Gabriele Lolli

COMPLETENESS

Being an Account of the Vain Search for Meaning in Logic

TABLE OF CONTENTS

Final conclusion, where it is explained that hot water is better than cold water

Reopening of the case, with a supplementary investigation on the incompleteness of certain logics, in particular of the notorious second order one

The strange case of Dr. Skolem and Mr. Gödel

Introduction, whose (intended) effect is likely to be that of bewildering and confusing the reader, according to a Socratic maxim: first realise that you don't know, contrary to the lazily received wisdom, then proceed to untangle the wool

In a logic course the completeness theorem is a point of no return; you cross it, as Caesar crossed the Rubicon, and you cannot go back to the lost innocence. But the pristine innocence was made up of a lot of sinful biases and misconceptions. After the theorem, you cannot cheat any more; logic changes the way it is used and conceived (hence its status).

The completeness theorem is the first, sometimes the only theorem one meets in an introductory course - so it must be important . Before it, there are a lot of proofs, but no theorem (some say there is a sequence of lemmata, paving the way for the theorem). The proofs are very tiresomely detailed inductive proofs of properties of the languages and of the interpretations, which parallel the inductive definitions of the syntactic and semantic notions (assignments, satisfaction and so on); for example the property that A[x/y][y/x] is A if y is free for x in A and it is not free in A; or the property that if σ satisfies A and σ' is such that σ'(y) = σ(x), again with a lot of strange conditions on variables, then σ' satisfies A[x/y], and so on and so forth. The completeness theorem on the contrary is a real theorem, though ironically sometimes it is not proved at all, especially in computer science or philosophy classes (when in math classes the maximal ideal theorem is used, no mention is done of their equivalence). What is the use of a proof after all? Part of the innocence is that logic is a technique. In a logic course you spend your time improving your natural abilities by mastering new techniques, through formalisation and derivation exercises, but you do not prove theorems - that is a mathematicians' affair (but for those painstaking and uninformative proofs of the lemmata, of course). Perhaps proving the theorem might lead you astray with mathematical spells (or might really teach you something).

The completeness theorem opens the Pandora's box of the distinction between syntax and semantics; but Pandora's box is full of riches as well as of sorrows. Our previous experience with natural language or mathematics is of no use to understand the theorem, because such distinction is not to be found in the use or in the study of any language. When learning a language beginners certainly do grammatical analysis, which is useful from many points of view, for example to disambiguate ambiguous sentences, or even to grasp the meaning conveyed by a sentence. If a sentence is not grammatically correct it has no sense, though it can sometimes be understood from the context, or thanks to extra linguistic circumstances; the parsing of the grammatical form is essential to understanding, as far as the sense is compositional, in its simplest and most intuitive sense - that is corresponding to the syntactic construction. However in the study of a language one does only grammatical analysis, without meeting an independent semantics, even if one learns to grasp, through correct use, the sense and meaning of discourse.

In the first part of the logic course, students learn the syntax of predicative (first-order) languages, which is very much akin to grammar, as well as the rules of correct speech: alphabets, formation rules for terms and formulae, deduction rules and their chaining in inferences (derivations). Formulae are viewed as painstaking and absolutely rigorous regimentation of phrases. The aim is that of learning the action of the formation rules which give origin to *schemata*. In this way one can recognise all phrases with the same form, that is constructed with the same sequence of applications of the grammatical formation rules.

The use of (variable) letters to express form, which goes back at least to Aristotle, is on a par with the primitive use of variables to be found at the beginning of modern arithmetic, when people began to use n for a generic number, in order to write down formulae or relations valid for all numbers. Sometimes people used instead a particular (albeit small) number, stating at the same time (without proof) that calculations and reasoning had however a general validity. If true, the effect was the same that with the generic n, the more so since the use of n wasn't accompanied by a real general, say inductive, proof.

History repeats itself with logical formulae. The use of letters in itself does not mark a novelty with respect to a rigorous, albeit informal study of language. The problem is to explain the why and wherefore of the quest for generality. The explanations are often off the mark and perversely twisted. The answer usually found in the introductions to logic textbooks is that a logical reasoning valid in a domain should be valid also in other domains. Aristotle had already remarked that some inferences are valid by virtue of form alone, hence if valid in a domain they are valid also in others. But it is not often that one explicitly notes that a certain kind of inference can be found also in other circumstances; this happens only with very special arguments, e.g. proofs by contradiction (which deservedly have a widely known name). Here, however, the reference to the form of inferences is a loose one, only the very general strategy and structure of the proof is actually pointed out. What happens in truth with logical inferences is that instead of exporting generality one imports it; the appeal to the formal validity serves the purpose of justifying a particular inference in terms of general formal laws. Reasoning becomes the application of general laws to particular cases.

The whole machinery of formal logic thus acquires an ambiguous status; when in action, it represents with its pedantic proceedings the famous chains which Poincaré contrasted to the wings of thought; as a theory, it carries along a debatable thesis on the nature of reasoning, namely that correct reasoning in each particular domain is just the application of universally valid (possibly innate, mental) rules.

Those who study reasoning - knowledge engineers, psychologists - maintain on the contrary that domain specific inferences, those where the speaker knows what he is talking about, are not logical inferences; though the reasoner may say: A, if A then B, hence B, actually what is at work is some kind of *intuition* (lat. *intuere*), or of *pattern-matching*, appropriate to the domain; one really means something like: "if you see A then you see (you can see, you should see) B", or "pattern A goes into pattern B" or the like.

Elementary mathematics as is learned at school is not different from any other knowledge domain. It is far from formal: it has a content (germ. *Inhalt*). The same is true of higher mathematics; professional mathematicians *see* structures and mathematical entities; they see them because they have learned at school to see the Lockian general triangle in pictures on the blackboard. To explain how this can happen is not an easy matter; Kant has built his whole transcendental system to tune together sensible intuition and logical abstraction, but he can hardly be said to have succeeded. It is a fact that mathematics begins with sensibility, and this is *the* hard problem, both for foundations and for teaching. In algebra, formal practice is accompanied by a contentual (but loosely related) description of what is being done; the justification is a realistic one, inherited from the material (in Carnap's sense) way of talking of the first numerical experiences. Hence the widespread dissatisfaction and malaise among students and teachers, but also a crack where one can put a wedge to upset established habits of (non-)thought. If everything were clear and satisfactory, there would be nothing to reflect upon.

Logic as a (theory describing thought as a) formal game cannot present itself as the solution of the riddle. Insistence on formal logic to celebrate the triumph of reason has the opposite effect of making a caricature of the thinking (wo)man. For example Richard P. Feynman claimed that «mathematics is language plus reasoning, a language plus logic, that is a tool to reason». But at the same time according to Feynman mathematicians cannot properly (be said to) *reason*. They «deal only with the structure of the reasoning, and are not interested in what they are talking about. They do not even need to know what they are talking about, or, as they say, if what they are saying is true [...] Logical deduction can be applied even if one doesn't know the meaning of those words [in the axioms]. If the statements of the axioms are carefully worded and sufficiently complete it is not necessary that he who makes the inferences has any knowledge of the meaning of the words in order to deduce new conclusions in the same language». Then the poor mathematician who works with the equations into which a phenomenon has been translated, since the symbols tell him nothing at all has no guide line for his reasoning, but the precision of mathematical rigour. The physicist on the contrary, more or less knowing the results he is expecting, can try a guess and so proceeds rather directly to the heart of the problem. (Feynman probably thought that mathematics is a tool for reasoning only if it is a language plus logic plus a physicist as a user.)

The teaching of mathematics pushes towards the mastering of the formal, which reaches its highest with logic. At the beginning of the logic courses students are informed, probably for the first time, that there is a distinction between syntax and semantics; grammatical correctness and meaning part their ways. This is the loss of innocence: numbers disappear - whatever we seemed to be talking about, also disappears. It really looks as if the final goal were that of becoming Feynman's cripple, or a monster, as far as reasoning is concerned. The semantic notions are used to show that the formal rules have general validity; thereafter rules can play the formal and apparently impossible game. But at the same time this is not the policy pursued in the parallel mathematical courses, where discourse continues to be an apparently meaningful one, concerning various types of entities. To fill the gap, reference and meaning are (said to be) the concern of another part of logic, namely semantics; semantics takes on the task of describing, or defining the strange things one seems to talk about in mathematics. The completeness theorem connects the two aspects in quite a satisfactory way - so it is said. Paradise is thus immediately regained after the loss of innocence. Is it so easy? Can the promises be fulfilled?

Enter the dramatis personae

The completeness theorem is actually a family of theorems. Not only has each logic - classical, intuitionistic, modal and so on - a completeness theorem, but moreover every logic has a family of completeness theorems, one for each particular pair of poles that are connected by the theorem. On the one hand there is a *calculus*, a rule system that generates *derivations*, viz. sequences of (correspondents of) phrases, or (representations of) reasonings; on the other hand there is a notion of *logical truth*, and correspondingly of logical consequence, logical validity, logical satisfaction - all related and interdefinable.

Although the notion of logical truth can also be conceived in many different ways, nowadays it is usually presented as *truth under all interpretations*, under a great variety of notions of “interpretation”. The latter becomes then the crucial notion. “Logic” is actually an ambiguous term, usually meant to denote both a system of (axioms and) rules of derivation and a semantics - viz., a class of interpretations for the languages of a certain kind. To stress the first aspect it is preferable to talk of a logical calculus, while by “logic” it is usually (but not necessarily) meant that also semantic notions are involved. But the latter need not be uniquely determined by the name of the logic. Intuitionistic logic, for example, has Kripke semantics as well as the semantics given by infinite matrices, topological spaces and so on.

Each completeness theorem states, for a given pair <calculus, semantics>, that the two notions of derivability and of logical consequence are (extensionally) the same relation on the set of sentences, and give rise to the same set of theorems (derivable sentences, logically valid sentences). When completeness holds, the calculus is said to be complete with respect to the semantics. The usual schematic notations for these concepts are well known; one writes

$$\models A \text{ if and only if } \vdash A,$$

to state that A is logically true if and only if A is derivable from the empty set of assumptions. More generally, for every sentence A and every set of sentences T one writes

$$T \models A \text{ if and only if } T \vdash A,$$

(*strong* completeness) where $T \models A$ means that A is a logical consequence of T, and $T \vdash A$ means that A is derivable in the calculus from the (sentences of the) set T. $\models A$ and $\vdash A$ are a particular case of $T \models A$ and $T \vdash A$ when T is the empty set $\emptyset$, and the completeness property is a particular case of strong completeness. As for $T \vdash A$, we assume that the reader is acquainted with one of the usual calculi presented in any introductory textbook, a Hilbert calculus, or natural deduction, the *sequent* calculus or the *tableaux* method, so that we know what it means $T \vdash A$ with reference to this calculus.

The big deception begins when one describes several calculi and just one semantics, tacitly *the* semantics. As a corollary of the different completeness theorems for the different calculi and the same semantics one has the equivalence of the calculi with respect to. derivability: the same theorems can be derived in any two of them. Giving one official semantics is easier for classical logic than for intuitionistic logic, where it is more difficult to ignore the several semantic possibilities. But we will try to unveil the deception sticking to the classical logic. (Boolean valued semantics is not a real alternative, owing to the representation theorem, but it should at least be mentioned.)

As remarked above, semantic notions are mutually interdefinable. They all depend on a basic one, namely the following: given an interpretation **M** and a sentence A, by $\mathbf{M} \models A$ one asserts the fact that A is *true in* **M**; **M** is then also said to be a *model* of A. If A is not true in **M**, A is said to be false, and one writes $\mathbf{M} \not\models A$.

Warning "A is true in **M**" is not defined by the notation $\mathbf{M} \models A$, which for now is only an abbreviation. The definition of $\mathbf{M} \models A$ depends on the languages and above all on the notion of *interpretation* for a language. We'll come back to this main issue after recalling some related notions.

For any set T of sentences, we let $\mathbf{M} \models T$ be an abbreviation of the statement that for every $A \in T$ $\mathbf{M} \models A$. Then **M** is said to be a model of T. $\mathbf{M} \not\models T$ mean that there is at least one $A \in T$ such that $\mathbf{M} \not\models A$. If $T = \{A_1,..., A_n\}$, $\mathbf{M} \models T$ if and only if $\mathbf{M} \models A_1 \wedge ... \wedge A_n$, where $\wedge$ is the conjunction symbol.

A sentence A is said to be *logically true*, or logically valid, if $\mathbf{M} \models A$ for every interpretation $\mathbf{M}$ of the language; one writes then $\models A$. A sentence A is said to be *satisfiable*, or semantically consistent, if there is at least one $\mathbf{M}$ such that $\mathbf{M} \models A$; A is said to be *unsatisfiable*, or semantically inconsistent, if there is no $\mathbf{M}$ which is a model of A. A sentence is unsatisfiable if and only if its negation is logically true. A set T is said to be satisfiable if and only if there is an $\mathbf{M}$ such that $\mathbf{M} \models T$, otherwise it is unsatisfiable.

A sentence A is a *logical consequence* of the sentence B, in symbols $B \models A$, if for every $\mathbf{M}$, whenever $\mathbf{M} \models B$ then $\mathbf{M} \models A$. One also says that B logically implies A. In classical logic, A is a logical consequence of B if and only if $\models B \rightarrow A$, where $\rightarrow$ is the implication symbol. A sentence A is a logical consequence of the set of sentences T, written $T \models A$, if for all $\mathbf{M}$, if $\mathbf{M} \models T$ then $\mathbf{M} \models A$. If $T = \{A_1, \ldots, A_n\}$, $T \models A$ if and only if $A_1 \wedge \ldots \wedge A_n \models A$ if and only if $\models A_1 \wedge \ldots \wedge A_n \rightarrow A$.

The rules of a calculus have by definition a very special character; they might be called *mechanical*, with reference to the effectiveness of their application and of the recognition of their applicability; they might be called *syntactical*, with reference to the fact that they are tuned on the syntactic structure of the formulae. They are usually finite in number, or in any case they form an effective set, to ensure that all possible derivations can be effectively generated. The decidability of the notion of proof and the semidecidability, also called partial effectiveness, of the notion of theoremhood, follow as corollaries. The fact that theoremhood is only semidecidable appears to be well suited to the limited means of any finite agent, as human beings, as well as machines, are supposed to be. Semantic notions on the contrary do not usually have an effective character, they are defined more or less at the same level of abstraction as mathematical infinity (e.g. the set-theoretic semantics). This may explain why logicians regard completeness as a prerequisite for a logic to be acceptable. As a matter of fact, the requirement has a double import: according to some people, it is debatable whether a logic can be considered such if there is no complete calculus for it. On the other hand it is doubtful that a logic can be given by a calculus which is not complete with respect to some (even weird) semantics.

Approaching the problem, which doesn't seem to be a problem of every-day experience, but at variance with it; it appears that is not a case of saving the phenomena, but of exploding them

Any educated person who is logically *naïve* (that is, who hasn't entered a logic course) does not see a clear cut distinction between derivation and logical consequence; these notions must be severed one from the other before being again unified at a higher conceptual level. Usually the non-technical term "deduction" is used to denote a chain of inferences. *Prima facie*, or before the Fall, deductions and inferences partake of both sides: they have to do with intuitive logical consequence, but at the same time they'd better be done in small steps. Students might remember the impression of their first geometric proofs, with the (small) step by (small) step structuring of the reasoning; each statement had to depend from previous ones, to be dutifully mentioned (the statements, not the way and the why of the dependence). But students and teachers are seldom aware of the existence of fixed systems of rules; of such rules, they have at best a practical and casual experience, no explicit presentation. Rules were never mentioned to them. Perhaps that is the right thing to do, for you cannot mention rules without entering in the kind of global discussion on which we are embarking.

A logical calculus explicitly contains all its rules of transformation. It is often said that a completeness theorem proves that the rules of the calculus are sufficient. This may be reassuring but it is by no means clear. Sufficient for what? What is the nature of the rules? what is the effect of following them? What if we had forgotten some? But is it possible, and how and where could we have forgotten a rule? If rules are in our mind, their working should be familiar. There is an empirical proof that they are sufficient in the fact that looking through all the reasonings registered through the history up to now (*à la* Peano) one doesn't find any other. One cannot say that one needs a new rule to make a reasoning which has already been done. One could speculate that with the evolutionary growth of the dimension of the brain some new capacity to do (now) inconceivable inferences could emerge; then the completeness theorem would falsify the speculation, unless also new notions of logical consequence should be invented (this type of shaky speculations on evolution may be found also in some debates on Church's thesis).

