $\mathcal{L}$. Now, we have a function $s : Vars \to A$, where $Vars$ is the set of all variables in $\mathcal{L}$ and A is the universe on which the structure is defined. Here, the function s assigns each variable in $\mathcal{L}$ to some element in the universe. This function can be modified to address situations when we want certain variables to take certain fixed values from the universe. The x-modification of the variable assignment function is defined as follows:*

$$s[x|a](v) = \begin{cases} s(v) & v \neq x \\ a & v = x \end{cases}$$

The $x-$modification of the variable assignment function asserts a fixed value to a particular variable. This can be done for multiple variables in the language. So, in a way, you can prove $x = y$! Though x, y are syntactically different, this does not necessarily imply that the formula $x = y$ is semantically false. It solely depends on the value assigned to this variable in a particular structure. Similar to the variable assignment function, we also have the term assignment function that is defined as follows:

Definition 7. *Consider $\mathfrak{N}$ to be the structure of some language $\mathcal{L}$. Now, we have a function $\bar{s} : T \rightarrow A$, where T is the set of terms in $\mathcal{L}$ and A is the universe on which the structure is defined.*

- $\bar{s}(t) = s(t); t$ *is variable*
- $\bar{s}(t) = c^{\mathfrak{N}}; t$ *is a constant symbol*
- $\bar{s}(ft_1t_2t_3..t_n) = f^{\mathfrak{N}}(\bar{s}(t_1), \bar{s}(t_2), \bar{s}(t_3), .., \bar{s}(t_n)); t :\equiv ft_1t_2t_3...t_n$, *where* $t_1, t_2, t_3, .., t_n$ *are all terms in* $\mathcal{L}$.

Remark 1. *Once the variable assignment function is defined, the term assignment function is defined by default.*

We now have all the prerequisites to conclude whether a sentence is true in a given structure. That is, we can explain the truth or the falsehood of a $\mathcal{L}-$ formula ϕ, or specifically the truth of sentences in $\mathfrak{N}-$structure.

Definition 8. *Consider the following to be predefined-$\mathfrak{N}$ is $\mathcal{L}-$structure and $s : Vars \rightarrow A$ is the assignment function into $\mathfrak{A}$. We can can conclude $\mathfrak{N} \models \phi[s]$ read as "ϕ satisfies $\mathfrak{N}$ with the variable assignment function s" or "ϕ is true in $\mathfrak{N}$ under the variable assignment function s" when:*

- *If* $\phi :\equiv = t_1t_2$ *then* $\bar{s}(t_1) = \bar{s}(t_2)$
- *If* $\phi :\equiv Rt_1t_2t_3....t_n$ *where* $t_1, t_2, t_3, ..., t_n$ *are* $\mathcal{L}-$ *terms*
  then $(\bar{s}(t_1), \bar{s}(t_2), \bar{s}(t_3), ..., \bar{s}(t_n)) \in R^{\mathfrak{N}}$.
- *If* $\phi :\equiv (\neg\alpha)$ *then,*$\mathfrak{N} \models \phi[s]$ *iff* $\mathfrak{N} \not\models \alpha[s]$.
- *If* $\phi :\equiv (\alpha \vee \beta)$ *then* $\mathfrak{N} \models \alpha[s]$ *or* $\mathfrak{N} \models \beta[s]$ *or both.*
- *If* $\phi :\equiv (\forall x)(\alpha)$, *then for each element* $a \in A$, $\mathfrak{N} \models \alpha[s(x|a)]$

Corollary 1. *Given a $\mathcal{L}-$ formula, ϕ and $\mathcal{L}-$ structure, $\mathfrak{N}$, $\mathfrak{N} \models \phi[s]$ for every assignment function s or $\mathfrak{N} \models \phi[s]$ for no assignment function s.*

There are a number of interesting consequences of satisfiability. A well-defined structure enables an effective mapping between the set of sentences to $\{0, 1\}$ where 0 implies false and 1 implies true. If a formula is true in every structure under every assignment function, the formula is said to be **valid** (eg: $x = x$).

Example 2.3.1. *Consider* $\mathcal{L}_{NT}$ *and* $\mathfrak{N}$ *to be the* $\mathcal{L}_{NT}-$ *structure as defined in* ***Example 2.2.1****. Consider the variable assignment function that sends all variables to 0. We will check the truth of two statements* $=(SSSS0)(E(SSS0)(x))$ *and* $>(SSS0)(SS0)$*. Intuitively, we can conclude that the first formula is not true while the second is true. But, let us show that our intuition is correct, given the variable assignment function.*

$\phi_1 :\equiv= (SSSS0)(E(SSS0)(x))$ *which is equivalent to* $\phi_1 :\equiv= t_1t_2$*, where* $t_1 = SSSS0$ *and* $t_2 = E(SSS0)(x)$*. So,* $\bar{s}(t_1)$ *is* $S^{\mathfrak{N}}(S^{\mathfrak{N}}(S^{\mathfrak{N}}(S^{\mathfrak{N}}(0^{\mathfrak{N}}))))$ *which is equal to* 4*. While* $\bar{s}(t_2)$ *is* $E^{\mathfrak{N}}(S^{\mathfrak{N}}(S^{\mathfrak{N}}(S^{\mathfrak{N}}(0^{\mathfrak{N}}))), \bar{s}(x))$ *which is* $E^{\mathfrak{N}}(3, 1)$ *which is equal to* 3*. Hence,* $\mathfrak{N} \not\models \phi_1[s]$

$\phi_2 :\equiv> (SSS0)(SS0)$ *which is equivalent to* $\phi_1 :\equiv> t_1t_2$*, where* $t_1 = SSS0$ *and* $t_2 = SS0$*. So,* $\bar{s}(t_1)$ *is* $(S^{\mathfrak{N}}(S^{\mathfrak{N}}(S^{\mathfrak{N}}(0^{\mathfrak{N}})))$ *which is equal to* 3*. While* $\bar{s}(t_2)$ *is* $S^{\mathfrak{N}}(S^{\mathfrak{N}}(0^{\mathfrak{N}}))$ *which is equal to* 2*. Now,* $>^{\mathfrak{N}}= \{(a, b)|a > b, a, b \in \mathbb{N}\} \subset \mathbb{N}^2$*. We have* $(\bar{s}(t_1), \bar{s}(t_2)) = (3, 2) \in >^{\mathfrak{N}}$ *Hence,* $\mathfrak{N} \models \phi_1[s]$*.*

We now have a clear sense of what a true statement is, in first-order logic. The contents of the subsequent chapter will understand what is proof.

