Gödel's Proof

2. Semantics of First Order Theory and Hilbert's Program

S M Srivastava

Basic aspects of model of a theory are reviewed. The well known Hilbert's program, that sought to set the agenda for 20th century mathematics, is presented.

S M Srivastava is with the Indian Statistical Institute, Kolkata. His current interests lie in descriptive set theory, mathematical logic and topology. He loves teaching mathematics.

1. Model of a Theory

In Part 1 (Section 2)[1], we presented the syntactical notions pertaining to first order theories. However, in actual practice mathematical theories are not developed syntactically. The syntactical description of a theory is needed to explain and examine the logical aspect of a theory.

In this section we give relevant definitions to connect the syntactical description of a theory with the setting in which a mathematical theory is developed in practice. Note that instead of defining group theory, in practice, one defines a group as a non-empty set G with a specified element e and a binary operation $\cdot : G \times G \to G$ satisfying the following three conditions:

1. For every a, b, c in G,
$$a \cdot (b \cdot c) = (a \cdot b) \cdot c.$$
2. For every $a \in G$,
$$a \cdot e = e \cdot a = a.$$
3. For every $a \in G$, there is a $b \in G$ such that
$$a \cdot b = b \cdot a = e.$$

Thus a group consists of a non-empty set G with 'interpretations' of the non-logical symbols $\cdot$ (a binary function symbol) and e (a constant symbol) such that all

[1] Part 1. An Introduction to Mathematical Logic, *Resonance*, Vol.12, No.2, pp.59–70, 2007.

Keywords
Structure or interpretation of a language, model, truth, tautology, Hilbert's program, logical axioms, proof, rules of inference, completeness, consistency, undecidable, conservative extension, interpretation.

the axioms of the group theory are 'satisfied'. Further a statement in group theory is called a theorem, if it is satisfied in all the groups. Now it should be obvious that, to give the connection we are looking for, first we should define the interpretation or a structure of a language L as a non-empty set A together with all the interpretation of all the non-logical symbols of L. Then, models of a theory T are those structures of the language of T in which all non-logical axioms are 'true'. We formally present these concepts now.

The year 2006 marked the birth centenary of the great mathematician-logician Kurt Gödel. Gödel's incompleteness theorem has far reaching implications concerning the foundations of mathematics. This article by S M Srivastava can be viewed upon as a prelude for an understanding/appreciation of Gödel's incompleteness theorem. The article is in three parts. The first part will be accessible to anyone with mathematical preparation at the level of II Year BSc/BTech/ BE. The second and the third parts presuppose a certain level of mathematical maturity; students at the master's level may find them useful.

The reader is also urged to look at the article 'Gödel's Explorations in Terra Incognita' by Vijay Chandru in the July 2001 issue of *Resonance*, featuring Gödel. One may also see the April 2006 issue of *Notices of the American Mathematical Society* devoted to Gödel.

Editors

Definition. A *structure* or an *interpretation* for L consists of (i) a non-empty set A (called the *universe* of the structure), (ii) for each constant symbol c of L, a fixed element $c_A \in A$, (iii) for each n-ary function symbol f of L, a n-ary map $f_A : A^n \to A$ and (iv) for each n-ary relation symbol p of L, a n-ary relation $p_A \subset A^n$ on A. The interpretation of $=$ is the equality in A.

For instance, any group is a structure for the language of group theory; the usual real number system with the usual 0, 1, $+$, $\cdot$ and $<$ is a structure for the language of ordered fields. Note that which statement of L is true in the structure and which is not, is not relevant in the notion of a structure of L.

We now define when is a formula of L true or when is it false in a structure of L. Suppose we have a structure of L with universe A and we would like to know whether there is an element $a \in A$ satisfying some property expressed by a formula $\varphi[x]$ for $x = a$. We are in a bit of a problem because φ is a syntactical object and elements of G are not. To circumvent this, given a structure of L with universe A, we first describe an extension L_A of the language L .

Given L and and a structure of L with universe A, let L_A be the first order language obtained from L by adding a new constant symbol i_a for each $a \in A$. The symbol i_a

is called the *name* of a. We regard the same structure as an interpretation of the language L_A by setting the interpretation of i_a to be a, $a \in A$.