It is true that mathematical proof have not always been the same, but when a recalcitrant proposition has at last been proved (e.g. Fermat's last theorem) the success has not been due to a newly invented rule, but to the intelligent use of new definitions. An example of an inference that at a certain time has appeared to be new and revolutionary is Saccheri's *consequentia mirabilis*, but later it was found that the rule had been already used by Euclid. Of course, there is always a first time for everything, but for logical inferences this first time seems to be always a rather ancient one, in a very limited period of the development of civilisation (as far as the western world is concerned). An important example of a demonstrative technique which has appeared later, in a relatively recent historical period, is the induction principle (disguised as infinite descent in Euclid and Fermat). Indeed, it took a great labour to clarify the issue of induction, but the difficulties concerned the concept of natural number, not purely logical questions.

Besides geometrical proofs, students meet another kind of proof in algebra, i.e. manipulations of equations. Although these are still stepwise procedures, they look different: *prima facie* they don't have the form of a sequence of statements, which makes it even more difficult to understand what is logical reasoning. As a justification of the steps, some properties of numbers are perfunctorily mentioned, but again the rules are hidden. In fact, they are even stranger that inference rules for plain sentences (in equational manipulations, such rules as the substitution of *equal* distinct terms is imposed as obvious by the prevailing teaching philosophy). Proofs, through a purported logical reasoning, are meant to establish the validity of a mathematical formula in a numerical system (let us say for convenience the real numbers **R**, though students do not have any idea of what it is, a fact that could ultimately prove a blessing for them).

To fix ideas, a typical and extremely simple example of reasoning is the following method to establish the identity

$$(x + 1)(x - 1) = (x^2 - 1).$$

Starting from

$$(x + 1)(x - 1) = x(x - 1) + 1(x - 1),$$

and obtains

$$(x + 1)(x - 1) = x^2 - x + x - 1,$$

whence

$$(x + 1)(x - 1) = x^2 - 1.$$

For the purpose of justifying the algebraic transformations - and not just applying them - people on both sides of the table (students and teachers) refer to certain facts true for R. I have said "true for R", and not "in R", as a logician or a structuralist mathematician would say, because the idea of a sentence "true *in* a domain" is not natural at all; the common wisdom is that a sentence may refer to a domain, and be true (or false). The distinction will appear meaningful later on. Explicit reference to truth may be avoided by saying that one uses certain "properties" of the real numbers (meaning of course some properties which hold for the real numbers). Every step of the proof gives a true statement, and of the properties used there is no trace, but their effect. The properties one makes (tacit) reference to are simply items in the general common background knowledge; one similarly knows that France is a presidential republic and has a certain well defined electoral system.

These reasonings, as they are called, are descriptions of facts, as in normal speech; if they use (implicitly) the notion of "truth", it is not that of "logical truth"; on the contrary, logical truth appears entirely off the mark; one is talking of precise things and (hopefully) telling the truth, without qualification. But then it is difficult to conceive that the truth of one fact *depends* on the truth of another (in what sense in real life do we make appeal to an intuitive causation?) and that one uses a rule, instead of cunning attention, to pass from one fact to another; this is one reason for our reluctance to mention rules. If the teacher talks of logical truth and of logical consequence, it is likely that the adjective is construed in an intensional sense. The sense of "intensional sense" is not clear at all, but to philosophers. To normal beings it is only a kind of rhetorical emphasis: "logical" evokes necessity, a state of affairs which imposes itself with irresistible force and a conclusion one cannot refute (perhaps for the sake of our common rationality).

To say that the single steps are particular instances of general logical laws is not much help, or it is worse, because then one does not see the logical (argumentative) development. It is only the specialist who can see the effect of the delicate substitution rule (which by the way is not logical at all, but a property of equality, and only in a devious way can be replaced by *modus ponens* on axioms). Thus, any appeal to logic are at best confusing; its effect is at most to force the mention of the relevant properties, but without seeing how they work, and the link between the steps. What remains is the yearning for precision, verging on pedantry; the received message is that of excluding intuition, in a broad sense. This in fact is the traditional role assigned to formalisation. But why should one want to bar intuition, if correct and useful? A straightforward geometric proof of the above algebraic identity (going back at least to Euclid) consists in drawing a square with side x, and superimposed to it a rectangle with sides $x - 1$ and $x + 1$; then one immediately sees that the difference of the areas is 1. History teaches that intuition was banned when it began to lead people astray by *not* allowing them to see - but this is another long and tangled story.

It is by no means easy to justify the notion of truth involved here ("for numbers it is true that..."). As a matter of fact, both at high school and at college level, one encounters no precise definition of numbers and truth. One works with systems of notations, representations without reference, and some rules of thumb learned by heart and accepted, often to please the teacher, according to Russell, or on empirically inductive grounds. Some other rules are obtained via the kind of reasonings described above.

As is well known to logicians, in order to proceed according to the book, one should explain that the aim of the above reasoning, as well as the aim of any proof, is not to show that the formula is valid in R, but that it has a logical generality whose boundaries are fixed only by the axioms. Certainly, it is not easy to explain that one is doing logic when appearances (as conveyed by language, side comments, motivations) point to the contrary. The question arises of the relationship between one particular concrete argument and its logical counterpart lurking behind. Is it a case of instantiation, inclusion, paradigm, generalisation? Is it the Hegelian spirit becoming flesh? Teachers have a tough job here and deserve all our sympathy.

The high road to the logical and axiomatic point of view should probably begin with the assertion that the given argument contains some hidden facts, or unspoken assumptions. That would come as no surprise to the students, but let us chase the hidden assumptions, as in a game. It would be easier to highlight them, if more stress were put on talking, instead of blind manipulation. The use of incomplete mathematical formalism and jargon has this paradoxical negative consequence, that the various formulae appear as snapshots of situations; one forgets that one is building a discourse. Making ourselves understood requires a minimum respect for grammaticality; one should not indulge in leaving truncated phrases hanging in the air, and other such sloppiness.

From careful inspection of the proof (*warning*: this is misleading, there is no proof to inspect, but only a proof to build) the following assumptions emerge: distributivity, 0 as neutral element of addition, $x + 0 = 0 + x = x$, the inverse property, either both $x + (-x) = 0$, and $(-x) + x = 0$ or one of them and the commutativity of addition, and 1 as neutral element of multiplication, $1x = x = x1$. If we call A the conjunction of these assumptions (the above mentioned "properties") then the conclusion of our reasoning is

$$\text{A implies } (x + 1)(x - 1) = (x^2 - 1).$$

This doesn't mean that if you believe A then you (have to) believe the identity, or that A has some long-distance effect on the formula, or other fancy versions of "implication". It means that A logically implies the identity, or that

$$[\mathrm{A} \rightarrow (x + 1)(x - 1) = (x^2 - 1)] \text{ is logically valid,}$$

according to the previous definitions - that is, in every interpretation in which A is true also the identity is true. Using the preferred terminology, A can be said also valid as an open formula, in the sense that the universal closure $\forall x\, \mathrm{A}(x)$ is true (same for the identity). The reasoning is a logical one in that it allows us to state the following kind of conclusion: that something is true in all possible interpretations. How the result can actually be achieved we will see later (in fact, A produces the above identity, via a number of manipulations of its components; this is why all components of A must be mentioned, but this is a discovery embodied in the completeness theorem). Now the first problem is: how to present to the students the idea of (all possible) different interpretations?

While variables range over numbers, operation symbols denote fixed operations. School algebra still pertains to a pre-Hilbertian epoch, in that the symbols + and × have a fixed, not a variable meaning. Hilbert in the twenties was the first one who explained that in mathematical logic every part of the alphabet should be construed as a variable (over objects of suitable kind).

Different interpretations could mean different algebraic structures; for instance, the above identity is also valid in the natural numbers, though one usually does not consider it when doing arithmetic. The identity is proved only after the development of methods of calculations involving numerical variables, as is systematically done in college algebra - with almost exclusive reference to rational or real numbers. Anyhow the other numerical structures that could be known to students are substructures of the encompassing one given by the reals; the operations on these substructures are not essentially different - they are mere restrictions of the original ones. It might be useful, though by no means easy, to consider also some non-numerical structure, thus opening the way to the study of (hints of) abstract algebra. Different interpretations might well involve such artificial and provoking examples as the tables, plates and glasses invented by d'Alembert and Hilbert for geometry. Too far-fetched at a first stage. A little bit of qualitative, non-numerical mathematics, like graphs and discrete structures could serve as a good introduction; they are representations of fragments of reality, and interpretations are introduced in mathematics through the many different systems of representations; the more we have, the better.

Talking of a whole class of interpretations of a language is not easy when, to be honest, the audience is not even acquainted with one. If one tries to define the reals one ends in a bees' nest. But the situation could be positively exploited, in a first stage: ignorance of the details of any interpretation could help in talking of the generic one, whose details do not matter, but only some very general features, which turn out to be common to all. This is what the set-theoretical notion of structure will achieve.

We are faced anyway with the problem of saying what an interpretation is. We must refer interpretations to languages, although no language *per se* is singled out in college algebra. Thinking of the above identity as relative to one or another type of numbers changes the interpretation, but only by slight corrections, refinements, not really *changes*. We didn't have an interpretation, but only some knowledge, though vague, and difficult to talk about - which is not surprising, because knowledge comes with and within the spoken language. The situation for mathematics is the one inherited from our natural language; linguists say that natural language is *descriptive*, to be taken at face value. There is no need to look for *interpretations* of the language as such, as a global structure, and we do not know what it could mean. When speaking, we are at most concerned with precision, with the elimination of vagueness and ambiguity of single special terms and constructions; this is the meaning of “interpretation”. At least in principle, refinements are introduced in scientific languages thorough careful definitions (although there is a certain indecision about scientific languages; they are either fragments of the natural language, polished as said, or else they are full-fledged formal languages of the kind treated in logic).

The descriptive nature of language is unproblematic, as far as talk does not involve mysterious things, the *arcana* and *invisibilia*, from which hermeneutics takes its origins; but the latter is the effort to interpret God’s *verbum*. Facts and stories of the Holy Bible are pointers to spiritual realities, or to a divine project whose working and code remains in the hands of God, whatever the interpretative efforts from the hearer. In their struggle to understand God’s word men have learnt much about language, especially about how to talk also through enigmatic sayings, which according to St. Augustin are made of things that do not exist; they have learnt perhaps more than about mathematics.

In naïve mathematics there seems to be the same kind of talk which is literally descriptive, as is natural language. Interpretations are not needed and not looked for, they simply are there, along with the stream of discourse. When one raises the problems of the denotation of terms (proper and common names, descriptions) and of the truth of statements, these problems look artificial and unnatural. But there are some differences which should be a warning; in normal speech one sometimes explains what one is talking about by reducing it to more familiar things. This is not the case with mathematical terms. Moreover, in mathematics meanings are not conveyed by the words as such, but by their rules of use, by the algorithms to which they are subjected.

Think of the way rational numbers are introduced, with a blend of practical intuitions about dividing a cake (fractions as operators) and numbers in decimal notation, obtained by the algorithm of division with quotient and remainder, all mixed and interrelated.

Knowledge contained in a set of algorithms on these disparate notations is acquired before the question "what are the rational numbers" comes to the fore of consciousness. This is an important fact, one that in a sense shall be saved by the completeness theorem, albeit through a devious route. For the present, we introduce a confusing complication if we ask the student to go and look for *other* interpretations. The invitation has no sense, if the interpretation is in the notations and in the rules for handling them.

Let us compare mathematics with the stories about non-existent beings, literature characters, or impossible animals of children's tales, such as the winged horse or the chimera. We often know everything there is to know of them, as if we had daily dealings with them. Could it be the same with numbers? If we are asked what imaginary beings are, the answer is a piece of information about theirs deeds, and when and where they have been living. This information is actually false, if our truth criteria depend on witnesses and certified documents. Sometimes one can also prove that they are false, but falsity is an extrinsic matter, the stories are consistent, and verisimilar. And when the information is true, as in the case of existing persons, the information has the same nature - it is expressed by statements of the same form. If you ask whether (or what it means that) John loves Mary, the answer will be a description of certain inner states and actions of someone who is in love: sighs, thoughts focused on the beloved object. However, the question and answer apply indifferently whether John and Mary are existing persons you are acquainted with, or they are from the pages of a novel.

Staying within the use of language has the effect that the explanation of the sense of a discourse (through synonyms) is unavoidably circular; a certain circularity is absolutely unavoidable with fundamental concepts. For example, in some definitions of life the condition of reproduction is included, i.e. of giving life to some new being. The same holds with the definition of consciousness. To break the circle one has to change level, or language. For example, in the case of love it can be informative to pass from the psychological or behavioural language to the physiological one; this is what is meant by a reductionist explanation. It is debatable whether this kind of explanation, though undoubtedly useful, is a source of meaning.

Mathematics is probably in the same predicament as the products of fantasy. If you have an interest in fantastic beings, that are the product of a fancy and coarse genetic engineering, you are not really interested in how it is possible that two eagle wings connect on the back of a horse; you are interested in the fact that the horse can fly and whether it will fly to the moon. If asked what it means that $0 < 1$, it is not sensible to answer recalling the set-theoretical (reductionist) construction of Dedekind cuts and of the order relation on them; an answer more useful and to the point will probably recall and explain some related facts, like $x < x + 1$, the notion of positive numbers and the like. What do we know about the objects and beings we talk about? Is it the stuff they are made of, or else it is the set of statements we can consistently utter about them?

These are just a few of the reflections we can make on the basis of every-day (including school) experience concerning the sense and meaning, or lack thereof, of the mathematical statements we use and produce in a naïve way: a lot of questions and doubts and few certainties. Let us pass now to the theorem and its proof, to see if some new light can be obtained, as promised. Science introduces distinctions which are not grasped by the naked eye, and of which one does not feel the need - to split the atom for example. The completeness theorem could help towards a better understanding of this muddle of problems; not to simplify it, but at least to have a more complete vision of all sides of the issue.

First tentative plunge into the proof, which comes to nothing because the two horns to be connected turn out to be one and the same, so there is nothing to prove

There are many proofs of the theorem, each to be recommended for one feature or another, each throwing light on a different corner; but the lesson we are interested in will emerge from any one of these proofs. First, one has to choose the two horns of the question: a system of rules and a notion of interpretation. Since we still do not have an inkling of the latter, we will choose for now a calculus and we will see how far we can go without committing ourselves in a precise way to what is to be an interpretation. Rules are man-made, while interpretations are tricky and slippery. As we have said, we do not have any experience of the latter through our usage of natural language. It is possible, however, to begin to talk of them by stipulating some conditions that any notion of interpretation must satisfy. This is the typical way of mathematics: in every field one has same partial information, coded by suitable axioms or by accepted assumptions. Building on its basis, one begins to say something and thus goes on increasing the knowledge. For example, we don't know whether (it is right to say that) points are - as we find in Euclid - entities with no dimension, and lines are entities with one dimension. We know that two points determine one line, and two incident lines meet in one point, and some other things of this kind, independently of a definition of dimension. By developing geometry on the basis of this knowledge we'll eventually meet with the problem of a metric, dimension and so on. By pursuing our study, we are confronted with new problems, acquire new knowledge and new wisdom, and sometimes it may happen that what was temporarily accepted as a necessary condition becomes also a sufficient, or a defining, one.

We know two or three things about interpretations, even if we still don't know what they are. Using our knowledge, and without saying what they are, we can already prove the correctness (or validity) of the rules - to fix ideas, let us refer to the rules of natural deduction. A rule such as "from $A \wedge B$ infer A" is correct, because in any interpretation - whichever it be and whatever they be, and whatever it means to hold in an interpretation - if $A \wedge B$ holds in the given interpretation then also A holds. Similarly, the rule "infer $A \wedge B$ from A and B" is correct, because in any interpretation in which the premises A and B are separately true then also $A \wedge B$ is true. The other logical particles are treated in a similar way. To find a situation, or an example of a meaningful speech in which the particle called "conjunction" does not behave in this way is not an easy matter, which we can forget for the moment. The case is different for a particle like negation, where it is rather easy to find alternative, non-classical, uses. It turns out that such alternative uses give rise in a rather natural way to alternative logics, but if you are a teacher you can cheat a little and with some rhetorical performance you can be convincing about the contrary. When experience is limited, the available examples are of a classical kind, and people are always ready to believe that their experience is general and valid for anybody (anyway, we shall only consider classical logic).

We are determined not to accept a notion of interpretation that does not satisfy the above conditions for conjunction, as well as the other usual conditions for the other logical particles. So our calculi will be correct by definition. Somebody claims that those conditions are a definition of the logical particles, but it seems more sensible to say that they are (also, at the same time) a partial definition of interpretation.

If we want our rules to be correct, interpretations must satisfy certain conditions; on the other hand, we accept the rules because we tacitly believe that a semantic justification must be formulated along those lines. But the belief is not so neutral; in assuming it we believe in fact that semantic valuations depend on the grammatical structure. Stated otherwise, we accept compositionality, and this is a first inkling that semantics tends to be identified with grammar, as in natural language.