Chapter 3

Proof is same as deduction

3.1 Proof: A prelude to Deductions

Consider that the following statement is known to you- If a cat tucks its tail then it is sad. A few moments later, you observe *Mili* sitting in the corner, curled up. If you happen to love cats the way I do too, you would probably consider letting the owner know that *Mili* is sad. This scenario is the application of classic inference rule-Modes Ponens. Here, essentially you have deduced that *Mili* is sad, using a collection of statements.

Taking a mathematical example, suppose that you have to show that given $a \in \mathbb{N}, a \cdot 0 = 0$. Consider the LHS, $a \cdot 0$ which is same as $a(1-1)$. This can also be written as $a(1+(-1))$ and because we know that addition and multiplication is distributive in $\mathbb{N}$, we can write it as $a(1) + a(-1) = a + (-a)$. Now, because $-a$ is the additive inverse of a, we can conclude that $a + (-a) = 0$. Hence, we have proved that $a \cdot 0 = 0$, that is, we deduced that $a \cdot 0 = 0$ using the distributive property of natural numbers and using the fact that every natural number has an additive inverse.

At this stage, we clearly understand what is a true statement and we have an intuitive idea that proving/deducing a statement is the process of arriving at the desired conclusion by working on some set of statements. But you may still ask, how is proof related to truth? To answer this question, we must go back to what we had started with. That is, we are interested in determining the truth of mathematical statements. For doing so, we first need to identify this collection of statements (which forms the content of the second chapter). One of the ways to check the truth is to give proof. So proving a statement confirms the fact that the statement is indeed true. That is, to prove a statement is equivalent to providing substantial evidence in order to accept the truth of the statement made. These statements that we use to arrive at the desired result are called axioms which are taken to be true. So, formally,

Definition 9. *Given a collection of $\mathcal{L}_{NT}$ formulas Σ, D is a finite sequence of $\mathcal{L}_{NT}$formulas $(\phi_1, \phi_2, ..., \phi_n)$ such that each of ϕ_i, where $1 \leq i \leq n$ is:*

1. *$\phi_i \in A$ where A is the set of logical axioms*
2. *$\phi_i \in \Sigma$ where Σ is the set of non-logical axioms*
3. *ϕ_i is inferred from the set of the inference rules, where (Γ, ϕ_i) is an inference rule such that $\Gamma \subset \{\phi_1, ..., \phi_{i-1}\}$*

We say $\Sigma \vdash D$, read as "D is a deduction from Σ" or "D is deduced from Σ".

The logical axioms and inference rules form the fundamentals of the deductive system. Depending on what we are studying or what we want to prove, the set of non-logical axioms vary and hence do not form a part of the deductive system. For instance, if we circle back to the proof of $a \cdot 0 = 0$, the axiom states that every natural number a has an additive inverse and the distributive property of natural numbers belongs to the set of non-logical axioms. The subsequent sections provide detailed explanations of what are these collections of statements and inference rules, which would in turn help us prove the deduction theorem.

3.2 Logical Axioms

Logical axioms can be categorized into types-axioms uses the equality symbol and axioms that use quantifiers. There are a total of 5 logical axioms where three illustrate the effective use of the equality symbol while the remaining two make use of quantifiers. We will now look at each of these in detail.

3.2.1 Axioms that use Equality

The three axioms are as follows:

1. $x = x$ for each variable x.

2. Given f is a function symbol of arity n, $[(x_1 = y_1) \wedge ...... \wedge (x_n = y_n)] \implies [f(x_1, ..., x_n) = f(y_1, ...y_n)]$

3. Given R is a relation symbol of arity n, $[(x_1 = y_1) \wedge ...... \wedge (x_n = y_n)] \implies [R(x_1, ..., x_n) \implies R(y_1, ...y_n)]$

3.2.2 Axioms that use quantifiers

The two axioms are as follows: Given a $\mathcal{L}_{NT}-$ formula α,

1. $(\forall x \alpha(x)) \implies \alpha(t)$. This axiom concludes the following- if a formula α holds true for every value in the domain then the α also holds true if we substitute x with a particular element in the domain.

   Example 3.2.1. *If all students at Azim Premji University have done the Field Immersion, then Tanu has done a Field Immersion.*

   Example 3.2.2. *If all natural numbers are greater than or equal to 0 then 1 is greater than or equal to 0.*

2. $\alpha(t) \implies (\exists x \alpha(x))$. This axiom concludes the following- if the formula α holds true for a particular element in the domain then there happens to be an element in the domain for which the α holds true.

   Example 3.2.3. *If Tanu has done a Field Immersion, then there is a student at Azim Premji University who has done the field Immersion.*

   Example 3.2.4. *If 1 is greater than or equal to 0 then there exists an element in the set of natural numbers that is greater than or equal to 0.*

3.3 Inference Rules

It is crucial to have the rule of inference set in place. This is because, without the rule of inference, the most trivial formulas cannot be concluded from A and $\sum$. For instance, let $\sum = \{A_1, A_2\}$, where,

A1: All cats are animals

A2: Animals have four limbs.

The claim to be proven is: All cats have four limbs. We can see that A_1 is a formula of the form, $P \to Q$ and A_2 is a formula of the form $Q \to R$. Intuitively, using the idea of transitivity, we can deduce the claim. That is, given- $P \to Q$ and $Q \to R$, we can conclude that $P \to R$. But, to be able to do this, it should comply with a rule of inference that is set. So, only if there exists $\Gamma = \{P \to Q, Q \to R\}$ and $\phi := P \to R$, we can conclude that all cats have four limbs. That is, given all premises (all formulas in Γ) are true, we can infer that the conclusion(ϕ) is also true.

To generalize conclusions of the above kind in the paradigm of first-order logic, we broadly have two types of inference rules: Propositional Consequence (PC) and Quantifier Rule (QR), where PC, explains best explains the example.

3.3.1 Inference rule of type (PC)

To check if a formula is a propositional consequence, the first-order logic problem must be first converted into a propositional logic problem, that is, construct a propositional formula. To construct a propositional formula, we need to identify the propositional variables. To do the same, we first prioritize formulas with quantifiers contained in them and replace them with propositional variables and then proceed to replace atomic formulas with new propositional variables. At the end of this, we have the α_p to the propositional formula $\mathcal{L}_{NT}-$ formula α.

Example 3.3.1. *Let* $\alpha :\equiv (S(x) \vee Q(x)) \implies (\forall x P(x) \vee S(x))$ *be a* $\mathcal{L}_{NT}-$ *formula. A step by step translation* α *to* α_p *is as follows:*

$\alpha :\equiv (S(x) \vee Q(x)) \implies (\forall x P(x) \vee S(x))$

$\equiv (S(x) \vee Q(x)) \implies (A \vee S(x))$

$\equiv (B \vee C) \implies (A \vee B) \equiv \alpha_p$

We then check if the propositional formula is a propositional consequence of another set of propositional formulas.