First we give the interpretation t_A of a variable-free term t of L_A in this by induction on the rank of t. Note that all such t's can be obtained starting from constant symbols and iterating function symbols on them. The interpretation of a constant symbol c is already given by the structure, namely c_A. If $t_1, \cdots, t_n$ are variable free terms whose interpretations have been defined and if f is a n-ary function symbol of L_A, then we define

$$(ft_1 \cdots t_n)_A = f_A((t_1)_A, \cdots, (t_n)_A).$$

The definition of the interpretation of each variable-free term of L_A in the universe is complete by induction on the rank of terms.

Example. Let L be the language for ring theory with 1. For each positive integer m, let $\underline{m}$ denote the term by 'adding' 1 to itself m-times. Let $P(x)$ be a polynomial expression whose coefficients are of the form $\underline{m}$. Let R be a ring with identity 1 and $a \in R$. Then the interpretation of $\underline{m}$ in R is the element $m \in R$ obtained by adding the multiplicative identity 1 to itself m-times and the interpretation of $P_x[i_a]$ in R is the element $P(a)$ of R.

We first define the notion of truth (of a sentence) in a structure of L. The definition is based on the well-known intended meanings of logical connectives $\vee$ and $\neg$ and that of the logical quantifier $\exists$. We fix a language L and a structure of L with universe A. Recall that formulae have been defined inductively starting from atomic formulae and iterating $\neg$, $\vee$ and $\exists$ on them. We define the truth of a sentence of L_A also by such an inductive method.

A sentence will either be true or will be false in the structure.

The notion of truth of a sentence is based on the well-known intended meanings of logical connectives.

Tautologies are true in all structures of the language.

A variable free atomic formula is of the form $pt_1 \cdots t_n$ where p is a n-ary relation symbol (including $=$) and $t_1, \cdots, t_n$ variable free terms. We say that $pt_1 \cdots t_n$ is true in the structure if

$$p_A((t_1)_A, \cdots, (t_n)_A)$$

holds, i.e.,

$$((t_1)_A, \cdots, (t_n)_A) \in p_A.$$

A sentence $\neg A$ is true if and only if A is false. A sentence $A \vee B$ is true if either A is true or B is true. Finally, a sentence $\exists v A$ is true if $A_v[i_a]$ is true for some $a \in A$. We say that a formula A is true in the structure if its closure is true in the structure.

One can easily check the following: A sentence $A \wedge B$ is true in a structure if and only if both A and B are true in the structure; a sentence of the form $\forall v \varphi[v]$ is true in a structure with universe A if and only if for every $a \in A$, the sentence $\varphi_v[i_a]$ of L_A, is true in the structure.

Note that some formulae, e.g., $v = v$, (v a variable), $A \vee \neg A$, (A a formula), etc. are true in all structures of the language. Such formulae are called *tautologies*.

A *model* of a first order theory T is a structure for L with universe M in which all non-logical axioms of T are true. We usually denote such a model by M itself. For instance, any group is a model of group theory. On the other hand, the set N of natural numbers together with usual 0 and + as the interpretation for e and $\cdot$ respectively is definitely a structure for the language of group theory but not a model of group theory.

In mathematics parlance, a valid sentence, depending on its importance, is called a 'Lemma', a 'Proposition', or a 'Theorem'.

A formula A of T that is true in all models of T is called *valid* in T. One writes $T \models A$, if A is valid in T. In mathematics parlance, a valid sentence, depending on its importance, is called a 'Lemma', a 'Proposition', or a 'Theorem'.

Notice that the notion of theorem given above is somewhat abstract: In order to decide whether a statement is valid in T or not, one has to show that it is true in all models of T. This is where the famous program of the great German mathematician *David Hilbert* enters. Hilbert proposed to write down a set of axioms and a set of rules of inference to infer a sentence from its syntactical structure, and call a sentence a theorem if one can infer it from axioms by a finite sequence of rules of inference.

Hilbert proposed to write down a set of axioms and a set of rules of inference to infer a sentence from its syntactical structure, and call a sentence a theorem if one can infer it from axioms by a finite sequence of rules of inference.

Hilbert's approach obviously has strong potential. First it makes the notion of a theorem finitary in contrast to the abstract notion of validity. Second, the question of whether a sentence is a theorem or not, now becomes amenable to mechanical verification. Gödel carried out this program and proved his famous completeness theorem.

2. The Definition of Proof

In the previous section, we defined the concept of Theorem of a theory: *all those sentences that are valid in the theory, i.e., that are true in all the models of the theory.* Is there an equivalent syntactical definition of a Theorem? In this section, we shall address this question.