One can anticipate an objection: when we introduced truth tables for connectives we used interpretations, so we know what they are. Could it be then that difficulties arise only with quantifiers? While it is true that interpretations for predicative languages pose peculiar problems, and that their presentation is apparently different from that of propositional interpretations, it is a fact that the differences are inessential. Indeed, when interpretations are used for truth tables, their scope never goes beyond what is necessary to prove the correctness property. An interpretation amounts to assigning "true" and "false" values to the propositional letters, and to uniquely extending these truth value assignments to all propositions, according to (truth-functionality) principles of the following type: if A and B have value "true" then also $A \wedge B$ has value “true”, and so on. We use abstract finite functions. We have written “true” in quotation marks because we do not want to consider the assignment of the value as the attribution of a well defined property; we have no right to do so since we have not anticipated any analysis of the notion of truth. For our assignments we are not using names (truth, falsity), or their corresponding adjectives, but labels. In effect other labels, such as 1 and 0, or T and F, or T and ⊥ are also used as an alternative to “true” and “false”.

Truth tables actually are nothing more than the conditions we want the “true” and “false” of the real interpretations to satisfy, independently of what they will turn out to be. It is a new corroboration of the chosen axiomatic treatment of this notion. Acceptance of the truth table for conjunction is equivalent to saying that the elimination and introduction rules for conjunction are correct. Since everything works smoothly for propositional logic on this skin-and-bone basis, the suspicion arises that this is all that matters, and we don’t need to know anything more. Notice that the same is true for interpretations of predicative languages, though they are a little bit more rich and complicated. If it is so, however, it is curious that some people (e.g. in Artificial Intelligence) claim that describing a structure for a language is tantamount to giving an ontology, or a description of the world.

As a calculus, we choose now the method of the *tableaux*, also known as semantic or analytic tableaux. The motive for the choice will become clear later; the name itself is a challenge. There is no misunderstanding here: we choose the semantic tableaux as our syntactic calculus. This method has the advantage of extreme ambiguity, consistently with the situation of ambiguity we are rejoicing in. The method is presented as a set of precise, syntactic, mechanical rules, as in any other deductive system, and the adjective “semantic” is a thorn in the back that need to be explained.

In order to describe this method, let us suppose we want to establish whether the sentence $A_1 \wedge \dots \wedge A_n \rightarrow B_1 \vee \dots \vee B_m$ is a theorem (a logical theorem; this word, just like “deduction”, is often used informally to denote without distinctions both a logical truth and the conclusion of a sequence of deductive steps). This is the concept we are interested in, and if it possible to get it without the syntax-semantics dichotomy so much the better. We open a table with two entries, two columns labelled I and II; at the initial stage we insert $A_1,\ldots, A_n$ in I and $B_1,\ldots, B_m$ in II. The cases that $n = 0$ or $m = 0$ are allowed. Both these initial sentences and those to be inserted later must be ordered in some way; the order is not essential for the global properties but only for the implementation of the procedure. By "the first sentence in a table" we mean the first sentence in this order, and when we speak of inserting new sentences at the end of a column we mean that the new sentences are assigned an ordinal number greater than those already used in the table. When a sentence is selected for consideration according to one of the rules stated below, it will be marked - say by a star - to prevent it from being considered more than once. For technical reasons, our tables will only contain sentences (i.e. variable-free, formulae) and open formulae will not be allowed.

At stage n, we first check that some sentence does not occur, starred or not, in both columns; if so, the table is (said to be) *closed* and the work is terminated. Otherwise, we take the first unstarred sentence in the table; if there aren't any, we are done, and the table is said to be *terminated*; if A is chosen sentence, we apply one of the following rules, according to the form of A and whether A is in column I or in column II:

- if A is atomic, mark it and go to the new stage;
- if A has the form $\neg B$ and is in I, mark it and add B at the end of column II;
- if A has the form $\neg B$ and is in II, mark it and add B at the end of column I;
- if A has the form $C \wedge D$ and is in I, mark it and add both C and D at the end of column I;
- if A has the form $C \vee D$ and is in II, mark it and add both C and D at the end of column II;
- if A has the form $C \rightarrow D$ and is in II, mark it and add C at the end of column I and D at the end of column II;
- if A has the form $C \wedge D$ and is in II, mark it and divide the table into two subtables, each of which has the same sentences as the given one, in the same columns and with the same order and the same marks, and moreover one of them has C added at the end of its column II and the other has D added at the end of its column II.

Warning If a table splits into two subtables, under the action of the above rule or of the others we are going to mention next, from that point on it is meant that both subtables are to be developed in parallel. Each subtable is a table on its own, when one works on it, and is called a table; it can divide into subtables, which are called also subtables of the original one (rather than being called subsubtables); the table generated by the procedure is the set of all subtables thus generated, some of them closed, some other not; there is a certain ambiguity, but from the context it should be clear whether by "table" one means a single two columns (sub)table or the set of all subtables.

Here are the remaining rules:

- if A has the form $C \vee D$ and is in I, mark it and divide the table into two subtables, each of which has the same sentences as the given one, in the same columns and with the same order and the same marks, and moreover one of them has C added at the end of its column I and the other has D added at the end of its column I;
- if A has the form $C \to D$ and is in I, mark it and divide the table into two subtables, each of which has the same sentences as the given one, in the same columns and with the same order and the same marks, and moreover one of them has C added at the end of its column II and the other has D added at the end of its column I;
- if A has the form $\exists x\, B$ and is in I, introduce in the language a new constant c not belonging to the original alphabet and not already introduced in the course of the procedure, mark A and add $B[x/c]$ at the end of I;
- if A has the form $\forall x\, B$ and is in II, introduce in the language a new constant c not belonging to the original alphabet and not already introduced in the course of the procedure, mark A and add $B[x/c]$ at the end of II;
- if A has the form $\forall x\, B(x)$ and is in I, the rule is more complicated to state: one wants to introduce in I all sentences $B[x/t]$ for every closed term t that comes to appear in the (sentences in the columns of the) table, also at later stages; then at stage n one begins to do all possible substitutions of the closed terms already present in the table, and adds sentences $B[x/t_i]$ at the end of I, unless they are already in I; then mark $\forall x\, B(x)$ but reintroduce it at the end of I, so that at some future stage, after having worked on the other sentences and possibly introduced in the table new closed terms thorough the application of other rules, you will come back to it and do the new substitutions and insertions. If there are no closed terms t such that $B[x/t]$ is not already in I, and there are no unstarred sentences in the table, then $\forall x\, B$ is not added at the end. The rule is long to state, but it should be rather obvious

(there are possible refinements to minimise the number of substitutions);
- if A has the form $\exists x\, B(x)$ and is in II, introduce at the end of II all sentences $B[x/t_i]$ for every closed term occurring in the table; mark $\exists x\, B(x)$ and introduce it at the end of II, with the same *proviso* as in the preceding rule.
After applying one, and only one of these rules go to stage $n + 1$.

A (two columns) subtable is closed if the same sentence occur in both columns; a table is *closed* if all its subtables are closed; the initial table is said to close if its development according to the described procedure leads at some finite stage to a closed table. A subtable is terminated but not closed if all sentences occurring in it are starred and it is not closed. A table is said to be terminated if all its subtable are terminated, whether closed or not closed. It is terminated but not closed if at least one of its subtables is terminated but not closed.

When the procedure is started on an initial table, three cases are possible: (i) after a finite number of steps the table is closed; (ii) after a finite number of steps the table is terminated but not closed; (iii) the procedure does not halt.

Propositional logic corresponds to the fragment of the above construction in which no rule for quantifiers applies. The method then always halts after a finite number of steps (we do not prove it, though the proof is easy and instructive, since it is a case where for a rather simple and intuitive property one needs transfinite induction). Hence there are only two cases: after a finite number of steps, when the procedure halts, either the table is closed or it is terminated and some of its subtables are not closed.

For a predicative language the procedure need not halt in general (however, it always halts for monadic languages). In case of nontermination, looking from the limit through König's spectacles, it is easy to see that there must be at least one subtable (one made of two columns) that is infinite and not closed, and in which every sentence is starred. Hence also in this case, the possible alternatives reduce to two, albeit in a non-effective way. Either the table closes down in a finite number of steps, or there is at least one subtable, either finite and terminated or infinite, which is not closed. In the first case we say that $A_1 \wedge \ldots \wedge A_n \rightarrow B_1 \vee \ldots \vee B_m$ is a theorem and in the other case that it is not.

Here is an example where we have an algorithm, which is quite easy to learn to use (or if you like, to implement), while it is less trivial to understand what it does and why it gives certain answers. By the above definition, *we* say that in the first case we have a theorem, but why? What does the algorithm say? Nothing, except that the table closes. The implementation could be such that when the table closes you see on your monitor the statement "$A_1 \wedge \ldots \wedge A_n \rightarrow B_1 \vee \ldots \vee B_m$ is a theorem". Or, we could modify the output so that on the monitor there appears the statement "$A_1 \wedge \ldots \wedge A_n \rightarrow B_1 \vee \ldots \vee B_m$ is derivable by means of the tableaux method", but that wouldn't be very informative. It would be the same as "the method say: YES", but yes to what? If we let the algorithm say "theorem", it looks like a definition of theoremhood, but not a very transparent and convincing one, since such an apparently logical notion turns out to depend on the method.

It is easier to see what is going on if we change the labels of the entries and instead of I and II we use **T** and **F**, respectively. **T** and **F** stand obviously for "true" and "false". Moreover, in the light of our technical definitions, let us reformulate the problem as that of determining whether $A_1 \wedge \ldots \wedge A_n \rightarrow B_1 \vee \ldots \vee B_m$ is logically true, or $B_1 \vee \ldots \vee B_m$ is a logical consequence of $A_1 \wedge \ldots \wedge A_n$. Then the method works as a search for counterexamples: the statement to be proved is a universal one, stating that for *all* interpretations (whatever they are) if all A_i are true then some B_j is true. Let us then investigate the converse problem - whether an interpretation exists in which all of the A_i are true and all of the B_j are false.

Though we don't know yet what an interpretations is, we can begin our search on the basis of the little we know about the (semantic) relations among sentences that are in a certain compositional (syntactic) relation. We know, or we want it to be so, that if in a hypothetical interpretation $C \wedge D$ turns out to be true then also C and D separately must be true; we similarly know that if $C \wedge D$ turns out to be false then the interpretation must have the property that C be false or else that D is false (different interpretations yielding different alternatives regarding C and D), and so on. By initially assigning **T** labels to the A_i's and **F** labels to the B_j's, we are just stating a hypothesis, namely that the counterexample exists: it is the interpretation that, if existent, would falsify the consequence relation. The existential hypothesis does not describe how the interpretation is, but how it has to behave with respect to the A_i and B_j if it has to falsify the consequence relation (or, equivalently, how the A_i and B_j have to behave with respect to the interpretation). One can see here also a proof by contradiction of the fact that $B_1 \vee \ldots \vee B_m$ is a logical consequence of $A_1 \wedge \ldots \wedge A_n$. Now in the development of the table, each subtable is a different search line, and at the same time an attempt - if not to build the interpretation - to increase by successive approximations the amount of information we have on the supposed interpretation. In the hypothetical interpretation, all sentences put into **T** up to any given stage should be true, and all the sentences put into **F** should be false.

It is quite easy to accept that if a subtable closes then there is an overt contradiction in the information concerning the tentative interpretation, so that search road is blocked. If all the subtables close, then an interpretation satisfying the initial conditions does not exist, whence one is entitled to conclude that $B_1 \vee \ldots \vee B_m$ logically follows from $A_1 \wedge \ldots \wedge A_n$, still knowing nothing about interpretations, except their compositional behaviour with respect to the logical particles embodied in the (motivations of the) rules. In particular, we assume that a sentence cannot be at the same time true and false in the same interpretation. (From now on for simplicity we will consider just one A and one B.)

Students are ready to accept this line of reasoning without perceiving that it is not wholly justified: while it is true that all explored roads lead to nothing, it is not so obvious that we have tried all possibilities. We have tried all the roads opened by the method, but it is conceivable that there are others. So there is still something to prove.

But this additional proof can still be done without a more precise definition of interpretation. The usual proof is as follows. We suppose to have an interpretation satisfying the initial conditions, and we show that at each stage there is at least one subtable such that all sentences in its column **T** are true in the given interpretation, and all sentences in its column **F** are false. Let us abbreviate this long stipulation by saying that the interpretation *respects* the subtable. It follows that this subtable is not closed at that stage, hence the table never closes. The proof is by induction on stages. The inductive hypothesis is that at stage n there is a subtable such that the original hypothetical interpretation respects this subtable; let us work on this subtable. If we apply the rule for $C \wedge D$ in **T**, then we add in **T** both C and D, which are true in the interpretation if $C \wedge D$ is true, hence the interpretation respects the new subtable at stage $n + 1$. If we apply the rule for $C \vee D$ in **T**, we get two subtables, one with C in **T** and the other with D in **T**; but if $C \vee D$ is true in the interpretation, then either C is true in it or D is true - say it is D; then the interpretation respects that of the new subtables at stage $n + 1$ which contains D. The same applies for the other rules and logical particles, using only the partial information we have accepted on the notion of interpretation.

If when the table closes there is no interpretation satisfying the initial conditions, that means that after trying all possible roads available from the tableaux method, we have failed; *in this sense* the method is a complete one, as far as the search is concerned. To avoid any misunderstanding, for the moment let us forget this notion of completeness; we will come back to it later, when comparing it with the logical one. Notice that the completeness of the search does not follow from the interpretations being finite in number - as far as we know, there might be infinitely many of them. However, the completeness of the tableaux search shows that it is enough to examine a finite number of (types of) them.

The other case is more delicate: what happens if the table does not close? Since this is the opposite of the first case, students again tend to think that there is nothing to prove: when the answer is of one type it is correct - therefore, when it is of the other type the opposite must hold, hence B is a consequence of A. But this again is a case of putting the cart before the horse; the underlying assumption here is that the algorithm is always correct, that what it does is always right. Students tend to assume for laziness that all algorithms are correct, just as physicists tend to assume that all series are convergent. To their justification, one must say that algorithms are usually presented accompanied by an intuitive motivation which is a kind of correctness proof. The usual presentation of an algorithm, as opposed to a formal program, is a mixture of syntax and intuitive semantics. Moreover, algorithms correspond to a well considered idea which embodies the solution of a problem (as series are representations of natural phenomena) so they must be right, unless some stupid error has inadvertently slipped in. But it could well be that the algorithm is entirely crazy, with no rational correspondence with our problem; it could thus be that in the second case no relationship holds between the answer and the fact that logical implication has been falsified - whence the method would of course be radically wrong.

Actually, the method is correct; but then, whenever a table is not closed there must exist an interpretation in which A is true and B is false; if such interpretation exists, we should be able to see it, grasp it, define it. Moreover, as a precondition to the proof of correctness, our construction should be independent of the method, since the latter only gives YES or NO answers. The method in fact allows us to see the interpretation. It has been so conceived in order to reach this goal - but we have a long way to go before we can see it, working out the *rationale* of the algorithm. It will turn out that going through the algorithm is the only way to see the interpretation, so the correctness proof is actually on shaky grounds, a kind of bootstrapping. What we have to prove is that for every subtable which is finite, terminated and not closed, or infinite, there is an interpretation which respects it, hence in particular the interpretation satisfies the initial conditions. Now such a strong existential statement cannot possibly be sustained without clearly saying what is an interpretation. We have delayed the answer up to now, but it should no more be possible to be evasive. We could say that the real proof is only now to begin; meanwhile, through the application of the tableaux method the problem has found the following reformulation: given a non-closed table, find an interpretation that respects the table. The completeness theorem can be newly stated in this form: for every non-closed (wholly developed) table there is an interpretation which respects it.

There are some clues for the proof; as it happens, the very description of the algorithm contains (a sketch of) the informal correctness proof. The rules were plausible as rules for the behaviour of sentences in interpretations. So convincing were the comments to the algorithm that one may have had the faint impression that the method did all the work to be done - and there is no need of the present appendix of the completeness theorem in the above reformulation. The reason was that in changing our labels from I and II to **T** and **F**, a meaning was given to our operations (as already noted, this often happens with commented algorithms). When we had labels I and II and didn't know what they were meant for, we built only sets of sentences; the formation rules for these sets were effective, but not transparent in their aim and effect. After re-labelling, we obtained a method for the determination of the truth and falsity of sentences (just as truth tables are such a method in the propositional case). It really looked like we were building an interpretation, evaluating all sentences down to the atomic ones. But if the method is a semantic one, as the very name suggested from the beginning, where is the gap that only the completeness theorem could fill? Perhaps it is better to start anew with more sharply defined distinctions.