Definition 10. *Let* Γ_p *be the set of propositional formulas and* ϕ_p *be the propositional formula that is inferred from* Γ_p*.* ϕ_p *is a propositional consequence of* Γ_p *when for every truth assignment that makes* Γ_p *true, also makes* ϕ_p *true.*

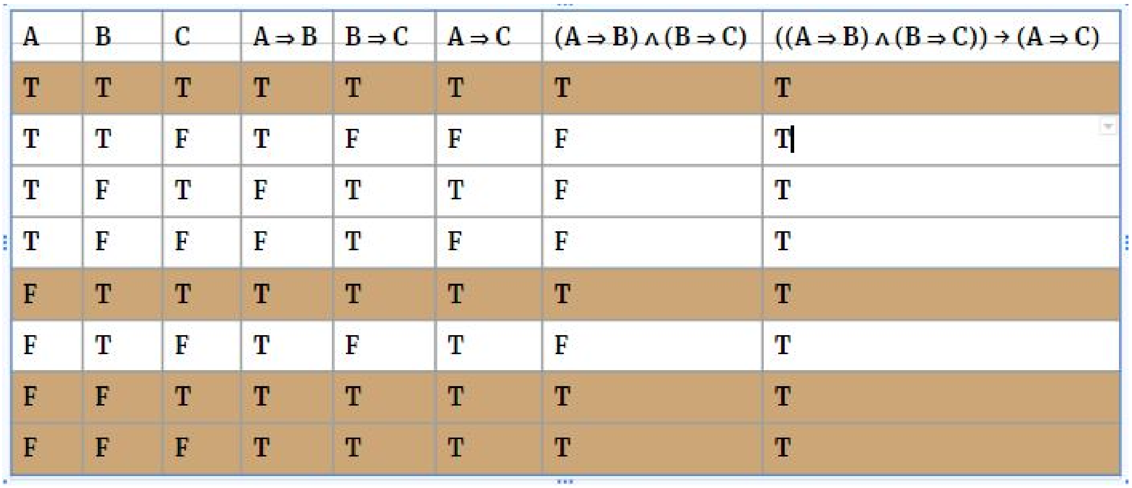
The above definition uses the idea behind a conditional statement. If a given statement $P \implies Q$ or If P then Q, we can infer that Q is a consequence of P. That is, whenever the antecedent P is true, the consequent Q is also true, the only difference being that the P in the above definition is a collection of statements.

The above definition also introduces a new word, that is *truth assignment* which is a function from the set of the proposition variables to $\{T, F\}$, that is a function that assigns a truth value to the proposition variables based on the truth value of the proposition variables, we can define another function that assigns a truth value to the propositional formula that follows from our usual interpretation of the logical connectives.

One of the easy ways to check if ϕ_p is a propositional consequence of Γ_p is given by:

Lemma 2. *Let* $\Gamma_P = \{\gamma_{p1}, \gamma_{p2}, ..., \gamma_{pn}\}$*.* ϕ_p *is a a propositional consequence of* Γ_p *is* $[\gamma_{p1} \vee \gamma_{p2} \vee ... \vee \gamma_{pn}] \implies \phi_p$ *is a tautology.*

Example 3.3.2. *Consider we need to if* $\phi :\equiv A \implies C$ *is a propositional consequence of* $\Gamma = \{(A \implies B), (B \implies C)\}$*. Following from the above lemma, we need to check if* $[(A \implies B) \vee (B \implies C)] \implies (A \implies C)$ *is a tautology. Verifying the same using a truth table, we would get the following:*



A	B	C	$A \Rightarrow B$	$B \Rightarrow C$	$A \Rightarrow C$	$(A \Rightarrow B) \wedge (B \Rightarrow C)$	$((A \Rightarrow B) \wedge (B \Rightarrow C)) \rightarrow (A \Rightarrow C)$
T	T	T	T	T	T	T	T
T	T	F	T	F	F	F	T
T	F	T	F	T	T	F	T
T	F	F	F	T	F	F	T
F	T	T	T	T	T	T	T
F	T	F	T	F	T	F	T
F	F	T	T	T	T	T	T
F	F	F	T	T	T	T	T

The highlighted rows explain the cases when the antecedent is true, the consequent is certainly true. The last column abides by the lemma, hence confirming that ϕ here is indeed a propositional consequence of Γ

Definition 11. *Given* $\Gamma = \{\gamma_1, ..., \gamma_n\}$ *and* $\phi :\equiv \phi$ *and* ϕ_p *is a propositional consequence of* Γ_p*, we can conclude that* ϕ *is propositional consequence of* Γ.

3.3.2 Inference rule of type (QR)

These can be further categorized into two types:

1. Type 1: Suppose you prove some property for an arbitrary element in a set. You should be able to conclude that this property holds for all elements of that set.

   Formally,

   Definition 12. *Assuming* x *is not free in* ψ*, we have* $\Gamma = \{\psi \rightarrow \phi\}$ *and* $\alpha :\equiv (\psi \rightarrow (\forall x)\phi)$ *and* (Γ, α) *is an inference rule of the type (QR)*

   Example 3.3.3.

2. Type 2: Suppose that a particular element of the set possesses a property. You should be able to conclude that there exists that particular element in that set for which the property holds.