Since all tautologies are theorems, to give a finitary notion of proof as proposed by Hilbert, in addition to the non-logical axioms, we should also have a set of tautologies, to be called 'logical axioms', in our initial premises. We then have to fix a set of 'rules of inference' which are logically correct. What does it mean? In mathematical practice, assuming a certain statement A true in a structure, if we show that B is true in that structure, we say that B is a *logical consequence* of A. In other words, whenever $A \Rightarrow B$ is a tautology, we infer B from A. Thus inferring B from A is logically correct, if $A \Rightarrow B$ is a tautology. Similarly, inferring a sentence

B from sentences $A_1, \cdots A_n$ is logically correct provided $A_1 \Rightarrow \cdots \Rightarrow A_n \Rightarrow B$ is a tautology. For instance, we infer $B \vee C$ from $A \vee B$ and $\neg A \vee C$.

To give a satisfactory syntactical definition of a Theorem, the following question arises: Is there a 'nice' set of tautologies (to be called *logical axioms*) and a 'nice' set of rules of inferences so that a formula is true in all models if and only if it can be inferred from logical and non-logical axioms by iterating the rules of inference in our list of rules of inference?

Thus to give a satisfactory syntactical definition of a Theorem, the following question arises: Is there a 'nice' set of tautologies (to be called *logical axioms*) and a 'nice' set of rules of inferences so that a formula is true in all models if and only if it can be inferred from logical and non-logical axioms by iterating the rules of inference in our list of rules of inference? By 'nice' here we mean the following: It should be mechanically decidable whether a formula is in our list of logical axioms or not. Similarly, it should be mechanically decidable whether a rule of inference is in our list or not. Based on this aim, we now proceed to give a definition of a Proof.

Here is a set of formulae of L, called *logical axioms.*

1) Propositional Axiom. These are formulae of the form $\neg A \vee A$.

2) Substitution Axiom. These are formulae of the form
$$A_v[t] \Rightarrow \exists v A.$$

3) Identity Axiom. These are formulae of the form $v = v$.

4) Equality Axiom. These are formulae of the form
$$v_1 = v_1' \Rightarrow \cdots \Rightarrow v_n = v_n' \Rightarrow f v_1 \cdots v_n = f v_1' \cdots v_n'$$
and
$$v_1 = v_1' \Rightarrow \cdots \Rightarrow v_n = v_n' \Rightarrow p v_1 \cdots v_n \Rightarrow p v_1' \cdots v_n',$$
where f is an n-ary function symbol and p an n-ary relation symbol of L.

Note that every logical axiom of L is a tautology.

We now introduce five *rules of inference*.

1) Expansion Rule. Infer $B \vee A$ from A.
2) Contraction Rule. Infer A from $A \vee A$.
3) Associative Rule. Infer $(A \vee B) \vee C$ from $A \vee (B \vee C)$.
4) Cut Rule. Infer $B \vee C$ from $A \vee B$ and $\neg A \vee C$.
5) $\exists$-Introduction Rule. If v is not free in B, infer $\exists v A \Rightarrow B$ from $A \Rightarrow B$.

Note that the conclusion of a rule of inference is true in a structure if its hypotheses are true in the structure.

Here is a definition of Proof as envisaged by Hilbert.

Definition. A *proof* in a first order theory T is a finite sequence $A_1, \cdots, A_n$ of formulae such that each A_i is either a logical axiom or a non-logical axiom of T or it can be inferred from $A_1, \cdots, A_{i-1}$ by one of the rules of inference. It is also called a *proof* of A_n.

Definition. A formula A is called a *theorem* of T if it has a proof in T. In this case one writes $T \vdash A$.

The following theorem is easily proved by induction on the length of a proof.

Validity Theorem. *Every theorem of T is valid in T.*

The following remarkable theorem is essentially due to Gödel.

Completeness Theorem. *A formula A of a theory T is a theorem of T if and only if it is valid in T.*

There is an important equivalent formulation of the completeness theorem. Call a theory T *inconsistent* if there

A theory T is inconsistent if and only if every formula of T is a theorem of T.

is a formula A of T, such that both A and $\neg A$ are theorems of T. We can prove:

Theorem. *A theory T is inconsistent if and only if every formula of T is a theorem of T.*

A theory T is called *consistent* if it is not inconsistent.