Some technicalities; since the method is not restricted to a particular initialisation of the table, it is useful to learn how to read the possible answers form a logical point of view, according to the form of the initial data. We can have A in **T** and B in **F**, or only A (possibly a finite conjunction) in **T** and nothing in **F** or only B (possibly a finite disjunction) in **F** and nothing in **T**. The idea is always the same: if the table closes, there is no interpretation satisfying the initial conditions, which means in the three different case: (i) there is no interpretation in which A is true and B is false, hence B is a logical consequence of A; (ii) there in no interpretation in which A is true, if there was only A in **T**, hence A is unsatisfiable; (iii) there is no interpretation in which B is false, if there was only B in **F**, hence B is logically true. It the table does not close, there should be an interpretation satisfying the initial conditions, if we are able to prove it in a convincing way; hence in this case we have respectively: (i) B is not a logical consequence of A, (ii) A is satisfiable, (iii) B is not logically true. We do not discuss the case of an infinite initial set of sentences, which if effectively given can be inserted in chunks; the conclusions are the same, with the extra bonus of compactness.

Second start, with a new calculus, while the ambiguous one is reduced to a technique

Let us then take as a system of syntactic rules some (deductively) unambiguous calculus, such as a natural deduction calculus. We can still use the tableaux method for a preliminary investigation of the possibility to arrive at semantic notions.

As an immediate consequence of the basic definitions, we have that $T \models A$ if and only if $T \cup \{\neg A\}$ is unsatisfiable, or semantically inconsistent; for the natural deduction calculus, as well as for the other usual calculi, there is a formally analogous relation for derivability: $T \vdash A$ if and only if $T \cup \{\neg A\}$ is syntactically inconsistent. A set is said to be syntactically inconsistent if a contradiction $X \wedge \neg X$ is derivable from the set (otherwise it is said to be syntactically consistent). The completeness problem is then equivalent to the following: if a sentence, or a set of sentences, is syntactically consistent then there is an interpretation of the language which is a model of the given sentence, or set of sentences. In this form the completeness theorem is also known as the *model existence theorem*.

Let us suppose that the given finite set $T = \{A_1,..., A_n\}$ is syntactically consistent; let us put $A_1 \wedge ... \wedge A_n = A$ in **T** and let us show that under the consistency hypothesis the table does not close. This is a preparatory lemma. To prove it, we take inspiration for a previous proof to show that at each stage there is something which is transmitted down and which obstructs closure. For a fixed subtable, let us denote by T_n the set of sentences that at stage n are in the **T** column, and by F_n the set of the negations of the sentences that are in the **F** column. $T_0 = \{A\}$, $F_0 = \emptyset$ and $T_0 \cup F_0$ is syntactically consistent. If at stage n a subtable closes, then its $T_n \cup F_n$ is syntactically inconsistent, because for a certain sentence E it contains both E and $\neg$E. By induction on n we prove that at each stage at least one of the subtables generated up to that stage has the property that $T_n \cup F_n$ is syntactically consistent, whence the subtable does not close. It follows that at no stage all subtables are closed, and the initial table does not close.

The proof of the inductive step splits according to which rule is supposed to be applied at that stage. Let us consider at stage n a subtable for which the inductive hypothesis holds, that is $T_n \cup F_n$

is syntactically consistent. It the rule applied is that for the conjunction in **T**, then in **T** there is a sentence $C \wedge D$ and at stage $n + 1$ in the augmented subtable there are also C and D, while the rest is the same. Then $T_{n+1} \cup F_{n+1} = T_n \cup F_n \cup \{C, D\}$. (To be pedantic, the T_{n+1} is that of the new subtable, while T_n is that of the subtable of stage n, but there is an obvious embedding; we don't want to be pedantic.) If this set were inconsistent, then $T_n \cup F_n \cup \{C, D\} \vdash X \wedge \neg X$. But $C \wedge D \vdash C$ and $C \wedge D \vdash D$, and $C \wedge D$ is in T_n, hence $T_n \cup F_n \vdash X \wedge \neg X$, contrary to the inductive hypothesis.

Let us consider one more propositional case: given the described situation, let us suppose that the applied rule is that for the disjunction in **T**; then in **T** there is a sentence $C \vee D$; the table splits into two subtables, one with C in **T** and the other with D in **T**, the rest being the same; let us suppose that the ($T_{n+1} \cup F_{n+1}$ of the) first of the two subtables is syntactically inconsistent, otherwise we are done; then $T_n \cup F_n \cup \{C\} \vdash X \wedge \neg X$; with some admissible derivation steps, $T_n \cup F_n \vdash C \rightarrow X \wedge \neg X$, hence $T_n \cup F_n \vdash \neg(X \wedge \neg X) \rightarrow \neg C$, hence $T_n \cup F_n \vdash \neg C$, since $\vdash \neg(X \wedge \neg X)$; but $C \vee D$ is in T_n, hence for the rule of elimination of disjunction $T_n \cup F_n \vdash D$. Granted this, let us consider the second subtable, the one with D in **T**; here $T_{n+1} \cup F_{n+1} = T_n \cup F_n \cup \{D\}$, but since D is derivable from $T_n \cup F_n$, and this set is syntactically consistent, then so is $T_{n+1} \cup F_{n+1}$.

If the rule is that for the existential quantifier in **T**, then in **T** there is a sentence $\exists x\, B$ such that $T_{n+1} \cup F_{n+1} = T_n \cup F_n \cup \{B[x/c]\}$, where c does not occur in $T_n \cup F_n$; if $T_{n+1} \cup F_{n+1} \vdash X \wedge \neg X$ then $T_n \cup F_n \vdash B[x/c] \rightarrow X \wedge \neg X$; an easy and purely syntactic lemma assures us that if z is a variable not occurring anywhere in the derivation then $T_n \cup F_n \vdash B[x/z] \rightarrow X \wedge \neg X$, from which $T_n \cup F_n \vdash \exists z\, B[x/z] \rightarrow X \wedge \neg X$; but $\exists x\, B \vdash \exists z\, B[x/z]$, whence $T_n \cup F_n \vdash X \wedge \neg X$. The lemma means that if a letter c does not occur in the assumptions (or anywhere else, but this further condition serves to other transformations, such as the renaming of x by z), hence nothing is said about (what is denoted by) c in the assumptions, then c is in effect a variable parameter, independently of how it is classified in the alphabet.

If the rule is that of the universal quantifier in **T**, then $T_{n+1} \cup F_{n+1} = T_n \cup F_n \cup \{B[x/t_1],..., B[x/t_m]\}$, with $\forall x\, B \in T_n$; if $T_{n+1} \cup F_{n+1} \vdash X \wedge \neg X$ then $T_n \cup F_n \cup \{B[x/t_1],..., B[x/t_m]\} \vdash X \wedge \neg X$, and $T_n \cup F_n \vdash B[x/t_1] \wedge ... \wedge B[x/t_m] \rightarrow X \wedge \neg X$; but the antecedent of the implication is derivable from T_n, whence again $T_n \cup F_n \vdash X \wedge \neg X$. The other cases are treated in a similar way, and we conclude that if A is syntactically consistent then the table initialised with A in **T** does not close.

Where does this leave us? Apparently at the same point as before, with the only difference that distinct things are now kept separate. On the one side we have the rules of natural deduction, and the related notion of syntactical consistency; on the other side we have the interpretations, which however we haven't handled yet. Still missing is the lemma that if a table does not close then there in an interpretation respecting it. The search for interpretations has been prepared with the method of semantic tableaux, but this time we don't want and there shouldn't be the danger of confusing the tool with the goal.

Still, the borderline is very thin; in the above proof the various cases were easily handled because there is a strong tie between the rules of natural deduction and the rules of the tableaux method: same number, same grouping in pairs. A certain attention is required not to get confused and to be sure that no vicious circle is involved. The analogy is so strong that one can take a closed table initialised with A in **T** and B in **F**, reverse it and with few adjustment get what is also considered a derivation of A → B, this sentence at the bottom instead of at the top. The reversed table would be a derivation in the so called *sequent calculus*, whose derivations are isomorphic in this way to closed tables, to the extent that the two methods cannot be really considered two different calculi. This is the reason why in our first attempt we could forget about natural deduction and use only the tableaux method. As a result however we ended in confusion, with the suspicion not to be able to extract an independent notion of interpretation. But couldn't the same situation reproduce itself in the new setting?

Without a definition of interpretation we were able to proof the completeness of the search, that is the fact that all possible (ways to construct) interpretations were taken in consideration through the tableaux. This circumstance should mean that interpretations are not something radically different from what this method affords, that they are not “totally other”. It means at least that the information gathered in the course of the development of the method is sufficient to establish an interpretation. It is not only the case that the answer is right, but that the interpretation should arise as a bonus from the method itself. On the other hand, from natural experience we know that interpretations are not something extrinsic to be added to the discourse, but emerge from and in the discourse, and they tend to identify themselves with the discourse.

Since the method of semantic tableaux which generates sentences from sentences according to their form is a purely syntactic one, it seems proper to use it for a preliminary processing of sentences on which only derivability conditions are known. The difference between its rules and those of natural deduction are practically negligible, but for labels attached to the generated sentences. The natural deduction rules better correspond to the natural discourse, since in the stream of phrases there are just phrases without labels, not meta-statements about the truth of what one is saying - unless one is swearing or rhetorically engaged. The thesis that since natural discourse is meaningful, then it develops at a semantic level is a misleading one; language is directly meaningful in itself but the discourse develops at a syntactic level, not through a series of semantic statements. The word “true” is never used, except for questions (“is it true that…”), or to stress a point (“I tell you it is true that…”), or similar. The word marks the passage to the meta-language. You can of course take a normal discourse and insert therein a lot of “it is true that…” and of “it is false that…”, when there is a negation, but the effect is the same, and the two utterances have (a kind of) isomorphic meaning.

So we see why one is prepared to accept without questioning the positive answer of the table - if the table does not close there is an interpretation - but doesn’t feel it necessary to look for the interpretation in a separate manner; the reason is the confidence in the correctness of the underlying algorithm, but it is also the analogy with normal speech.

Solution with a cut of the Gordian knot: what we were trying to explain becomes a definition

Let us pause for a moment and try to extract what an interpretation could be from all we have said so far. If when the table is not closed we (want to) have an interpretation, then the interpretation could be the table itself. But what is the table? It is just a pair of sets of sentences **T** and **F** that are (separately and mutually) closed with respect to the following syntactic downward operations (and are, moreover, disjoint if the table is not closed):

- if $\neg A$ is in **T**, then A is in **F**;
- if $\neg A$ is in **F**, then A is in **T**;
- if $B \wedge C$ is in **T**, the both B and C are in **T**;
- if $B \vee C$ is in **F**, then both B and C are in **F**;
- if $B \vee C$ is in **T**, then either B is in **T** or C is in **T**;
- if $B \wedge C$ is in **F**, then either B is in **F**, or C is in **F**;
- if $B \rightarrow C$ is in **T**, then either C is in **T** or B is in **F**;
- if $B \rightarrow C$ is in **F**, then B is in **T** and C is in **F**;
- if $\exists x\, B$ is in **T**, then there is a closed term t such that $B[x/t]$ is in **T**;
- if $\forall x\, B$ is in **F**, then there is a closed term t such that $B[x/t]$ is in **F**;
- if $\forall x\, B$ is in **T**, then for every closed term t occurring in sentences of $\mathbf{T} \cup \mathbf{F}$, $B[x/t]$ is in **T**;
- if $\exists x\, B$ is in **F**, then for every closed term t occurring in sentences of $\mathbf{T} \cup \mathbf{F}$, $B[x/t]$ is in **F**.

Sets of sentences satisfying the above closure properties, possibly with inessential variants, are also known as Hintikka sets, or with other names, in the literature. Let us call them interpretations. The closure properties of a tableau follow from the fact that when a sentence is marked (and each sentence is eventually marked, because all sentences in a terminated or infinite table are starred) then other connected sentences are put into the table. Actually, for the universal quantifier in **T**, or the existential quantifier in **F**, the closure property is not so evident, without a closer look at the procedure. The intuition is as follows: when $\forall x\, B$ in **T** begins to be considered, its unstarred occurrence keeps being repeated, unless at some stage all substitutions with closed terms are already done, and all other sentences are already starred; if t is a term which for the first time occur in the table at stage m, then when $\forall x\, B$ is again considered at a stage greater then m then $B[x/t]$ is inserted in **T**. If this is true, as it was our original aim, then the formulation of the rule was right.

If we assume the above definition of interpretation (as a pair of suitable closed sets) we are finished, the completeness theorem is proved with perfect satisfaction, and nothing is left out. The proof of the lemma on (the existence of interpretations respecting) terminated or infinite tables has been replaced by a definition.

When dealing with truth tables for propositional logic, our procedure was not much different; there we went on assigning the values “true” and “false” to propositions, while respecting compositionality - that is, we formed the sets of true and false propositions, satisfying the above closure properties. But we never explained what it is meant to be true or to be false. Interpretations were defined as assignments of values respecting compositionality according to the truth tables. We (said that we) did semantics, but we didn’t define truth (or denotation), we just divided the sentences into two classes. Interpretations for predicative languages are analogous (along these lines, one could also further pursue a reduction of predicative to propositional logic, thorough the Skolem-Herbrand theorem). With reference to semantic notations, $\mathbf{M}$ would now be the pair $<\mathbf{T}, \mathbf{F}>$ and $\mathbf{M} \models A$ would just mean $A \in (\mathbf{M})_1 = \mathbf{T}$, the first component of the pair $\mathbf{M}$.

This however looks really like cheating, a perverse and typically mathematical trick: assume as axiom a theorem you are not able to prove, take as primitive a concept you are not able to define. What is lacking, and is a source of uneasiness, is the *sense* of the statements. If we want to say that the sense is in the discourse, when consistent, well, we would like to grasp it, to feel it, while we have only in $\mathbf{T} \cup \mathbf{F}$ a set of artificial signs (they are not even hieroglyphics, otherwise there would be a sense); where is the discourse?

Second ending, where a discourse is a discourse

Let us then replace the formal schemata of sentences by authentic statements, as given by some meaningful discourse. This is quite feasible: one can substitute the atomic sentences in $\mathbf{T} \cup \mathbf{F}$ with statements of the English language (in a consistent way, i.e., taking care that the same statements replace occurrences of the same atomic sentences, and - for additional safety - also taking care that different statements correspond to different atomic sentences). Upon reading all statements thus obtained from column $\mathbf{T}$, and reading the negations of all statements in column $\mathbf{F}$, we then obtain a meaningful discourse (except possibly for the order of statements). For example, suppose at the beginning we were given in column $\mathbf{T}$ the conjunction of $A \to B$ and $\neg B$, with A and B atomic. Applying the above procedure, our table splits into two subtables, one of which closes, while the other contains the initial sentences plus A in column $\mathbf{F}$. The resulting schema of discourse is given by $\{A \to B, \neg B, \neg A\}$; if A is replaced by "the sun shines" and B by "it is warm", then we can read: "if the sun shines then it is warm, it isn't warm, the sun does not shine", which makes perfect sense.

We have thus come across another notion of interpretation, the so called *substitutional* one: an interpretation is a (consistent and systematic) replacement as described above of meaningful statements of a real language to the schemata represented by the formal sentences. The mapping must also respect compositionality, but this automatically follows from our policy of replacing only atomic sentences, and letting the already developed tableau take care of the rest. With reference to the semantic notations, let M be the mapping into the English language, or into the fragment of meteorological forecasting, sending A to "the sun shines" and sending B to "it is warm". Let $\mathbf{M}$ be the extension of M obtained by respecting compositionality as above. Then $\mathbf{M}$ would be a model of $A \to B$. By writing $\mathbf{M} \models A \to B$ we now mean that if the sun shines then it is warm; similarly, $\mathbf{M} \models (A \to B) \wedge \neg B \to \neg A$ may state today's gloomy forecast; $\mathbf{M} \models (A \to B)$ stays for "$\mathbf{M}(A \to B)$ is true", in the sense of the English language. A rather devious but correct way of talking.

Under this notion of interpretation, a sentence is logically true if every substitution with meaningful statements transforms it in a true statement. Different interpretations can be thought of as different application of the formal schema to different knowledge domains.

The substitutional notion of interpretation is perhaps the oldest one. This notion had a fundamental role in the axiomatic thought at the end of the nineteenth century (Pasch, Peano, Enriques, Hilbert): people began to say then that theorems were invariant with respect to variations of the sense of the concepts involved in the axioms. As we shall soon see, although old and respectable, this notion has some drawbacks.