   Formally,

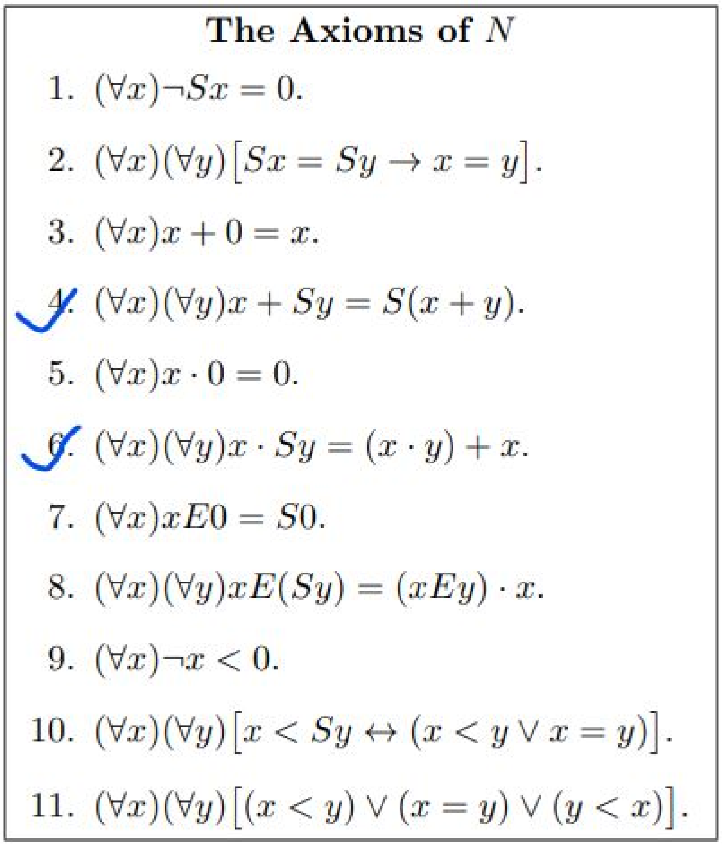
Definition 13. *Assuming x is not free in ψ, we have $\Gamma = \{\phi \to \psi\}$ and $\alpha :\equiv ((\exists x)\phi \to \psi)$ and (Γ, α) is an inference rule of the type (QR)*

Example 3.3.4.

In both definitions, the "property" is described by ψ.

3.4 Non-logical axioms

As mentioned previously, the non-logical axioms do not form a part of the deductive system. Depending on what we want to prove, the set of non-logical axioms changes. To study number theory, the non-logical axioms or the axioms of N, more commonly as Peano's axioms are as follows:



The Axioms of N

1. $(\forall x)\neg Sx = 0.$
2. $(\forall x)(\forall y)\left[Sx = Sy \to x = y\right].$
3. $(\forall x)x + 0 = x.$
4. ✓ $(\forall x)(\forall y)x + Sy = S(x + y).$
5. $(\forall x)x \cdot 0 = 0.$
6. ✓ $(\forall x)(\forall y)x \cdot Sy = (x \cdot y) + x.$
7. $(\forall x)xE0 = S0.$
8. $(\forall x)(\forall y)xE(Sy) = (xEy) \cdot x.$
9. $(\forall x)\neg x < 0.$
10. $(\forall x)(\forall y)\left[x < Sy \leftrightarrow (x < y \lor x = y)\right].$
11. $(\forall x)(\forall y)\left[(x < y) \lor (x = y) \lor (y < x)\right].$

Figure 3.1: From the textbook, pg no: 68

In layman's terms, The axioms of N can be explained as follows:

1. There does exist any number $n \in \mathbb{N}$ such that $S(n) = 0$.
2. The successor function is an injective map. No two numbers have the same successor
3. 0 is the additive identity
4. Given $n.m \in \mathbb{N}$, adding m to n is same as incrementing m by 1, n times or adding 1 to m, n times.
5. Any number multiplied by 0 is equal to 0
6. Given $m, n \in \mathbb{N}$, multiplying n with m is same as adding n with itself m times.
7. Any natural number power 0 is equal to 1.
8. Given $m, n \in \mathbb{N}$, m^n is same as multiplying m with itself n times.
9. 0 is the smallest natural number.
10. Given two numbers $n.m \in \mathbb{N}$, if $n < m$ then either m is the successor of n or there exists another natural number $k > 1$ such that $n + k = k$
11. The law of trichotomy

3.5 The Deduction Theorem

Definition 14. *Consider* Σ *to be a collections of* $\mathcal{L}_{NT}-$*formulas and* θ *is* $\mathcal{L}_{NT}-$ *formula. Then* $\Sigma \cup \theta \vdash \phi$ *if and only if* $\Sigma \vdash (\theta \to \phi)$

Proof: Because the definition is a biconditional statement, we have to prove the following:

1. If $(\Sigma \cup \theta) \vdash \phi$ then $\Sigma \vdash (\theta \to \phi)$
2. If $\Sigma \vdash (\theta \to \phi)$ then $(\Sigma \cup \theta) \vdash \phi$

To prove the first statement let us assume that there is a proof of ϕ from $\Sigma \cup \theta$. We will use induction to show that the above assumption proves $\Sigma \vdash \theta$. Considering the base case to be a one-line deduction of ϕ. By this, we can infer that one of two holds true: either ϕ is a deduction from Σ which also implies that there is a deduction of $\phi \rightarrow (\theta \rightarrow \phi)$ from Σ as $\phi \rightarrow (\theta \rightarrow \phi)$ is a tautology. So, we have $\Sigma \vdash \phi$ and $\Sigma \vdash (\phi \rightarrow (\theta \rightarrow \phi))$. Using modes ponens, we can conclude that there is a proof of $\theta \rightarrow \phi$ from Σ or $\Sigma \vdash (\theta \rightarrow \phi)$. Hence, for the base case, the statement holds true.

The induction hypothesis assumes that there is a deduction of n lines for ϕ from $\Sigma \cup \theta$ which implies that there is proof of $\theta \rightarrow \phi$ from Σ.

To prove: Given that there is a deduction of $n+1$ *lines of* ϕ *from* $\Sigma \cup \theta$*, then this implies that there is a deduction* $\theta \rightarrow \phi$ *from* Σ. So, given the assumption, there must be proof of some formula α of n lines which implies there is a deduction of $\theta \rightarrow \alpha$ from Σ. Now, given $\alpha \rightarrow \phi$, the proof of $\alpha \rightarrow \phi$ is also a n lines deduction from $\Sigma \cup \theta$ which implies there is a deduction of $\theta \rightarrow (\alpha \rightarrow \phi)$ from Σ. So, now we have $\Sigma \vdash (\theta \rightarrow \alpha)$ and $\Sigma \vdash (\theta \rightarrow (\alpha \rightarrow \phi))$. Now, because $(\theta \implies (\alpha \rightarrow \phi)) \rightarrow ((\theta \rightarrow \alpha) \rightarrow (\theta \rightarrow \phi))$ is a tautology, we have a deduction of $(\theta \rightarrow (\alpha \rightarrow \phi)) \rightarrow ((\theta \rightarrow \alpha) \rightarrow (\theta \rightarrow \phi))$ from Σ, that is, $\Sigma \vdash (\theta \rightarrow (\alpha \rightarrow \phi)) \rightarrow ((\theta \rightarrow \alpha) \rightarrow (\theta \rightarrow \phi))$. Using $\Sigma \vdash (\theta \rightarrow (\alpha \rightarrow \phi))$ and $\Sigma \vdash (\theta \rightarrow (\alpha \rightarrow \phi)) \rightarrow ((\theta \rightarrow \alpha) \rightarrow (\theta \rightarrow \phi))$, we can conclude $\Sigma \vdash ((\theta \rightarrow \alpha) \rightarrow (\theta \rightarrow \phi))$ by modes ponens. Using this result and $\Sigma \vdash (\theta \rightarrow \alpha)$, we can conclude $\Sigma \vdash (\theta \rightarrow \phi)$, by modes ponens again, under the inference rule (PC)

Under the inference rule (QR), the similar argument applies but in addition to this, we use the axiom $\alpha \rightarrow \forall x \alpha$ in accordance with the first quantifier rule.