Completeness Theorem. *A theory T is consistent if and only if T has a model.*

Here is another important notion. A formula A of T is said to be *decidable* in T if either A is a theorem of T or $\neg A$ is a theorem of T, in symbols, $T \vdash A$ or $T \vdash \neg A$. Otherwise, A is *undecidable* in T. Thus, A is undecidable in T if and only if neither A nor $\neg A$ is a theorem of T. It can be shown that a <u>sentence</u> A of T is undecidable in T if and only if both the theories $T[A]$ and $T[\neg A]$ are consistent (or equivalently, both of them have models), where for any formula B, the theory $T[B]$ is obtained from T by adding B as a new non-logical axiom.

A sentence A of T is called *independent* of the non-logical axioms of T if it is undecidable in T, i.e., both $T[A]$ and $T[\neg A]$ are consistent. By completeness theorem, *a sentence A is independent of the non-logical axioms of T if and only if both $T[A]$ and $T[\neg A]$ have a model.*

A theory T is called *complete* if it is consistent and if every closed formula A is decidable in T. Otherwise, the theory is called *incomplete*. Note that the completeness theorem is not about a theory being complete but about a theory being consistent.

A is *undecidable* in T. Thus, A is undecidable in T if and only if neither A nor $\neg A$ is a theorem of T.

3. Extension by Definition of a Theory

A first order theory T' is called an *extension* of another theory T if every non-logical symbol of the language of T is a symbol of T' of the same type and if every non-logical axiom of T is a theorem of T'. It is easy to check that if T' is an extension of T, then every theorem of T

is a theorem of T'. Indeed, if $A_1, A_2, \ldots, A_n$ is a proof in T, then replace each axiom of T in this sequence by a proof in T'. The sequence thus obtained is a proof of A_n in T'. It follows that if T' is consistent, so is T.

An extension T' of T is called a *conservative extension* if every formula of T that is a theorem of T' is also a theorem of T. We have

If T' is a conservative extension of T, then T' is consistent if and only if T is consistent.

To see this assume that T' is inconsistent. Then every formula of T' is a theorem of T'. In particular, every formula of T is a theorem of T'. Since T' is a conservative extension of T, this implies that every formula of T is a theorem of T, i.e., T is inconsistent.

Next we need the concept of *extension by definition* of a theory T. Let A be a formula of T in which no variable other than $v_1, \ldots, v_n$, $n \geq 1$, (v_i's distinct), is free. We form an extension T' of T by adding a new n-ary relation symbol p and adding a new non-logical axiom

$$pv_1 \ldots v_n \Leftrightarrow A.$$

For instance, in the language of set theory, if we add a binary relation symbol $\subset$ and a new axiom

$$x \subset y \Leftrightarrow \forall z(z \in x \Rightarrow z \in y),$$

we get an extension by definition of ZF in which 'subset' is a defined concept. If T' is defined from T as above, then T' is a conservative extension of T.

Let $v_0, \ldots, v_{n-1}$ and y, y' be distinct variables and A a formula of T in which no variable other than $v_0, \ldots, v_{n-1}$ and y is free. Further, assume that

$$T \vdash \exists y A, \tag{1}$$

and

$$T \vdash (A \wedge A_y[y']) \Rightarrow y = y'. \tag{2}$$

Informally speaking, conditions (1) and (2) say that for all $v_0, \ldots, v_{n-1}$, there is a unique y 'satisfying' A. We form T' from T by adding a new n-ary function symbol f and a new non-logical axiom

$$y = f v_0 \ldots v_{n-1} \Leftrightarrow A.$$

For instance, in set theory, (after defining $\subset$), consider the following formula $A[x, y]$:

$$\forall z(z \in x \Rightarrow z \subset y).$$

Using power set, comprehension and extensionality axioms, one shows that the formula A satisfies (1) and (2) for $T = ZF$. So, one adds a new unary function symbol $\mathcal{P}$ (traditionally called *power set*) and a new non-logical axiom

$$y = \mathcal{P}(x) \Leftrightarrow A.$$

Note that if $n = 0$, this method adds a new constant symbol to T. As an example, we can add a constant symbol 0 (called *emptyset*) in an extension by definition of ZF. This can be seen as follows: Let A be the formula

$$\forall x \neg(x \in y).$$

Using set existence, extensionality and comprehension axioms of ZF, one can show that A satisfies conditions (1) and (2) for $T = ZF$. One then defines an extension by definition of ZF by adding a new constant symbol 0 and a new axiom

$$y = 0 \Leftrightarrow A.$$

Again one can show that if T' is an extension by definition of T by adding a new function symbol, then T' is a conservative extension of T. We say that T' is an *extension by definition* of T if T' is obtained from T by a finite number of extensions of the two types which we have described.