If we are not just content with examples, but we want to establish general facts, we have to be more precise with the notion of arbitrary substitution. We cannot simply say: replace arbitrary statements. The vagueness and undecidability of natural language must be taken into account. If a table is initialised with an atomic sentence A in **T**, the table is terminated and it is not closed; but if we replace now A with an inconsistent statement, such as for example "the sentence in **T** is false" - the liar paradox - it is no more true that we get a meaningful statement. It's no use to require to substitute only atomic statements to atomic sentences, since what counts as atomic in natural language is relative - "it is warm" could be a highly complex statement. The fault is with the natural language, which is all but consistent. Its inconsistency comes precisely from the use of the untrustworthy concept of truth.

We want to replace the atomic sentences in **T** with true statements, but - as we know from Tarski - it is impossible to define true statements in natural language. In order to avoid semantic paradoxes, we may try to resort to fragments of the natural language. Upon restriction, the notion of truth may become less naïve, and we may have more control on it - possibly scientific control, well beyond sensorial evidence (could our forms of truth-control be different for different knowledge domains?). Again the question arises: by putting such restrictions, do we still have a sensible notion of logical truth? Undoubtedly, restrictions are necessary: for instance, statements about statements should be avoided, to dodge self-reference.

Since we have to put restrictions, we could make a bold decision and make replacements in such a way as to always obtain discourses referring to the same subject, for example meteorology.

Then we wouldn't have to continuously change the notion of truth. There is still a potentially infinite variety of statements (for example, by use of temporal and geographical indexicals as in "the sun shines in the South"). We can choose to talk about the weather, atoms, particles, or numbers. If the latter are chosen, *prima facie* this is only for technical reason: indeed we cannot exclude the possibility that the table is infinite. If the table contains infinitely many terms, then our discourse shall have to implicitly refer to an infinite domain - to which purpose there is nothing better than the natural numbers. Moreover, if we want to deal with real proofs, we have to consider a domain whose appropriate notion of truth is sufficiently precise to be treated mathematically.

The completeness theorem we arrive at along this route reads as follows: a sentence F is derivable in the natural deduction calculus if and only if all arithmetic statements which result from F by substituting in all possible ways all relational and functional symbols of F with arithmetic relations and functions, turn out to be true. *Mutatis mutandis*, and with a contrafactual rewording anticipating the (arithmetic) model existence theorem, this is the statement of the theorem first proved by Gödel.

The layman will hardly find immediately convincing this reduction only to discourses about natural numbers; their generality is evident only to mathematicians, who (after Pythagoras) know that numbers are a sort of universal universe, in which every entity is representable by suitable codings. But each coding is a filtering process, and this already poses some obstacles to complete generality: for instance, how can we be sure that codes respect truth? Actually, the choice of the arithmetic language is the residue of a time when people believed that all of mathematics is reducible to arithmetic, and that natural numbers - as a direct emanation of the spirit, or creation of the mind - are not further reducible to anything else. The same kind of irreducibility would also apply to arithmetic truth.

The proof of the completeness theorem based on arithmetic substitutions still has some gaps to be filled, in order to become fully convincing, and fully corresponding to our initial intuition of logical generality. Given a non-closed table, the constants c_1, c_2,... occurring in it are identified with the numbers 1, 2,... We thus obtain an arithmetic interpretation upon defining, for every r-ary relation symbol R, an arithmetical relation R' as follows: R' holds exactly for those r-tuples $<n_1, ..., n_r>$ such that $Rn_1, ..., n_r$ occurs in the **T** column. If we know what it means for a relation to hold or not to hold for numbers, we have an interpretation.

Before generalising this idea, let us pause for a moment to point out a lesson it teaches. In order to get the meaning out of symbols, we had to *see* the constants c_1, c_2,... *as* numbers. We wanted the interpretation, or the meaning, to be conveyed by formal sentences. The completeness theorem should teach - and in fact does teach - that we have a meaningful discourse as soon as we have a syntactically consistent set of formal sentences. The completeness theorem is *Saccheri vindicatus*, the very Saccheri who developed his geometry with the negation of the fifth postulate, formally looking for a contradiction he could not find, and all the time he was building in fact the description of a non-Euclidean geometry. His description was indeed a geometrical one, not an arithmetical one; people who later read it as meaningful didn't have to pass through the arithmetical interpretation; they would see points and lines, or rather see circles were the description talked of lines. In order that our syntactically consistent phrases are able to generate a meaningful discourse about certain objects, we first have to specify the objects we are talking about, we must be able to see a model; in the same way, when we talk of the flying horse or of the chimera, we see them.

Third ending, where again the discourse disappears, leaving only the phantoms of the things we want to talk about

When we talk, we talk of something. We can also step aside and describe what we are talking about. This shift is a meta-theoretical move, though it can be hidden by the fact that the meta-language coincides with the language.

As already noted, when we are just talking, telling a story, the interpretation is implicit and participated in by the hearers. When we embark on talking of the interpretation, again there is an interpretation of what we are saying - and it usually refers to a richer language.

If I am telling that John loves Mary, and somebody ask me what I am talking of, it is likely that I begin to explain that I am talking of two persons, whose names are respectively “John” and “Mary”, and then go on with further details. The first answer isolates the type of objects, which in this case are human beings instead of, say, molecules; this answer is itself a discourse, referring to several types of entities, individuals and their names. If I say “John loves Mary”, I am talking of two individuals whose names are “John” and Mary”; if I say “I am talking of a man named ‘John’ who loves a woman named ‘Mary’”, I am talking of some individuals - myself included - of their names and of the attribution of names to individuals. All of these entities are (to be construed as) real, since one normally talks about them; some of them, e.g., names and attribution functions, have an abstract character. If pressed to go on, I will probably illustrate a particular relation between these individuals, a relation also possessing a name represented by the verb “to love”, and so on. This part of my speech will probably use the word “relation”, or something synonymous, so that new entities enter the stage, to constitute what is called the “ontology”.

In the statement “I am talking of two individuals named respectively ‘John’ and ‘Mary’ between which a relation holds such that...”, I am pointing out the existence of a link between individuals and relations, on the one hand, and linguistic elements on the other. I could also phrase my explanation in this way: “I am talking of two individuals named respectively ‘John’ and ‘Mary’ and between John and Mary there is a relation such that when such a relation holds...” and go on describing the effects of love; in this case John is the value of the assignment of a denotation to the name "John", in symbols, John = σ("John"). But such an assignment to linguistic items of parts of the non-linguistic reality does not exhaust the content of my statement; one must consider the implication that "John loves Mary" is true, either explicitly ("I am saying that... and 'John loves Mary' is true"), or implicitly, because "saying that..." is stronger that just "uttering", and it implicitly conveys the intended truth of what is said. All these levels and aspects, which will be carefully separated in our discussion below, are mixed in the natural language (and they are also mixed in the semantic meta-language, as soon as the latter is activated).

In short, through the analysis sketched above we arrive at a new notion of interpretation, pertaining to denotational semantics, or to set-theoretical semantics. An interpretation is now a pair formed by a structure and a function from the language to the structure. A structure is what is meant by such a term in mathematics: a non empty set with relations and functions. The set is the universe of discourse, having the conceptual riches and possibilities afforded by the functions and relations. On the other hand we have the linguistic elements, which come to be connected to the structure by means of a function, which is also called interpretation: proper names (which in formal languages are represented by individual constants) are made to correspond to, or are interpreted on special individuals - deserving a name; verbs (represented by relational symbols) are interpreted on relations; common names are usually interpreted on properties, i.e., one place relations; descriptions (represented by functional symbols) are interpreted on functions.

Now to be more precise; if the language has an alphabet containing a set of individual constants $\{c_i : i \in \mathrm{I}\}$, a set of relational symbols $\{\mathrm{R}_j\!: j \in \mathrm{J}\}$, each with an *a*-arity (number of arguments) n_j, a set of functional symbols $\{\mathrm{F}_h : h \in \mathrm{H})$, each with an *a*-arity n_h, then an interpretation adequate for the language is a structure

$$\mathbf{M} = < \mathrm{M}, \{\mathrm{R}_j^{\mathrm{M}}\}_{j \in \mathrm{J}}, \{\mathrm{F}_h^{\mathrm{M}}\}_{h \in \mathrm{H}}, \{c_i^{\mathrm{M}}\}_{i \in \mathrm{I}} >$$

where M is a non empty set, $\mathrm{R}_j^{\mathrm{M}}$ is an n_j-ary relation on M, $\mathrm{F}_h^{\mathrm{M}}$ is a function from M^{n_h} in M, and c_i^{M} is an element of M.

This is the definition of structure as given in mathematics; the usual numerical systems in algebra are structures according to such definition; for example $\mathbf{Z} = <Z, +, \times, 0, 1>$ is a presentation of the integers as a structure with two binary operations and two special elements, while $<Z, <, +, \times, 0, 1>$ is another presentation of the integers as an algebraic ordered structure. Teachers have learnt at college the structuralist language, but they tend to forget it when teaching; they haven't obviously seen the reason for it. The presence of 0 and 1 in the presentation of the structure does not mean that there are only these elements in the universe; rather, these are the only elements with a proper name in the language. According to this presentation, the element "two" doesn't have a proper name, but only the indirect description $1 + 1$. In the usual expositions one does not distinguish between relations and relational symbols, functions and functional symbols, special elements and constants; it would be more correct to keep the distinction active; thus if 0, 1, + and so on are symbols of the language, then the structure should be more correctly presented as $<Z, +^Z, \times^Z, 0^Z, 1^Z>$. Of course, 2 may become a new constant symbol added to the alphabet to abbreviate the composite term $1 + 1$, whose value in the structure is the element $1^Z +^Z 1^Z$.

A good reason to avoid superscripts is of course that when considering the ring structure of the rationals, $\mathbf{Q} = <Q, +, \times, 0, 1>$, one never introduces new symbols for zero, one, addition and multiplication. Accordingly, instead of viewing the integers as embedded or embeddable in the rationals, one sees them as a subring of the rationals. Dropping the superscripts is here tantamount to the transformation of the embedding in the identity function.

But another reason is that one tends to identify the name and the thing named; now this is strange, since nobody would identify John with the name of John; the fact is that in the natural language there is no straightforward way of expressing the difference. Rather recently, logicians have introduced quotation marks, but they are uncomfortable and silent - that's why we are systematically ignoring them in this paper, even if it looks kind of snobbish. One doesn't talk of names, except when saying, e.g., that one prefers "Mary" to "Ann", but in that case one usually takes care that the context makes the quotation marks clearly audible. This amounts to using the construct "the name of".

As already remarked, $<Z, +^Z, \times^Z, 0^Z, 1^Z>$ and $<Q, +^Q, \times^Q, 0^Q, 1^Q>$ are different structures for the same language; even if one identifies the first with a substructure of the second, one must still maintain the difference between the entities in these structures and their linguistic counterparts. So the usual symbols (of the pseudo-natural language) are reserved for the entities, and new signs are introduced for the linguistic counterparts; thus, e.g., if 0 and 1 are the elements, then **0** and **1** are the corresponding constants.

In mathematics, relations and functions are sets of ordered pairs, or n-tuples. The definition conforms to the extensional notion of relation. There would be a lot to say on the matter, but suffice to remark that a different notion - the so called intensional one, which was harboured by the first modern logicians - would not make much difference with respect to the problems we are discussing. It is commonly accepted that to any relation (whatever that be) there corresponds its extension - the set of n-tuples in the relation. Although not necessarily equivalent, the extension can to all practical purposes replace the relation in most considerations. Try to spot a place in the following pages where this ceases to be true.

An interpretation does not however ends with the presentation of a structure for the language; this is just the interpretation of the alphabet; then one has to interpret, or to transport the interpretation to the more complex parts of the language, such as terms, formulae, sentences. We say for example that <Z, +, ×, 0, 1> is a ring, while <Q, +, ×, 0, 1> can be considered both a ring and a field. These further characterisations of structures of the same type (for the same language) depend on the sentences of the language that are true or false in the structure.

Every closed term will denote an element in the same sense in which c denotes c^M; every sentence will be true or false according to a suitable inductive definition; the inductive step compels us to take into consideration also formulae with free variables (since for example descent from $\exists x\, A(x)$ to $A(x)$ leads from a sentence to a formula). An exception is represented by certain special interpretations to be discussed later on. This technical point drives us to define a more general notion than truth, i.e., satisfaction of a formula by elements of the structure. The notion of truth for sentences will be obtained from the particular case of formulae with no free variables. The technical point is actually also a practical one, corresponding to a common informal terminology:

to decide whether an equation as $x^2 - 2 = 0$ has solutions in <Q, +, ×, 0, 1>, that is whether $\exists x\,(x^2 - 2 = 0)$ is true in <Q, +, ×, 0, 1>, one looks for elements which satisfy the formula $x^2 - 2 = 0$.

A formula with free variables cannot have a definite truth value, because variables have a generic, indeterminate denotation, while each constant c denotes just and always the fixed element c^M. Strictly speaking, according to our standing definition of interpretation, variables have been left without any denotation at all; in order to give variables a definite (temporary) denotation, one has to fix the elements they (temporarily) denote. To this purpose, a function σ from the set of individual variables into the universe M is called an *assignment* in **M**. Any assignment σ can be uniquely extended to all terms - including nonclosed terms - in the natural way, by inductively defining the element t^σ assigned by σ to t in **M** as follows:

$x^\sigma = \sigma(x)$
$c^\sigma = c^M$
$(Ft_1...t_n)^\sigma = F^M(t_1{}^\sigma,..., t_n{}^\sigma)$.

The value t^σ depends only on the variables occurring in t; if t is closed, all assignments give t the same value, which is written as t^M. Let x be any variable; then an assignment which differs from σ at most for the value assigned to x is called an x-variant of σ.

At last we can define the notion of satisfaction; one talks of satisfaction in **M** of a formula A by (or under) an assignment σ, a relation usually denoted by **M**, $\sigma \models$ A. Let us suppose for simplicity that the language has only one binary relational symbol R; the definition of satisfaction for atomic formulae is then

$$\mathbf{M}, \sigma \models Rt_1t_2 \quad \text{if and only if} \quad \langle t_1{}^\sigma, t_2{}^\sigma\rangle \in R^M;$$

then by induction,

M, $\sigma \models A \wedge B$ if and only if **M**, $\sigma \models$ A and **M**, $\sigma \models$ B,
M, $\sigma \models A \vee B$ if and only if **M**, $\sigma \models$ A or **M**, $\sigma \models$ B,
M, $\sigma \models \neg A$ if and only if it is not the case that **M**, $\sigma \models$ A (written also **M**, $\sigma \not\models$ A) ,
M, $\sigma \models A \rightarrow B$ if and only if it is not the case that **M**, $\sigma \models$ A and **M**, $\sigma \not\models$ B,

$\mathbf{M}, \sigma \models \forall x\, A(x)$ if and only if $\mathbf{M}, \sigma' \models A$ for all x-variants σ' of σ,
$\mathbf{M}, \sigma \models \exists x\, A(x)$ if and only if $\mathbf{M}, \sigma' \models A$ for at least an x-variant σ' of σ.

It is easy to see (with the help of one of those tiresome proofs mentioned at the beginning) that the satisfaction of a formula A by an assignment σ depends only on the values assigned by σ to the free variables of A. If A is a sentence, then either all assignments satisfy A or none; in the first case we say that A is true in $\mathbf{M}$, and we write as anticipated $\mathbf{M} \models A$.

A special interest have those structures whose universe is in a one to one correspondence with the closed terms (briefly: each element has exactly one name) and in which the functions F^M are defined in such a way that $F^M(t_1{}^M,..., t_n{}^M) = (Ft_1...t_n)^M$ (the sophisticated reader will certainly perceive here some categorical niceties). These interpretations are called Skolem-Herbrand interpretations. In Skolem-Herbrand interpretations one can work only with sentences, by defining truth directly for sentences in this way: for atomic sentences, $\mathbf{M} \models Rt_1t_2$ if and only if $<t_1{}^M, t_2{}^M> \in R^M$; clauses for connectives are the same; then $\mathbf{M} \models \forall x\, A(x)$ if and only if for all closed terms t, $\mathbf{M} \models A[x/t]$, and $\mathbf{M} \models \exists x\, A(x)$ if and only if there is at least a closed term t for which $\mathbf{M} \models A[x/t]$.

A simple example of a Skolem-Herbrand structure is $<N, S^N, 0^N>$, where N is the set of natural numbers, S^N is the successor function, 0^N the only element which is not a successor, $0^N \notin \mathrm{range}(S^N)$; the elements $n \in N$ are in a one to one correspondence with the numerals $\mathbf{n}$ of the language of the structure, i.e., with terms of the form S ... S0, with n occurrences of S.