To prove the second statement, let us assume that there is a deduction of $\theta \rightarrow \phi$ from Σ. This means that given θ, we can conclude that ϕ is true. We know that there is a deduction of θ from $\Sigma \cup \theta$. Therefore, by our antecedent, we can conclude that there is a deduction of ϕ from $\Sigma \cup \theta$. Hence, proved. Wikipedia n.d.(a)

Chapter 4

Completeness and Compactness

4.1 Questions on proof and truth

We now know to how to prove a formula, or how to arrive at the desired formula and we clearly understand, given a language what is a true statement in that language and how to check if that statement is true. But, one can ask, "Does proof guarantee truth?" Yes, indeed. This is explained by the **Soundness Theorem**. Or "Does truth imply deductibility?" which forms the content of the **Completeness Theorem**. Soundness and Completeness together assert that our deductive system can prove only and all true statements. We will not look at soundness and completeness more closely.

4.2 Soundness

Definition 15. *If* $\Sigma \models \phi$ *the* $\Sigma \vdash \phi$*, where* Σ *is a collection o* $\mathcal{L}-$ *formulas and* ϕ *is* $\mathcal{L}$ *formula*

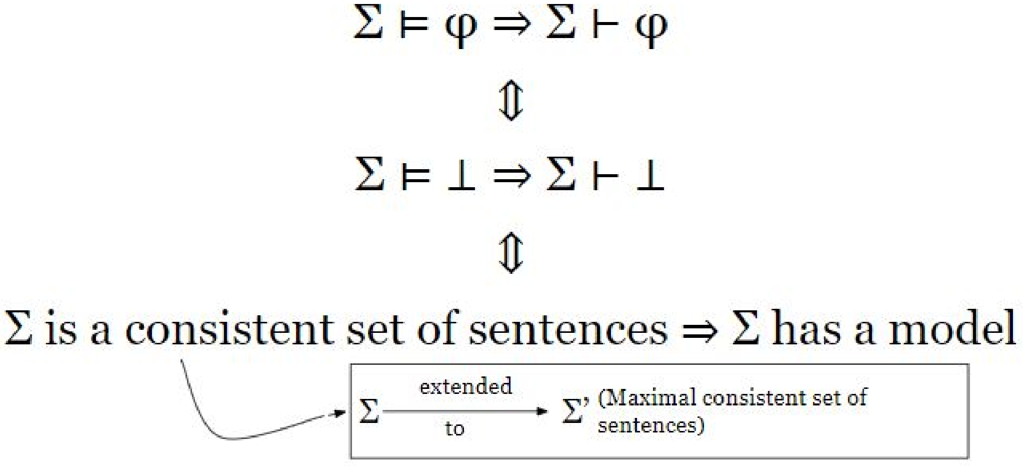
That is, syntactically when a statement can be proven, it must be valid semantically. This proves that our deductive system is sound. We have to show that if a formula is deduced from the set of non-logical axioms, it must belong to the set of all the valid formulas in the given system.

Proof idea: Let $C = \{\phi \,|\Sigma \models \phi\}$ We have to prove that any formula that is deduced, also satisfies Σ that is, we have to show that the set of logical axioms and the set of non-logical axioms are contained in C. We also have to show that set C is closed under each inference rule, that is under each (Γ, ϕ). To show that the set of logical axioms and the set of non-logical axioms satisfy Σ, we will use the fact that logical axioms are valid and $\Sigma \models \sigma$, $\sigma \in \Sigma$ respectively. To prove closure under each (Γ, ϕ), we have to show that given all the formulas in Γ) are true, that is $\Gamma \subset C$, then$\phi \in C$.

4.3 Completeness

Definition 16. *Given- Σ is a set of $\mathcal{L}-$ formula and ϕ is also a $\mathcal{L}-$ formula. If $\Sigma \models \phi$, then $\Sigma \vdash \phi$.*

One way proposed to go about the proof, is as follows:



$$\Sigma \models \varphi \Rightarrow \Sigma \vdash \varphi$$

$$\Updownarrow$$

$$\Sigma \models \bot \Rightarrow \Sigma \vdash \bot$$

$$\Updownarrow$$

Σ is a consistent set of sentences $\Rightarrow$ Σ has a model

$\Sigma \xrightarrow[\text{to}]{\text{extended}} \Sigma'$ (Maximal consistent set of sentences)

Currently, there are sentences in the language whose truth is not in the scope of Σ. That is, more precisely, the truth and falsity of the sentence are consistent with Σ. A consistent deductive system ensures that contradiction is not provable. To explain the equivalent formulations using an analogy with the system of linear equations, when the number of the equations is less than the number of unknowns, we consider the system to be underdetermined, which is exactly the case here. Presently, there are some sentences in the language, irrespective of their truth or falsity, they do not make the deductive system inconsistent. Now, consider the system of linear equations when the number of the equations is the same as the number of unknowns. We consider such a system to be “exact”, in some sense. We will extend the language by adding some evidence until no more can be added, causing the system to be inconsistent catering to those sentences that make the deductive system underdetermined. We will then construct a model $\Sigma^{'}$. Now, because $\Sigma \subset \Sigma^{'}$ and we have a model for $\Sigma^{'}$ all sentences of Σ will also be true in the same model.

4.3.1 Henkin Constants and Henkin axioms

Consider the following situation: $\mathcal{L}$ is a language with no constant symbols and $\phi :\equiv \exists x P(x)$ is an element in Σ. The universe of structure $\mathfrak{N}$, is the collection of all the variable free terms of $\mathcal{L}$. As ϕ is an existential statement, one of the ways to prove ϕ, is to find a c such that $\mathfrak{N} \models \phi(c)$.

We want to prove all the existential statements of the above kind and to be able to do that we must find some constant for which the statement stands true. So we add one such constant for each existential statement that becomes the witness for that sentence to be true. Now, in the practice of doing so, we also need axioms to state that if the above statement is true then it is enough to find c such that $\phi(c)$ is true $\mathfrak{N}$ or $\mathfrak{N} \models \phi(c)$.