Theorem. *If T' is an extension by definition of T, then T' is a conservative extension of T. In particular, T is consistent if and only if T' is consistent.*

4. Interpretations in a Theory

We now define the notion of an *interpretation of a theory* T in another theory T'. Informally, this would mean that T has a model in T'. For instance, starting from Peano axioms, we construct rational numbers and show that they form a field. Thus we can say that field theory has a model in Peano arithmetic.

We can say that field theory has a model in Peano arithmetic.

Let L an L' be first order languages. An *interpretation* I of L in L' consists of:

• a unary predicate symbol U_I of L', called the *universe* of L;

• for each n-ary function symbol f of L, an n-ary function symbol f_I of L';

• for each n-ary relation symbol p of L other than $=$, a n-ary predicate symbol p_I of L'.

An *interpretation* of L in a theory T' is an interpretation I of L in $L(T')$ such that

$$T' \vdash \exists x U_I x$$

and

$$T' \vdash U_I x_1 \to \cdots \to U_I x_n \to U_I f_I x_1 \cdots x_n$$

for each n-ary function symbol f in L. The first condition requires that T' proves that the universe U_I is non-empty; the second requires that in the theory T', f_I is a n-ary function whose restriction to U_I takes values in U_I. An interpretation I of L in T' may be thought of as a structure of L in T' where the underlying universe is U_I.

Let I be an *interpretation* of L in $L' = L(T')$ and A a formula of L. We define a formula A^I of L' such that $T' \vdash A^I$ will mean that A is 'true' in the interpretation I.

Let A_I be the formula of T' obtained from A by replacing each non-logical symbol u occurring in A by u_I and also replacing each subformula of A of the type $\exists x B$ by $\exists x(U_I x \wedge B_I)$. More precisely, we define A_I by induction on the rank of A. For atomic formulae, we obtain A_I from A by replacing each non-logical symbol u occurring in A by u_I. If A is $\neg B$ or $B \vee C$, A_I is $\neg B_I$ or $B_I \vee C_I$ respectively. If A is $\exists x B$, A_I is $\exists x(U_I x \wedge B_I)$.

Finally, if $v_0, \cdots, v_{n-1}$ are all the variables free in A (and hence in A_I) in alphabetical order, then A^I is the formula

$$U_I v_0 \rightarrow \cdots \rightarrow U_I v_{n-1} \rightarrow A_I.$$

An *interpretation* of a theory T in a theory T' is an interpretation I of $L(T)$ in the language of T' such that for every non-logical axiom A of T, $T' \vdash A^I$, i. e., A^I is a theorem of T'. In this case, we may think of U_I as a model of T in T'.

We have the following results.

Theorem. *If T has an interpretation in a consistent theory, then T is consistent.*

Theorem. *The Peano arithmetic has an interpretation in an extension by definition of Set Theory.*

Indeed, most interesting theories have interpretations in extensions by definition of set theory. In this sense, set theory is a theory on which most mathematical theories rest. Also, if set theory is consistent, so are most mathematical theories.

We have now presented an overview of the first order logic. In the third and last part of this article, we present Gödel's incompleteness theorem. We shall first present three very important concepts: Recursive Functions, Gödel numbers and Representability.

Suggested Reading

[1] Martin Davis, Elaine Weyuker and Ron Segal, *Computability, Complexity and Languages: Fundamentals of Computer Science*, Academic Press, 1994.

[2] Douglas R Hofstadter, *Gödel, Escher, Bach: An Eternal Golden Braid*, Vintage Books, 1989.

[3] K Kunnen, *Set Theory – An Introduction to Independence Proofs*, North-Holland, 1980.

[4] J Shoenfield, *Mathematical Logic*, Addison-Wesley, 1967.

Address for Correspondence
S M Srivastava
Stat-Math Unit
Indian Statistical Institute
203, BT Road
Kolkata 700 108, India.
Email:smohan@isical.ac.in