Instead of considering a set in one to one correspondence with the closed terms, one can take as universe the very set of closed terms. These are in fact the structures that are called Skolem-Herbrand structures, with functions defined by $F^M(t_1,..., t_n) = Ft_1... t_n$. The simplification is a common one in mathematics - any set is just the same as any other isomorphic to it. Then each t^M coincides with t, and one says that every closed term denotes itself. In Skolem-Herbrand structures, assignments have now closed terms as values, and an interesting consequence is that $\mathbf{M}, \sigma \models A$ if and only if $\mathbf{M} \models A[x/\sigma(x)]$.

Here one can see a *rationale* for the students' way of talking, when they claim to (be able to) substitute a number for a variable in an equation; they do not grasp the possibility of talking of the satisfaction of an equation by a number without literally substitute the number in the equation (which is impossible, because the result of such a substitution would no longer be a formula). Satisfaction is an abstract notion, in contrast with the material operation of writing. Students tend to talk as if they were only handling Skolem-Herbrand structures (which is not really so; even when every element has a name, the correspondence need not be one to one). For them, every number has a name, which is certainly false, but the truth is difficult to accept; if a number does have a name, or a description, how do we know its existence, and how can we say anything about it? Students are natural born Skolem-Herbrand structuralists; for them quantifier elimination, reduction to propositional logic and the like come entirely natural. If every number has a name, then checking satisfaction amounts to performing certain substitutions, and then applying standard normal form techniques and basic identities. The quite reasonable simplification afforded by Skolem-Herbrand structures can be proved to give sufficient elbow room for semantics, but it poses new disturbing problems.

Let us go back to the set-theoretical definition of interpretation, to see what it means for the completeness theorem. Starting from a consistent set of sentences, we had constructed a nonclosed table, and we were looking for an interpretation. We had also seen that starting with a spoken language with which we communicate meanings, it is also possible - by reflecting on the language itself - to describe the kind of things we are talking about, i.e., to define the universe, the associated structure, and a mapping of the language in the structure. This mapping is a counterpart of our intended meaning. By contrast, in the case of set-theoretical structures we have to reverse the direction; we start with a morass of artificial symbols which do not speak of anything; our symbols have only a grammatical correct form, but no meaning or reference; the latter must be imposed on them *ex novo*. The structure must be created from nothing, not from meanings, as was the case with natural language. No meaningful discourse is acting here as a mediator, but there is to be a direct link between formal expressions and structure. The latter might be the reference, but meaning is still missing. Even the substitutional version of interpretation is of no help - rather, as already noted, we would be led to an unsatisfactory situation: as a matter of fact, when pressed for precision, the substitutional arithmetic version tended to be formally the same as the definition of an arithmetic structure.

So again we lean on the usual strategy; *if* we had a true meaningful (but precise) language, we could have an associated structure, and in it some sentences would be true and some sentences would be false, according to the definition of satisfaction in a structure. Now the closure properties of the sets of sentences in a table are quite similar to the properties of two sets of sentences respectively coinciding with the sets of true and false sentences in a given structure. We must go in the opposite direction, defining the structure in such a way that the sets of its true and false sentences in it actually coincide with the sets we have in the table. We define a structure, and then we show that its true and false sentences are exactly those in the table. We can define a structure directly, without passing thorough a meaningful discourse, because we already have the sets of true and false sentences notably the atomic sentences, which make it possible for us to define functions and relations.

First, the universe: the closure conditions on the sets **T** and **F** of the table suggest a Skolem-Herbrand interpretation; this is what we will actually be doing, also for economy reasons, putting in the strictly minimum amount of elements (we are supposing that there is no equality sign in the language; thus there are no equality axioms to impose identifications of different names). So let M be the set of closed terms occurring in the table, with the natural definition of functions. As for relations, supposing for simplicity that the language only consists of one relational binary symbol R, we define

$$<t_1, t_2> \in R^M \quad \text{if and only if} \quad Rt_1t_2 \quad \text{is in column } \mathbf{T}.$$

For complex sentences, we straightforwardly apply the clauses of the truth definition. Having thus defined the interpretation, one may reasonably ask which sentences are true. Not surprisingly, it turns out that the interpretation respects the table, whence the conclusion follows: the original consistent sentences has a model.

The proof is easy, by induction on the complexity of sentences; if $\neg B$ is in **T**, then B is in **F**; by induction hypothesis, $\mathbf{M} \models B$ if and only if $B \in \mathbf{T}$, whence $\mathbf{M} \not\models B$, that is $\mathbf{M} \models \neg B$. If $B \wedge C$ is in **T**, then also B and C are in **T**, and by induction hypothesis **M** $\models$ B and **M** $\models$ C, whence **M** $\models$ B $\wedge$ C. The other cases are similar way. Since we are in a Skolem-Herbrand interpretation, for quantifiers it is enough to observe that if, say, $\forall x$ B is in **T**, then for every closed term t (viz., for every element of the structure) B$[x/t]$ is in **T**, and by induction hypothesis **M** $\models$ B$[x/t]$; it follows that for every σ, **M** $\models$ B$[x/\sigma(x)]$, i.e., **M**, σ $\models$ B(x), which is the same as **M** $\models$ $\forall x$ B. Same argument for the existential quantifier.

Now that the proof is complete, it wouldn't be honest not to ask whether it isn't just a trick. On the one hand, the interpretation whose search has given us so much trouble is there, as a well defined mathematical structure. On the other hand, it is by no means clear whether we have made any progress: the interpretation doesn't encode any more information than is contained in the two sets of sentences **T** and **F**, satisfying certain closure properties that - only *a posteriori* - will be gratified with such names as satisfaction, or truth. What seemed to be the more interesting and promising notion of interpretation (once we agree that the substitutional notion is elusive and unstable as a carrier of meaning) is in a sense squeezed and thorn apart between structures and formal sentences; it impinges both on syntax and on semantics.

A seemingly idle digression on algorithms and their correctness and completeness, where one is in trouble to distinguish which is which, i.e., which is the syntax and which is the semantics

Notice the paradoxical fate of semantic tableaux, and their metamorphosis. When they first made their appearance, they were a method to manipulate formulae, an effective mechanical method whose rules were form-driven; it was legitimate to think of it as a syntactic method. However, one can use it as well as a self contained method, whose output answers a semantic question: depending to the initial data, the method allows us to determine whether a sentence is a logical consequence of another, whether a sentence is satisfiable, or whether it is logically true. When we use the method as we did in the proof of the completeness theorem for natural deduction, semantic tableaux loose their independent meaning: they reduce to an auxiliary pre-processing technique yielding suitably closed sets of sentences. But, with a final twist, tableaux take their revenge: if those suitably closes sets are all that matters, this means that the method is by itself sufficient for the original purpose, it is the beginning and the end of the story, as it embodies both semantic answers and syntactic effectiveness. The fact that such a method must exist could be the corollary of the completeness theorem - the provably true coincidence of the two notions. But the (possibility of proving the) theorem follows, in turn, from the invention an application of such an encompassing method as tableaux; if we had used some other technique, e.g., maximally consistent sets, the remark would still hold, up to questions of effectiveness of the proof.

At the beginning of our discussion we observed that the tableaux method could be said to be complete with respect to search, in that all roads were essentially explored during the development of a table. This search completeness of the procedure is a counterpart of the correctness of tableaux as a logical method - indeed, the property of search completeness has the following effect: whenever a table initialised with A in **T** and B in **F** closes, an essentially syntactic event, then we have $\models A \rightarrow B$, an essentially semantic conclusion. Logical completeness is the converse fact: if we have $\models A \rightarrow B$, then the table must close. Both these logical properties depend on what is called correctness of the algorithm. Let us try to put some order in these notions.

When one looks for an algorithm to solve a problem, the problem involves an infinite class of instances of the same question; a property is given, and what is wanted is a method to decide, for each of infinitely many possible candidates, whether or not the property holds for the candidate. One can always reduce the problem to that of deciding membership in an infinite set - for, if there is a finite number of candidates it is always possible to enumerate them. A typical example, let us consider the primality property: here we have to decide whether a natural number is prime or composite; the associated set is the set of prime numbers; the set of candidates is the set of all natural numbers.

An algorithm is said to be correct if it always gives the right answer for any instance of the problem. As soon as an algorithm is presented, the first question is whether it is correct; the problem predates the algorithm, and the algorithm is invented in order to solve the problem. From the very beginning, problem and algorithm are kept quite distinct, and the problem (meaning) comes before the algorithm (syntax). Any algorithm is a syntactic method - recall the tableaux with anonymous labels I and II; its answers have to be right, that is, they have to respect the semantics. While the interest of the syntactic method lies in its effectiveness, its justification is that it gives right answers. This is not always evident: as a rule, the execution steps of the (efficient) algorithm do not parallel the problem's steps. A primality checking algorithm is not a straightforward application of the relevant arithmetical definitions, but relies on hidden and seemingly unrelated properties. If the algorithm were only an application of the definition, we wouldn't require a correctness proof. The proof often tries to show that the definition - also called specification - is satisfied.

When giving an algorithm to solve a problem, it is not enough to give the algorithm, we must carefully explain that what we are saying is right; there is here a meta-theoretical shift, which goes hand in hand with the objectification of discourse (the program) and its relationship to the domain problem. It is not easy to say why it should be so, or why has such duplication proved so effective: perhaps this has to do with the fact that the problems we are talking about are mathematical problems; they do not refer to real things, they are formulated in a language which seems meaningful (to those who understand the problem), but which on deeper analysis is just hanging on a *flatus vocis*; so one looks for equivalents, until one finds an equivalent version in another language that has the sturdiness of a machine.

With logic (and logics, in general) the situation is reversed, with respect to mathematical problem solving: there is a syntactic method in search of its semantic justification. It is true that the effective syntactic method of generation of chains of inferences is not an invented artefact; rather, it is modelled on our common practice of discourse - with a little polish. Still, what is lacking is the semantic counterpart. This is so because crude speech is not meant to solve any specific problem, but it is used to communicate; upon this mysterious function of speech there comes to be superimposed the distinction between problem and solution. In the case of pure reasoning, formalisation (that is, exclusive attention to syntactic constraints) transforms the discourse - its skeleton - in an algorithm. One may now wonder where is the problem to be solved. We will say more on this later. Suffice to say that logical reasoning is not involved with the solution of any specific problem, but with the solution of a whole class of formally similar problems; this follows from the established relationship between derivations and logical consequence.

A decision algorithm gives always one of two answers, YES or NO. Correctness can then be divided into two directions, correctness with respect to. YES and correctness with respect to. NO. These two can be called correctness and completeness, respectively. The reason for the terminology is as follows. To be specific, assume the problem to be represented by the set X, and suppose the algorithm to be correct: thus, if the algorithm answers YES for input x, then $x \in X$; if the answer is NO then $x \notin X$; so the answer is YES if and only if $n \in X$. The equivalence has two directions, one leading from the answer YES to the membership in X, the other from membership in X to the answer YES; the former goes from the answer of the algorithm to the facts, the latter goes from the factual reality to the answer of the algorithm. In general, one can call correctness the direction from syntax to semantics, and completeness the opposite direction from semantics to the syntax.

Although the choice of terminology may seem arbitrary, it is justified when there is a certain asymmetry between YES-candidates and NO-candidates. Asymmetry can arise from several sources. It may happen that there are sharp procedural differences in elaborating the answer for the two cases.

For instance, suppose the answer depends on searching over some infinite space, as it is often the case: then one type of answer could be the result of actual finding some witness - in which case, it is not just the answer that is positive but the very character of the method; the negative answer can come from (or with) a sort of proof that no witness was found. If the negative answer includes a statement on the exhaustiveness of the search that was carried over, then this justifies the name of completeness. As a matter of fact, if the negative answer comes when one is looking for something and doesn't find it, then we can always suspect that the search method could be an incomplete one, and sometimes is unable to find what actually exists. The completeness property (of the algorithm) precisely denies such an allegation.

A different, though related, type of asymmetry arises when for YES-candidates the procedure always halts in a finite number of steps, while for NO-candidates may loop forever without terminating. (It is usual to distinguish between algorithms, which always terminate, and procedures, which do not necessarily terminate; we do not stick consistently to the prevailing terminology, since the non-terminating case for tableaux is subsumed, for logical reasons, under the existence of an interpretation.)

Still it is a fact that a certain arbitrariness holds in deciding which of the two directions should be called call correctness and which one completeness. The arbitrariness, however, is not wholly conventional, but depends on what we require from the algorithm. Great care must be put in distinguishing the purely syntactic answer of the algorithm from the information we want to decode from it; otherwise we risk endowing the algorithm with the capacity to create meaning. We do attribute meaning to the algorithm when (and only when) we decode its output.

Thus, for instance, in a Turing machine working with an alphabet {*, |}, and computing a numerical algorithm, as a prerequisite to recognising correctness of a result, we must first verify whether the output is a number; then, assuming the output to be written as a sequence of strokes, say |||...|, we still must transform this output in some base greater than 1, so as to be able to be read and communicate it. Was the meaning already there in |||...|, before decoding?

As another example, the output of an algorithm might well have, instead of the YES-NO answer, some other content, in which that YES-NO dichotomy is only implicit, or deducible from it; thus a primality algorithm could just give the list of all factors of the input *n*, and leave to us to conclude whether or not *n* is prime, by counting or inspecting these factors.

Now suppose we use the tableaux method as an algorithm to look for interpretations (satisfying the initial conditions); the question is whether there is an interpretation or not; if the table does not close, then, after some supplementary reasoning, we conclude that the interpretation exists. In this case, a syntactic phenomenon has a semantic counterpart; since the original problem to which the algorithm gives YES-answer is the existence problem for interpretations, we are led to say the algorithm is correct. *Vice versa*, as we have seen, if an interpretation exists then the answer is YES, and this direction from the problem to the YES-answer is the completeness of the algorithm.

Notice that the logical problem is an asymmetric one, with the asymmetry hidden by the passage to the limit; in the case where an interpretation is drawn from an infinite table, since the answer is never YES at any finite stage, we might be tempted to say that the method is incomplete; but in another sense, it behaves rightly so, in order not to say YES to the question about logical consequence.

Suppose in fact that we take tableaux as an algorithm to answer questions of the form $\models A \to B$. The procedure is exactly the same. But now the proof that *if* the table closes *then* (there in no interpretation, whence) $\models A \to B$, goes in the syntax-to-semantics direction. We are thus led to regard this implication as expressing the correctness of the algorithm - with the converse implication expressing its completeness. You might object that this is a case of much ado about nothing; it is only the trick of switching YES and NO labels. But deeper issues are involved.

We should not confuse semantics for logical languages with semantics in the sense of the problem the algorithm is intended to solve. If we expect the algorithm to produce a closed table, whenever our problem $\models A \to B$ has a positive solution, then our original aim is to find a closed table. The same aim confronts our algorithm, precisely as other calculi aim at finding a derivation. The problem for which the algorithm has been devised, its semantics, is now the search for a derivation. Are we looking for derivations or for interpretations?

The answer of the algorithm is meant to give us insights into logical semantics; but the decoding of the answer depends on what we mean by semantics; the latter is influenced and refined, among other things, by our use of tableaux. Such considerations are another way of signifying our ambiguous feelings towards the syntactic/semantic character of the method of semantic tableaux, as well as towards any other logical calculus.

Conclusion, where again everything gets mixed up, because if you persevere with splitting hairs you end up with nothing in your hands

The interpretations given by the completeness theorem (in the form of the model existence theorem) are completely determined by the set of linguistic specifications arising from the development of the table; the latter provides us with sufficient information to single out the domain and the structure of the discourse. Skolem-Herbrand structures go a step further in this direction, since they are a way of saying that everything arises from and ends in the language. This does not mean - as in the old argument with Bishop Berkeley - that solid things are not there as obstacles to crash against. Rather, this means that words cannot prove that there are things that aren't words; briefly, uttering words produces words, not things. Recall the primitive wisdom of students who identify numbers and numerals; in our sophistication we try to convince them that the two are different, as they are, but only to conclude that the distinction lives up in the lightness of naming: numbers are different from numerals, but they are numerals in another notation, as far as we can name or exhibit them.

Which of the two is better suited to clarifying the nature of mathematics, the set-theoretical structure or the syntactic consistent set of sentences?

The interpretation-as-structure idea arose from a meta-linguistic reflection on what we are talking about when talking in a meaningful language; in the completeness proof the meaningful language is still missing, there is only the universe of discourse and the form of the discourse (the form we would also have if we had a meaningful discourse). Words do not create things; still, they are powerful enough to create meanings, as it is clear with the interpretation-as-substitution version, which however is a point of unstable equilibrium which still evades us. The meaning associated to a model must be the meaning of a whole class of formally similar discourses, or what is common to all the meanings of all interpretations. The set-theoretical structure takes care of that.

Warning The model existence version of the completeness theorem is somewhat deceptive, in that it suggests that a consistent set of sentences has just one interpretation (up to inessential variations, of a rather technical mathematical nature, such as cardinality - as if the universe for “John loves Mary” could be indifferently chosen to be the set of their acquaintances, or their fellow-citizens, or humankind). To avoid such a misunderstanding, attention must be paid to the special nature of set-theoretical language.