Henkin constants are the extensions (or witnesses) added to $\mathcal{L}$, in order to be able to prove all such existential statements, and Henkin Axioms are the sentences that are also added in the extension that forms the proof of the witnesses. But, in the practice of extending the language, there is a chance the new constants and/or axioms produce a proof of contradiction. In which case it would make Σ inconsistent. But, it so happens that extending the language does not affect the consistency of Σ, as given by the two lemmas

So, we add constants and then add axioms to justify this substitution and repeat, repeat, repeat for every such existential $\mathcal{L}_{NT}-$ statement. Add infinitum. That leads us to obtain $\mathcal{L}^{'}$ which is the extension of $\mathcal{L}_{NT}$ and that is also a countable union of the countable sets. We will construct $\Sigma^{'}$ by including only the sentences that are true in the given model, that is, depending on whether σ or $\neg\sigma$ is deducible from $\Sigma^{'}$, we will add σ or $\neg\sigma$ to $\Sigma^{'}$ thereby constructing $\Sigma^{'}$ that is a maximally consistent set of sentences.

4.3.2 Construction of a model for $\Sigma^{'}$

We must first define all the $\mathcal{L}^{'}$ elements precisely before constructing the model for the same. The universe of a model should be a collection of the variable-free terms of the new language. But there is a possibility that we have the sentence $= t_1t_2$ in $\Sigma^{'}$. If terms form the universe of the model, we realize $-t_1t_2$ no longer remains true as the two terms are syntactically different. So, the universe of the model needs to be constructed taking all this into account. One way proposed to solve this problem is to quotient out these terms, that is, express the universe of the model to be as the equivalence classes of the terms that are equal. The meanings of the function symbols and relation symbols can be taken to be similar to our recursive definition of the structure explained in chapter (02), but obviously with respect to the universe defined in the preceding paragraph. It so follows that all the sentences in $\Sigma^{'}$ are indeed true in the above-proposed model. Hence, we can conclude that the same structure satisfies all sentences in Σ.

4.4 Compactness

Definition 17. *Given that Σ is a set of axioms, there is a model of Σ if and only if every finite subset Σ_0 of Σ has a model.*

Proof: Because the above statement is a *if and only if* statement, we have to prove:

1. If Σ has a model then every finite subset Σ_0 of Σ also has a model.
2. If every finite subset Σ_0 of Σ also has model or then Σ has a model.

To prove the first statement consider that Σ has a model. That is, there exists a structure such that every formula in Σ is true in that structure. So, every finite subset Σ_0 of Σ will also be true in that structure. Hence, every finite subset Σ_0 also has a model. Hence, proved.

Now, we know that $\neg q \to \neg p$ is equivalent to $p \to q$. So, to prove the second statement we will prove that if Σ is not satisfiable then every finite subset Σ_0 of Σ is also not satisfiable. So, consider that Σ is not satisfiable. That is, there does not exist any structure such that every formula in Σ is true in that structure. That is, a contradiction is true in Σ. Because our deductive system is complete, we can conclude that there is a deduction of the contradiction from Σ. Let this deduction be $\Sigma_0 \subset \Sigma$. So, by definition, this contradiction can also be deduced from Σ_0. Now because our deductive system is sound, we can conclude that this contradiction is true in Σ_0. Hence, there does not exist any structure that satisfies Σ_0. Hence, proved.

One important corollary, following from the theorem asserts that if Σ has a model then ϕ also has a model is equivalent to saying that this is possible only if there is a finite subset Σ_0 of Σ that has model implying ϕ has a model. The intuition here is that the finite subset Σ_0 of Σ is the finite deduction of ϕ. So,

Corollary 3. *Let Σ be a set of $\mathcal{L}$ formulas and ϕ is a $\mathcal{L}$ formula.* $(\Sigma \models \phi) \Leftrightarrow (\exists \Sigma_0 \subset \Sigma, \Sigma_0 \models \phi)$

4.4.1 Compactness Theorem: Application 1

Proving the existence of the non-standard model of arithmetic.

Proof outline: Consider a language $\mathcal{L}$ that is defined as follows: $\mathcal{L} = \mathcal{L}_{NT} \cup \{c\}$ such that c is a constant that takes the role of a non-standard number. Also, consider Σ^* to be a collection that is defined as follows: $\Sigma \cup \{0 < c, S0 < c, ..., SSSS....S0 < c, ...\}$ Every finite subset of Σ^* has a model and hence there exists a model that satisfies Σ^*. That is, there exists a structure in which all the statements in Σ^* are true. Therefore, there must be an c in the universe that is greater than any number obtained by applying the successor function on 0 finite number of times. Hence, proved. Wikipedia n.d.(b) David n.d.

4.4.2 Compactness Theorem: Application 2

Proving the claim: There does not exist any set of formulas that can describe the family of graphs with infinite diameters. All connected graphs will have a finite diameter, while any disconnected graph will have an infinite diameter.

Proof outline: Assume Σ_a to be the set of formulas that describes the family of graphs with finite diameter and we have Σ_d to be the set of formulas that describes the family of curves with infinite diameter. We will show that every finite subset of $\Sigma_a \cup \Sigma_d$ has a model. That is, take a subset of $\Sigma_a \cup \Sigma_d$ and construct a model for this subset by considering its maximum diameter obtained by taking its intersection with Σ_d. So by applying the compactness theorem, we arrive at a contradiction that $\Sigma_a \cup \Sigma_d$ has a model. David n.d.

Chapter 5

Finally! Gödels's First Incompleteness Theorem

We are now at a stage where we understand the notion of completeness. But, how can completeness and incompleteness co-exist in the given paradigm? What is it that is complete (or incomplete)? Are we considering the completeness and incompleteness of the same entities? We have concluded that the deductive system is sound and complete. So, when we talk about completeness, we are discussing the completeness of the deductive system which constitutes the set of logical axioms and the set of the inference rules. The deductive system does not comprise a set of non-logical axioms. So, when we are concerned with incompleteness, we are actually referring to Σ being incomplete, which is not a part of the deductive system. So, if we go on with the set of non-logical axioms N to axiomatize the set of natural numbers, we run into trouble.

5.1 Overview

As mentioned previously, Gödel's First Incompleteness Theorem statements states that there exists a true statement that cannot be proved using the axioms of N. If this sentence is viewed as a statement in the language of sets, it would be as follows: The set of deductions from N is not equal to the set of true statements in $\mathcal{N}$. Therefore, essentially, checking for the existence of a proof is the same as checking for the belongingness of an element in a set. To simplify the problem even further, we will use coding to associate every deduction and a true statement to a natural number, which is called the gödel number of true statement/deduction. So, now the set of deductions from N and the set of true statements in $\mathcal{N}$ are now subsets of natural numbers. Hence, restating the problem: The set of natural numbers corresponding to deductions is not equal to the set of natural numbers that correspond to true statements and precisely there exists a natural number in the set of true statements(in $\mathcal{N}$ that does not belong to the set of deductions (from N).