Here comes the “third man” at last, the meaningful language - the set-theoretical language (relativized to the structure at hand). Set-theoretical language has replaced in our times the arithmetic language of the early substitutional completeness. It is a great improvement from many points of view: sets are a more logical notion than numbers, and they are also well suited as an interface with common discourse.

Let us go back to John and Mary. The structure **M** serving as background for their love can be thought of as formed by a certain set of persons (living humankind, or the persons we are acquainted with, or other); there is a relation corresponding to the verb “to love”, a relation we can suppose to be reflexive, not symmetric and not transitive, and that we have specified rather well in psychological or physiological terms in order to be able to decide, for every pair of individuals, whether one element of the pair loves the other or not. We call this relation *love*, and we write $c_1 \mathit{love} c_2$ (instead of $<c_1, c_2> \in \mathit{love}$) if the pair $<c_1, c_2>$ belongs to the relation. There are possibly also other relations and functions (e.g., for relatives and friends), but that depends on the kind of gossip we want to spread on our characters. For the interpretation of the statement “John loves Mary” we say that there are two individuals c_1 and c_2 whose name are “John” and “Mary” and such that “$c_1 \mathit{love} c_2$”. There isn’t a great difference between “John loves Mary” and “$c_1 \mathit{love} c_2$” - after stipulating that “John”$^M = c_1$ and “Mary”$^M = c_2$; this is even more evident if we consider more complex statements. “There is someone whom Mary loves and whom John, who loves Mary, does not love” becomes “there is an x such that $c_2 \mathit{love} x$ and $c_1 \mathit{love} c_2$ and not $c_1 \mathit{love} x$”. Once the substitution of the atomic statements has been done, the form is the same, as is well known.

Here we have adopted a notation different from the usual set-theoretical one, and more akin to the spoken one (and also to informal mathematics). Let us now consider another example from arithmetic, where the set-theoretical language is preserved. Let S be the functional symbol for successor, and S^N be the successor function in the structure $\mathbf{N} = \langle N, S^N, 0^N \rangle$ of the natural numbers; let further $\mathbf{0}$ be the constant symbol for the first element 0^N, and let E be the equality symbol, to be interpreted by the relation =. Then the following two statements are equivalent:

(i) $\mathbf{N} \models \forall x \exists y (\neg Ex\mathbf{0} \rightarrow Ex\, Sy)$

(ii) for all a in $\mathbf{N}$ there is a b in $\mathbf{N}$ such that if $a \neq 0^N$ then $a = S^N(b)$.

Suppose we use special symbols for quantifiers also in the meta-language, taking care that these symbols are different from those in the language, for example $\wedge$ and $\vee$ for the universal and for the existential quantifier, respectively. Further, let us use the symbol $\Rightarrow$ to denote implication in the meta-language. Then we have the equivalent statements:

(i) $\mathbf{N} \models \forall x \exists y (\neg Ex\mathbf{0} \rightarrow Ex\, Sy)$

(ii') $\wedge \mathrm{a} \in \mathbf{N} \ \vee \mathrm{b} \in \mathbf{N} \ (a \neq 0^N \Rightarrow a = S^N(b))$.

We are now left with an evident identity of form between the arithmetic formula and its interpretation (the identity of form would be even more evident if we had used the same sign for E and =, as usual).

Here are some consequences of this identity. The first is that the meaningful language is there - the language with which we talk of what is happening in the structure. Of structures, of their elements and relations we are talking. Linguistic elements are also found on our sheet of paper: "*love*" is the name of a relation, S^N is the name of a function, and c_1 and c_2 still are names - though not to be found in the registers - which we use to describe the structures in the meta-language, not in the language. Between a language and the structures in which it is interpreted there is always the filter of another language, which is a meaningful language, conveying a direct meaning. To explain the above transformation (i)-(ii') we use this latter language by showing, even without mentioning, the new statements. This is the semantic meta-language, by means of which we not only describe the structures and the interpretation functions, but we also translate formulae into statements - of the same form. This is done by passing through the definition of $\models$, expounding it along its the inductive clauses. The result is that of saying true and meaningful things by translating formulae into statements talking of structures.

The second remark, already mentioned, but worth repeating, is that formulae such as (i) and their associated meaningful statements (ii') have the same form, up to inessential variations; certainly, the grammatical form is the same.

Again, we are left with a substitutional interpretation, which is made possible by a uniform and precise language, the language of structures. Since the latter is rightly considered a mathematical language, it follows that, in order to explain the mathematical language, passing through the formalisation, we end up with a mathematical informal language. We might embark on formalizing the latter, too, as it is often done, but to no evident purpose.

The proof that the language of structures is an informal one is given by the fact that in front of (ii') there is no $\models$ (no assertion sign, as Frege would have it). (i) is a meta-theoretical statement about the formal statement $\forall x \exists y (\neg \mathrm{E} x 0 \rightarrow \mathrm{E} x \mathrm{S} y)$, and (ii') is the same expressed in the meta-language.

The conclusion is that the meaning of a sentence is the sentence itself, just as at the global level, an interpretation of a set of sentences is the set of sentences itself. Saying that the meaning of a sentence coincides with the sentence is equivalent to saying that meaning is undefinable: whenever we try to define the meaning, also for (fragments of) the natural language, and we go through formalisation - possibly thinking of it simply as regimentation - we are forced to conclude that we need another meaningful language, one that is able to talk of more abstract entities than the original language. If we do not want to define the meaning, but only to show it, or to communicate it, then it is enough to talk.

Let us again consider the equivalence between

(i) $\quad \mathbf{N} \models \forall x \exists y (\neg \mathrm{E} x 0 \rightarrow \mathrm{E} x \mathrm{S} y)$

(ii') $\quad \bigwedge a \in \mathbf{N} \ \bigvee b \in \mathbf{N} \ (a \neq 0^{\mathrm{N}} \Rightarrow a = \mathrm{S}^{\mathrm{N}}(b))$.

On the one hand we have a meaningful statement (ii'), and on other hand, a statement (i) saying that the formal structure of (ii'), its name, is true in an interpretation. As remarked above, in the natural language one seldom uses the term “true”: one usually prefers uttering statements, as is done in (ii'), avoiding meta-statements on the truth of statements. Nevertheless, (i) and (ii') are equivalent; the corresponding phenomenon for our formal languages and calculi is that the syntactic rules reproduce the relation of logical consequence.

Upon reading the above equivalence in the suitable direction we also learn that it is possible to think in formal terms, i.e., by manipulating uninterpreted symbols and building derivations. However, also in this case, it is as if we were talking of something; not only machines can make purely formal manipulations, though only machines can make only formal manipulations.

The things one is talking about when doing formal manipulations are the structures. The language of structures is used according to the same rules one finds in the formal calculi; the language of structures can coincide with the language it interprets. Skolem-Herbrand structures are the medium and fusion between syntactic form and structures; here all distinctions disappear, and also the meta-language is eliminated. So is semantics.

New conclusion, where one sees that the findings of sophisticated logical theorems correspond to the naïve experience: the discovery of hot water

The meaningful statement (ii') corresponding to the formal sentence (i) is a mere typographical variant of it. Therefore, it comes as no surprise that the student - however oriented to became a mathematician - and the accomplished mathematician he has become, both are unwilling to duplicate the language.

We have already observed that students are by nature Skolem-Herbrand animals. It may be useful to recall the experience gained by extending number systems, e.g., passing from natural numbers to integers.

We are given an operation $n - m$, the difference, which is recursively defined as the iterated predecessor, though it is only defined, in a natural way, for n greater than or equal to m. Then one intuitively introduces the idea of counting backwards from 0, as when one is going down a stair, so that $-n$ could be the number from which one has to count n to reach 0; stated otherwise, $n + (-n) = 0$. Then $n - m$ is defined also for $n < m$, by stipulating that it is equal to $-(m - n)$. But this definition is merely an operative one; one always tries to reduce oneself to the difference, in the old sense, where 1 is repeatedly subtracted from a larger number; since $n - m$ must be the same as $(n-1) - (m-1)$, one obtains $n - m = 0 - (m - n)$. For example 3 - 4 equals 0 - 1, written -1.

From the set-theoretical viewpoint, the integers are defined as the equivalence classes of pairs $<p, q>$ with respect to the following equivalence relation: $<p, q> \sim <r, s>$ if and only if $p + s = q + r$. The addition operation is defined pointwise: $<p, q> +^{Z} <r, s> = <p + r, q + s>$. There is a sort of isomorphism between the set-theoretical construction and the algebraic formulae that have inspired the construction: $<3, 0> +^{Z} <0, 4> = <3, 4> = <0, 1>$ is the structural fact that allows us to say that the sentence $3 + (-4) = 3 - 4 = -1$ comes true. Our guide is precisely given by the set of sentences (equations) we want to be true.

As noted above, more complicated is the situation with rational numbers, where - in our teaching experience - we have the superimposition of several historical phases: fractions as (cake slicing) operators, numbers resulting from the division operation, and - finally - numbers in decimal notation. What such numbers denote (but themselves) is hard to say: one cannot say that each of them denotes a fraction, because one first needs some simple fractions, with numerator less than denominator, to justify the notation.

This confusing plurality of notations invites to reflect a bit more on the notion of Skolem-Herbrand structure. Up to now we have been a bit sloppy; to say that each term denotes itself has been a means to say that it doesn't denote anything non-linguistic; in particular, different terms have different denotations. Such structures are rare. As soon as the arithmetic language includes other functional symbols, besides the successor, a Skolem-Herbrand structure for the language cannot be isomorphic to the natural numbers. The same applies to integers and rationals. So, according to the strict sense of the definition, it is not right to say that students are *naturaliter* Skolem-Herbrand (SH) animals. Still, there is an intuition here to be cultivated; one wonders whether it wouldn't be preferable to encourage students towards the SH disposition; instead of imposing platonistic number systems why not teach them to directly work on notation systems, in a steady, free, easy way, that also appears to be more correct from the formal viewpoint.

Remarkably enough, what students actually do in high school (in some countries) is called algebra of expressions, rather than numerical calculus or algebra of numbers. The ordered chain of integers stands alone in one compartment. In another compartment one finds a profusion of expressions, sometimes to be evaluated to numbers, sometimes to be transformed into other expressions, without any explicit purpose. This is the compartment where student labour. The two terms 1/2 e 0.5 are conceived as different; they *transform* each into the other, but they express different concepts. The same with expressions; the sign = has to be read in the only sensible way, as a meta-theoretical sign of reduction of one term to another.

When decimal (including rational) numbers with infinitely many digits are introduced, a new phenomenon shows up, the limit, only to be generally disregarded. Curiously enough, students are quite happy to accept it: their intuition is that, if a generation principle looks reasonable, then it is convergent. This *credo* is at the basis of all infinite procedures; it is the same thing as the belief-principle that if a discourse looks consistent then it is so. The same bold idea was at the origin of modern abstract mathematics. At the beginning it took a great courage - witness Cantor's hesitation: he too began with (ordinals as) notations, such as ∞, $\infty + 1, \ldots$, but then he had the cheek to announce that behind the notations there lurked a thought, and the notation expressed the idea of infinite numbers. We have only our words to communicate, we have to do with them. Human thought translates in rules and axioms for symbols.

Final conclusion, where it is explained that hot water is better than cold water

We don't want to suggest that we gain nothing by the completeness theorem. To critically revise what one does unconsciously is always useful. To sum up, we have the confirmation that it is the same to consider a form empty of all content, or else all contents. Saying that something works not only for all interpretations, but, moreover, for whatever notion of interpretations, is a nice way of expressing logical generality. Everybody has the freedom to concoct concrete suitable definitions and applications of the interpretations.

But what of semantics, of this *doppelgänger*, or reflecting looking glass we counted so much on, and that seems to disappear with the proof? Our predicament is not without hope. A duplicate is always more than the single, and an isomorphism is not perfect identity, it is always an interpretation. To pass from

$$\forall x\, \exists y\, (\neg \mathrm{E} x 0 \rightarrow \mathrm{E} x \mathrm{S} y)$$

to

$$\bigwedge a \in \mathbf{N} \ \ \bigvee b \in \mathbf{N}\ (a \neq 0^{\mathrm{N}} \Rightarrow a = \mathrm{S}^{\mathrm{N}}(b))$$

means to change the language, and changing the language entails also a change of the theories one can refer to. The set-theoretical sentence $\bigwedge a \in \mathbf{N} \ \ \bigvee b \in \mathbf{N}\ (a \neq 0^{\mathrm{N}} \Rightarrow a = \mathrm{S}^{\mathrm{N}}(b))$ is just a fragment that can be inserted in a wider context without discontinuity or jumps; for instance, in the set-theoretic context S^{N} becomes an object of the same type as *a* and *b*. The set-theoretical language allows one to tell much more about S^{N} than one can tell in arithmetical language using the functional symbol S. Indeed, by definition of first-order logic, S can never be the subject of a proposition. The logical isomorphism is not also a grammatical isomorphism. The set-theoretical meta-language is just like the natural language, with its compression of types (which are kept apart in such specialistic languages as arithmetic), and thus offers amazing elbow room for richer thoughts. The fruitful development of set-theoretical semantics, and its usefulness for understanding the relationship between languages and interpretations - as witnessed by the achievements of model theory - is the practical proof that interpretations, even when squashed on syntax, have a multiplicative effect of meaning modulations.

The arithmetic meta-language, too, has the same function, using different means, and resorting to different solutions. Here, instead of blurring types, one makes great use of codings; instead of using arbitrary functions, one uses, e.g., recursive functions, together with numbers to code their programs. As a counterpart to these restrictions, one usually obtains results of greater complexity and finesse.

Another benefit of the completeness theorem is that, starting from the proof, one can take inspiration for generalisations and extensions. First of all, one can investigate much more complex situations, where a completeness theorem exists for different systems of rules, for example the rules of certain non-classical logics. Let us mention just one example: suppose we receive data from a network of observation points, whose nodes send us possibly conflicting results. Accordingly, contradictions may arise in our discourse based on these pieces of information. Nevertheless, these contradictions may look acceptable to us, in the sense that we can make a reason of them, and we don't want to eliminate them. In such a context it may be natural to attach more than two truth-values to a proposition P, say four, according to the following possible alternatives: "P was received", "notP was received", "both P and notP were received", "neither P nor notP were received". Rules complete wit respect to this kind of semantics have been found, and the proof of the completeness theorem is exactly the same as in the classical case - sentences being now partitioned in four classes. One may wonder now about the underlying intuition. In classical logic, if our data include both P and ¬P, then (we say that) we are not able to imagine anything, and no interpretation can possibly exist. Well, people who can reason with a four-valued logic are perfectly able to imagine an interpretation. What kind of intuition do they have?

It is only by looking at the language that we are able to understand and to explain what we are talking of; but a language can be looked at in several ways, corresponding to its many practical uses. When reflecting upon a language, we are burdened with a lot of knowledge, decisions and prejudgements which flow together in the imagination by which we form an interpretation, and fix the logic. A further complication is due to the fact that language is never looked at in private. Rather, we become interested in language when using it in a dialogue with other people: how do we know the logic of the opponent?

For some non-classical logics and semantics, the translation into the meta-language is not an isomorphism, not even a local one - although here also the proof of the completeness theorem generalises that of classical logic. Think for example of logics with possible world semantics, and their much more complex translations of sentences. Here the interpreting sentence contains fragments of the (translation of the) interpreted sentence, mixed with other information on the structure of the worlds of the model.

Finally, the equivalence between syntax and semantics suggests the possibility of reversing the procedure, thus taking seriously the symmetry property of equivalence. This is not only reasonable from the logical, but also from the psychological point of view. Starting, for historical reasons, with systems of rules that could be thought of as coming from an analysis of discursive practice, we embarked on the search of (notions of) interpretations that could justify the assumed logical character of the rules. Conversely, we might start from (discourses about) interpretations, and look for the rules.

This is common practice in mathematics: for simplicity, let us focus attention on single concepts instead of languages. Concepts are introduced via some kind of definition. To fix ideas, let us take the concept of "group": we must then ask about the existence of axioms (and possibly, the existence of logics) which are complete with respect to the concept. In this way, completeness amounts to searching, not for interpretations, but for the effectiveness of semantic notions. There are cases when this search is doomed to failure, i.e., no system of effective rules is available for a concept.