The First Incompleteness Theorem provides a systematic method to explicitly construct this statement that is true in $\mathfrak{N}$ but not provable from N. This is done using the self-reference lemma. The self-reference lemma wants you to be able to substitute the gödel number of a formula. But to be able to do this, we need the formula to possess the property of it being representable. What is representability? How to go about coding? The subsequent sections explore each of these ideas broadly.

5.2 Representability and Definability

Definition 18. *Let $A \subset \mathbb{N}^k$. If we can find a n $\mathcal{L}_{NT}$formula $\phi(x)$ where ϕ is of arity k such that:*

$$\forall a \in A, N \vdash \phi(\bar{a})$$

$$\forall b \notin A, N \nvdash \phi(\bar{a})$$

then we can conclude that A is representable and it represented by ϕ.

The main idea behind representability is to confirm that the formula corresponding to this function(which is a relation) is indeed deducible from N when substituted with the different values in the domain. That is, consider:

Example 5.2.1. *Let $A = \{1, 2, 3, 4, 5\} \subset \mathbb{N}$. Intuitively, we can infer that the $\mathcal{L}_{NT}-$ formula that represents A is $\phi :\equiv x > 0$.*

Example 5.2.2. *We have $A = \{(x_1, x_2, x_3)|x_3 = x_1 + x_2\}$. Intuitively, the $\mathcal{L}_{NT}$ formula that* *can potentially represent A would be* $\phi :\equiv= (+x_1x_2)((x_3)$. *If* $N \vdash \phi^{x_i}_{t_i}$, *where* $t_i \in A$ *and this holds true for all elements in A, then we can conclude that A is representable and* ϕ *represents A.*

Definition 19. *Let* $A \subset \mathbb{N}^k$. *If we can find an* $\mathcal{L}_{NT}$*formula* $\phi(x)$ *where* ϕ *is of arity k such that:*

$$\forall a \in A, \mathfrak{N} \models \phi(\bar{a})$$

$$\forall b \notin A, \mathfrak{N} \not\models \phi(\bar{a})$$

then we can conclude that A is definable and ϕ *defines A*

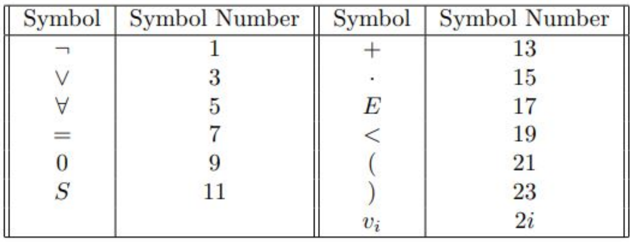
The main idea behind definability is to confirm that the formula corresponding to this function (which is a relation) is indeed true in $\mathfrak{N}$ when substituted with the different values in the domain. That is, consider:

Example 5.2.3. *We have* $f : \mathbb{N}^2 \to \mathbb{N}, f(x_1, x_2) = x_1 + x_2 + 1$. *So, following from our definition,* $A = \{(x_1, x_2, x_3) | x_3 = x_1 + y_2 + 1\}$. *The* $\mathcal{L}_{NT}$ *corresponding to the function is* $\phi :\equiv= (S(+x_1x_2))((x_3)$. *If* $\mathfrak{N} \models \phi^{x_i}_{t_i}$, *where* $t_i \in A$ *and this is applicable all elements in A, then we can conclude that A is definable and* ϕ *defines A.*

5.3 Coding Deductions

Recall: A deduction D is a finite sequence of formulas $(\phi_1, \phi_2, ..., \phi_n)$. So, first, we need to code each formula in the deduction sequence. This transforms D to $(\ulcorner \phi_1 \urcorner, ..., \ulcorner \phi_n \urcorner)$. The final step is to code this sequence of gödel numbers.

Recall: A $\mathcal{L}_{NT}$ formula is a string of $\mathcal{L}_{NT}$ terms and symbols. So, we first fix the code for the symbols.



Symbol	Symbol Number	Symbol	Symbol Number
$\neg$	1	$+$	13
$\vee$	3	$\cdot$	15
$\forall$	5	E	17
$=$	7	$<$	19
0	9	(	21
S	11	)	23
		v_i	$2i$

Using this table, we will obtain a sequence of n numbers from the above, given the formula contains n symbols.

Given a sequence of length k: $(n_1, n_2, n_3, ..., n_k)$, one way to code this - $\prod_{i=1}^{k} p_i^{n_i+1}$, where p_i or $p(i)$ is a function that maps i^{th} natural number to i^{th} prime number. But, what is so special about the proposed method of coding symbols? The answer to this - it allows a clear separation between a set of gödel numbers and its complement with the set of natural numbers. From the method proposed, we can see that not all natural numbers will be gödel numbers (the converse, however, is true). Consider the number $100 = 2^2 \cdot 5^2$. Notice that this number cannot be expressed as a product of k_i powers the first n prime numbers. However, this can also be written as follows: $100 = 2^2 \cdot 3^0 \cdot 5^2$ which decodes to the sequence $(2, 0, 2)$. But the above can also be written as $100 =^2 \cdot 3^0 \cdot 5^2 \cdot 7^0$ which decodes to sequence $(2, 0, 2, 0)$. So, to avoid this irregularity, we follow the method proposed above.

Example 5.3.1. *Consider a* $\mathcal{L}_{NT}$ ϕ *is* $\phi :\equiv= E(SS0)(+S0S0)$*. So, according to what we have fixed,* ϕ *corresponds to* $(7, 17, 21, 11, 11, 9, 23, 21, 13, 11, 9, 11, 9, 23)$*. Using this sequence,* $\ulcorner \phi \urcorner = 2^8 \cdot 3^{18} \cdot 5^{22} \cdot 7^{12} \cdot 11^{12} \cdot 13^{10} \cdot 17^{24} \cdot 19^{22} \cdot 23^{24} \cdot 29^{12} \cdot 31^{10} \cdot 41^{12} \cdot 43^{10} \cdot 47^{24}$ *which, with no further guesses, we will let the machines compute!*

We now have sequence of gödel numbers: $(\ulcorner \phi_1 \urcorner, ..., \ulcorner \phi_k \urcorner)$. Coding this sequence completes our task of coding deductions. One way proposed is the following:

$$\ulcorner D \urcorner = \prod_{i=1}^{k} p_{i+2}^{\ulcorner \phi_k \urcorner}$$

Example 5.3.2. *Suppose the sequence of the gödel numbers corresponding to the deductions is as follows:* $(134, 12, 421, 903, 86, 32)$*. The gödel number corresponding to this deduction is given by* $5^{134} \cdot 7^{12} \cdot 11^{421} \cdot 13^{903} \cdot 17^{86} \cdot 19^{32}$*. Not surprising if the number computed fills the rest of the page or probably more!*

5.4 Formulas with/without bounded quantifiers

5.4.1 Bounded quantifiers

Definition 20. *Let ϕ be a $\mathcal{L}$ formula and t be a $\mathcal{L}-$ term. x is a variable that is not in t. The set of bounded quantifiers is as follows:*

- $(\forall x < t)(\phi) \Leftrightarrow (\forall x(x < t \to \phi))$
- $(\forall x \leq t)(\phi) \Leftrightarrow (\forall x((x < t \vee x = t) \to \phi))$
- $(\exists x < t)(\phi) \Leftrightarrow (\exists x(x < t \to \phi))$
- $(\exists x \leq t)(\phi) \Leftrightarrow (\exists x((x < t \vee x = t) \to \phi))$

We classify $\Sigma-$formulas to those $\mathcal{L}_{NT}-$formulas with unbounded existential quantifiers while $\Pi-$formulas to be the one with unbounded universal quantifiers. The intersection of the two sets gives rise to a set of $\mathcal{L}_{NT}$ formulas with bounded existential and universal quantifiers and call these $\Delta-$formulas.

5.5 The Self-Reference Lemma

Definition 21. *Let $\psi(v_1)$ be an $\mathcal{L}_{NT}$ formula with only (v_1) free. Then there is a sentence ϕ such that $N \vdash [\phi \leftrightarrow \psi(\phi)]$.*

In layman's terms, the sentence says that some property, precisely ψ is applicable to itself.

Proof Outline: We construct such a ϕ and verify that it indeed satisfies the lemma. Precisely,

$\phi :\equiv \forall y \forall z[[Numf(\ulcorner\gamma(v_1)\urcorner, y) \wedge Subf(\ulcorner\gamma(v_1)\urcorner, \ulcorner v_1\urcorner, y, z)] \to \psi(z)]$, where

$\gamma(v_1) = \forall y \forall z[[Numf(v_1, y) \wedge Subf(v_1, \ulcorner v_1\urcorner, y, z) \to \psi(z)]$. Here, *Numf* and *Subf* represents the relation corresponding to $y = \overline{\mathrm{Num}(a)} = \ulcorner\overline{a}\urcorner$ and $y = \overline{\mathrm{Sub}(\phi, x, t)} = \ulcorner\overline{\phi^x_t}\urcorner$ in addition to which, $N \vdash [Numf(a, y) \leftrightarrow y = \overline{\mathrm{Num}(a)}]$ and $N \vdash [Subf(\phi, x, t, y) \leftrightarrow y = \overline{\mathrm{Sub}(\phi, x, t)}]$ All these details of $Numf$ and $Subf$ follow directly from the lemma.

5.6 Gödel's First Incompleteness Theorem

Definition 22. *Consider* A *to be a consistent and recursive set of axioms in the language* $\mathcal{L}_{NT}$*. Then there is a sentence* ϕ *such that* $\mathfrak{N} \models \phi$ *but* $A \nvdash \phi$

Before escalating to the proof of Gödel's First Incompleteness Theorem, there are theorems and definitions which we would take it to be true and use it freely in order to arrive at the conclusion. A detailed proof of each of these theorems/definitions have provided in the book I have used to study this subject. The definitions/theorems are as follows:

1. All $\Sigma-$ formulas are definable.
2. If $\phi(x)$ is $\Sigma-$ formulas and $\mathfrak{N} \models \phi(t)$, then $N \vdash \phi(t)$, where t is substituted in place of x.
3. $THM_A = \{\ulcorner \phi \urcorner \mid A \vdash \phi\}$. We have Thm_A defines THM_A

Proof: We will use the self-reference lemma to construct the statement that concludes of not having a deduction of itself. That is, $N \vdash [\phi \leftrightarrow \neg Thm_A(\ulcorner \bar{\phi} \urcorner)]$. Because $Thm_A(v_1)$ is a $\Sigma-$ formula with v_1 as the only free variable, by (1), we can conclude the following: $\mathfrak{N} \models [\phi \leftrightarrow \neg Thm_A(\phi)]$. Using the definition of truth, the above statement can be written as follows: $\mathfrak{N} \models \phi \iff \mathfrak{N} \not\models Thm_A(\phi)$. Only considering the latter, $\mathfrak{N} \not\models Thm_A(\phi) \implies \ulcorner \phi \urcorner \notin THM_A$, which is the definable set. From the consequent, we can infer $A \nvdash \phi$. So, we have concluded: $\mathfrak{N} \models \phi \iff A \nvdash \phi$.Consider the scenario when $\mathfrak{N} \not\models \phi$ and $A \vdash \phi$. Therefore, by definition,

$\Rightarrow \ulcorner \phi \urcorner \in THM_A$

$\Rightarrow \mathfrak{N} \models Thm_A(\ulcorner \phi \urcorner)$ (As THM_A is defined by Thm_A)

$\Rightarrow N \vdash Thm_A(\ulcorner \phi \urcorner)$ (By the application of (2))

$\Rightarrow N \vdash \neg\phi$

$\Rightarrow A \vdash \neg\phi \Rightarrow\Leftarrow$

Therefore, $\mathfrak{N} \models \phi$ and $A \nvdash \phi$ is the only possibility that makes the biconditional statement true. Hence, proved.

Bibliography

Leary, Christopher C and Lars Kristiansen (2015). *A friendly introduction to mathematical logic.* Lulu. com (cit. on p. 5).

Wikipedia (n.d.[a]). *Deduction theorem.* `https://en.wikipedia.org/wiki/Deduction_theorem` (cit. on p. 25).

— (n.d.[b]). *Existence of non-standard models of arrithmetic.* `https://proofwiki.org/wiki/Existence_of_Non-Standard_Models_of_Arithmetic` (cit. on p. 30).

David, Shai-Ben (n.d.). *Logic and Computation.* `https://www.youtube.com/playlist?list=PLPW2keNyw-utXOOzLR-Wp1p0eE5LEtv3N` (cit. on pp. 30, 31).