Let us stick to the groups and suppose for the moment that groups are our interpretations. Let us correspondingly relativize all semantic notions: thus, $\models A$ means that A is true in all groups. To avoid ambiguities, and to remind the restriction, let us write $\models_{gr} A$ for validity in all interpretations, given that our interpretations are the groups; more generally, for any given class *i* of interpretations, let us write $\models_{i} A$ for the corresponding notion of validity. We have already encountered one such restricted class, the Skolem-Herbrand structures, but that proved in a sense equivalent to the total class. Nowadays, all mathematical definitions can be cast in set-theoretical terms, every concept corresponds to a class of structures, and so there are many interesting classes of structures different from the total one. Every mathematical theory can be viewed as the study of a class of structures.

A class j of structures is said to be *elementary* if its definition is given by a set T of sentences of a first order language; stated otherwise, j is the class of models of T. The class of groups is elementary, because there is a set of sentences T such that for every structure **G** (equipped with a distinguished element, a unary operation, and a binary operation), **G** is a group if and only if **G** $\models$ T. With the help of T, the resulting new notion of logical validity is reducible to the usual one: as a matter of fact, for every A, $\models_{gr}$ A if and only if T $\models$ A if and only if $\models \wedge T \rightarrow A$. The set T is said to be a set of axioms of the theory of groups.

The strong completeness theorem includes the special theorem for all elementary classes i of structures: if T is a set of axioms for i, then $\vdash_i$ A if and only if T $\models$ A if and only if T $\vdash$ A, where $\vdash$ is derivability in a complete calculus.

A class of structures is said to be *axiomatizable* if there a set of rules (and axioms, to be thought of as rules with zero premises, in order not to have to treat them separately), not necessarily in a first order language, which is complete with respect to the class (i.e., is complete with respect to the new validity notion). The logic of a class of structures is the set of sentences true in all the structures of the class, or logically true with respect to this notion of interpretation. If a logic is not axiomatizable, then it is also improperly said to be incomplete (meaning that no set of rules for deriving its valid sentences is complete).

A more general, or precise, definition of axiomatizability for a particular notion of interpretation i is the following: i is axiomatizable if the associated notion of logical validity $\models_i$ is at least semidecidable. In this case, we also say that the logic associated or defined by the notion of interpretation i is axiomatizable, or that the logic defined by the semidecision method is complete with respect to i. Semidecidability can be obtained either by means of (search for) derivations in an extension of a usual logic, or else by any partial effective method.

The direction here is the opposite with respect to when we have a calculus in search of a semantics (as is also the case for computer languages and systems). We have argued that the search is, in a sense, hopeless. Here we have a semantics in search for a calculus - any calculus. This explicit formulation in terms of effective methods throws some further light on the issue of syntax and semantics, namely: the difference is not of substance, but of procedure. In one case there is the systematic exploration of all possibilities, in the other case, we are looking for (or, we are trying to construct) just one witness, namely a computation, or a proof. Only in this sense syntax and semantics remain essentially different.

Validity is a universal statement $\forall x\, (\mathrm{I}(x) \rightarrow \mathrm{V}(x, \mathrm{A}))$, to be read: for every interpretation, A is true in the interpretation. By contrast, syntactic validation is an existential statement $\exists x\, \mathrm{D}(x, \mathrm{A})$: there exists a positive answer for A (i.e., in the usual calculi, there exists a derivation of A). If both $\mathrm{D}(x, \mathrm{A})$ and $\mathrm{I}(x) \rightarrow \mathrm{V}(x, \mathrm{A})$ were decidable, the fact that $\exists x\, \mathrm{D}(x, \mathrm{A})$ and $\forall x\, (\mathrm{I}(x) \rightarrow \mathrm{V}(x, \mathrm{A}))$ are equivalent (when they are so, thanks to the completeness theorem), would imply that validity is Δ_1, that is decidable, recursive in a generalized sense. A Δ_1 set is recursive because both it and its complement are existentially expressible. To decide membership, one activates both search procedures for the set and for its complement, knowing that one of them will successfully find the answer. In case $\mathrm{I}(x) \rightarrow \mathrm{V}(x, \mathrm{A})$ is decidable, as for propositional logic, validity becomes decidable; but this is only possible if $\mathrm{V}(x, \mathrm{A})$ is decidable. In the case of propositional logic, interpretations are finite in number, so the universal quantifier is actually a fake, but in general it is not finiteness that is important. For predicative languages, it is $\mathrm{V}(x, \mathrm{A})$ that is not decidable, by Tarski's results, so validity is not Δ_1, notwithstanding the completeness theorem; thanks to it, validity is however at least (only) Σ_1, hence semidecidable. Ironically, it is the notion of truth that undermines the full-fledged import of the completeness theorem.

An important restriction of the usual notion of interpretation is represented by normal interpretations for languages with equality (i.e., structures where the interpretation of a special relational binary symbol E is the equality relation - the diagonal - in the structure). For normal interpretations, $\models_n$ is axiomatizable through the addition of the equality axioms, or by means of (some of them and) new rules such as paramodulation (in the context of resolution) or substitution and replacement (in the equational calculus). Hence equality is (said to be) axiomatizable. Notice that the proof of this fact is nontrivial, it is not just an adaptation of the usual completeness proof, because the class of normal interpretations is not elementary. As a matter of fact, the proof requires the algebraic construction of the quotient structure.

In order to ensure completeness, it is not enough to require that the notion of interpretation is a restriction of the classical one. Completeness depends on the definition of interpretation: a nonaxiomatizable mathematical concept is for example that of torsion groups; a notion for which there is no completeness theorem is the notorious second order logic. Here we find new mysteries, witnessing that, luckily, we have not reached the end of the story. It is only when a property does not always hold that it becomes interesting, and one understands it better through counterexamples.

Reopening of the case, with a supplementary investigation on the incompleteness of certain logics, in particular of the notorious second order one

To find examples of logics which are semantically defined through a class of structures, and which turn out to be incomplete, it is not necessary to change the language (as would be the case for second order logic). However, some preliminaries from the mathematics of infinity are required. First of all, it is useful to bear in mind some consequences of the completeness theorem, as is always the case when one wants to understand a mathematical result. One such consequence is the so called Löwenheim-Skolem theorem, LS for short: if a sentence has a model, it has also a finite or denumerable one (i.e., one whose universe has finite or denumerably infinite cardinality). In fact, if a sentence F has a model at all, then it is syntactically consistent; recalling now the model existence theorem, from its proof we get a finite or denumerable model **M** of F (recall that the universe of **M** is the set of closed terms built from the original constant and functional symbols of F, plus a possibly denumerable list of new constants). The same is true for a denumerable set of sentences.

The construction of the model existence theorem can be modified and made more abstract, to the point that “construction” becomes only a metaphor: we can add as many constants to the alphabet as we like, and with the construction of the proof we get a model of any cardinality (it is better to use a proof which lends itself easily to transfinite extensions, such as a proof using maximal consistent sets). In particular, there is a model which is exactly denumerable. There exist more refined versions of LS referring to the cardinality of the language, but we will not consider them; as a matter of fact, all our languages in this paper shall always be denumerable; they have a finite number of extra-logical symbols, with the (possible) exception of constants, and of course denumerably many variables. Although this is a restrictive assumption, let it pass as a technical tool. Moreover, unless explicitly stated, we never include equality in our languages - nor we consider the related subject matter of normal models.

LS has, in turn, the following consequence; let us suppose we define as interpretations only structures which are finite or denumerable - perhaps because we are not able to conceive of anything else, and talking of higher infinities is incomprehensible to us. But we live in a world where people - our colleagues - understand them, and perhaps our Best Friend assures us that they are useful. Then, we can re-interpret LS in the following way: if a sentence A is true in all our interpretations, then it is true in all the interpretations of our Best Friend; for otherwise, in one of them the negation ¬A would be true, whence, by his LS, A would be false in one of our interpretations. Now, by our Best Friend's completeness theorem, the sentence is derivable (in his favourite complete calculus). His derivation is also good for us. We conclude that our restriction to finite or denumerable structures is an admissible one, like the restriction to Skolem-Herbrand structures. We knew that already, since the completeness theorem is still valid for our interpretations by its very proof (while LS reduces to the following senseless tautology: if a sentence has a model then it has a model). But also our Best Friend must acknowledge it.

Eve if we only accept infinite denumerable interpretations, the above conclusion is still true, thanks to the refined version of LS yielding a truly denumerable structure, possibly by forced introduction of dummy constants. But suppose now that we decide to restrict to the class of finite structures. *Then we obtain an incomplete logic*: subtraction of a semidecidable set from a semidecidable set need not give a semidecidable set. In effect, if a sentence is true in all finite structures it is not necessarily true in all (the denumerable ones, which are sufficient to get all). Take for example the conjunction I of the following assertions about a binary relation R:

$$\forall x\, \neg Rxx, \quad \forall x \exists y Rxy, \quad \forall x,y,z\, (Rxy \wedge Ryz \rightarrow Rxz).$$

The interpretations are structures of the type **M** = <M, R^M>, where R^M is a binary relation on M. **M** is a model of I if every element of M is in the relation R^M with some other element, and there are no cycles, because of transitivity and antireflexivity. If **M** |= I, then M must be infinite, hence if M is finite then **M** |= ¬I. Assuming now the existence of at least an infinite set, it follows that there is a model of I. By the completeness theorem, the fact that ¬I is not universally valid with respect to the usual semantics means that it is not derivable in the usual calculi. This doesn't imply yet that the logic of finite structures is not axiomatizable: there might well exist other systems of rules, other partial effective methods yielding the sentences valid in all the finite structures.

However, this is not the case, and no such rules can exist. To see this, let us consider the notion of satisfiability in finite structures; fix a language, having only one binary relational symbol for simplicity. For every natural number n, up to isomorphism there are only a finite number m of structures for that language with universe of cardinality n. Specifically, m is the number of possible binary relations on a set with n elements. In order to verify whether a sentence B is satisfiable, i.e., whether it has a finite model, it is enough to inspect all possible interpretations, for increasing n. In each structure **E** of cardinality n, B is equivalent to a quantifier free sentence in some enriched language obtained by adding n new constant symbols for the elements of **E**, and by replacing the universal quantifier by an obvious conjunction indexed by the elements of **E**, and the existential quantifier by a disjunction. Verification of satisfiability of B in **E** reduces then to a propositional task that can be carried over in a finite number of steps. This procedure is a semidecision method for satisfiability in finite structures: if a sentence is satisfiable, a model will eventually be found; if not, the procedure in general will not halt.

So much for satisfiability; now, if also the validity problem were semidecidable, then we would actually have a decision method. To this purpose, we would exploit a well known technique of recursion theory: given a sentence A, we would start simultaneously and in parallel the semidecision method for validity on input A, and the semidecision method for satisfiability on input ¬A; since validity of A is equivalent to unsatisfiability of ¬A, precisely one of the two procedures will certainly halt; thus giving us the required decision method. (If the procedures are three-valued, with YES, NO and UNDEFINED values, both could halt, but then their answers would be consistent and both right, by the assumed correctness of the algorithms; we can take as answer that which comes first.) We know however that neither validity nor satisfiability in finite structures are decidable; therefore, there cannot be even a semidecision method for the logic of finite structures.

The above argument was implicit in what we said before on the Δ_1 character of logical validity (assuming completeness); in case of finite structures, the semantical definition of validity would have its universal quantifier ranging over (finite interpretations, hence, under suitable coding, over) natural numbers; it would be a genuine Π_1 arithmetical predicate and, if equivalent to a Σ_1 predicate given by a system of rules, it would be decidable.

To complete the proof and finally see that the notion of satisfiability in finite structures is undecidable, one can use the following argument, which we shall only sketch here, because a full proof would require too many details on arithmetic codings. One classical problem known to be undecidable is the halting problem for Turing machines. When a Turing machine **T** halts, **T** and its tape are in a particular configuration, only depending on the instructions of **T**, the input and the particular operating conventions for the halting of **T**. It turns out that the sequence of configurations leading **T** to a halting state can be fully described by a sentence H. The environment of **T**, its various tape configurations, the history of states of **T** leading to the present state of affairs, all of this amounts to assuming the existence of a certain a finite structure in which sentence H is satisfied. If there were a method to decide satisfiability in finite structures, then the method could be used to decide whether there is a halting configuration for a Turing machine over a certain input, thus making the halting problem decidable, a contradiction. This concludes our proof sketch that the logic of finite structures is incomplete.

Any infinite cardinality *k* can be used to impose the following kind of restriction on the class of admissible structures: take only structures of cardinality *k*. By an application of (the appropriate refinement of) LS, one sees that the resulting logic is axiomatizable, with the usual calculi. This fact has found some interesting applications in algebra, e.g., in the theory of fields. Axiomatizability of (the logics of) these classes does not depend on the classes being elementary, which in fact they are not. As another application of LS (or by an equivalent axiomatizability argument), for any two infinite cardinals *k* and *k'*, any denumerable set of sentences which is satisfied in the class of all structures of cardinality *k* is also satisfied in all structures of cardinality *k'*.

Also the class of all finite structures is not elementary, as a consequence of another corollary of the completeness theorem, the so called compactness theorem. The latter states that a set of sentences is satisfiable if and only if every finite subset of the set is satisfiable. To immediately derive the compactness theorem from the completeness theorem, it is sufficient to note that in any derivation there only occur a finite number of assumptions; this observation is known as the syntactic compactness theorem. There are also different direct proofs of the compactness theorem, e.g., proofs using algebraic methods. From compactness it follows that, if a sentence has arbitrarily large finite normal models, then it has also an infinite normal model. These results show that the completeness property has delicate and intricate consequences on definability.

For any fixed finite cardinal *n*, restricting to structures of cardinality *n* gives rise to complete logics - as implicit in the above mentioned reduction to propositional logic. It is only the concept of arbitrary unbounded finite set that produces an incomplete logic. An obviously related notion is that of natural number, the latter being a well known source of other kinds of incompleteness. Even assuming that the structure of natural numbers is a neat and uniquely specified concept (a neat interpretation of the arithmetic symbols $+$, $\times$), there is a strong noncompleteness result, which is different but related to the incompleteness phenomenon, and which can be stated as follows: the set of true sentences is not semidecidable, and is in fact not even arithmetical. Here we have no contradiction with substitutional arithmetic completeness, where one considers (not one but) many different interpretations on the universe of numbers; substitutional arithmetic completeness is more akin to the completeness of the logic obtained by restriction to denumerable structures.

Finally, to introduce second order logic we have to go back to the fundamentals, namely to the very definition of formal languages, and to the analysis of discourses from which alphabets and logics are extracted. In some statements, people use the so-called quantification of properties (or quantification of relations or functions). The simplest example that may cross one's mind is the ontological argument: the perfect being is a being that possesses all (good?) properties; admittedly, this is not to be heard every day discourse, still it is heard of. Formally, one can write $\forall X\, Xg$, where Xg is to be read "g has property X", just as in predicative languages one writes Pg for "g has property P".

Apart from theology, mathematics is the only other place one encounters the ontological proof - I mean, the quantification of properties. This is customary, for example, when giving such impredicative definitions as for natural numbers (every non empty set has a least element), or when stating the completeness of the reals (every bounded set has a sup). It is not by chance that these are the fundamental structures of classical mathematics. Instead of the quantification of properties, one talks here of the set of all subsets of an infinite set.

In the new definition of the language, we shall then introduce predicative, relational and functional variables, together with constants of the same type. The latter coincide with the old relational and functional symbols. In this new language, individuals and relations are thus treated more symmetrically than in first order languages. The old variables of first order languages will now be called individual variables. The new variables, for functions and relations, will have each a number attached, specifying the number of arguments; predicative variables are the same as one-place relational variables. Variables, when they are present, whatever their type, are there to be quantified; the sentence $\forall X\, Xg$ is now a correct one, once the definition of atomic formulae has been so extended as to include relational variables followed by the right number of terms - and the definition of terms now includes functional variables followed by the right number of terms.

When we presented the structures for predicative languages as

$$\mathbf{M} = < M, \{R_j^{M}\}_{j \in J}, \{F_h^{M}\}_{h \in H}, \{c_i^{M}\}_{i \in I} >,$$

the individual constants denoted some special elements deserving a proper name; now, more symmetrically, also the relational and functional constants will denote special relations (including properties) and functions, those exhibited in the presentation of the structure. To stress the symmetry, we will adopt italics for them as for individual constants; as before with individuals, these higher-order elements will not exhaust those to which reference can be made by the discourse, and which must be in the structure. In fact, the universe M is replaced by a set $M \cup U$ where U is a set of subsets, relations and functions on M. Thanks to the possibility of coding relations and functions by sets, and without bothering with details, from now on we will take for simplicity U to only consist of subsets of M. We shall thus avoid the recurring necessity of triple specification. A structure will then be presented as

$$\mathbf{M} = < M \cup U, \{R_j^{M}\}_{j \in J}, \{F_h^{M}\}_{h \in H}, \{c_i^{M}\}_{i \in I} >,$$

where $U \subseteq \wp(M)$, $\wp(M)$ = the power set of M, or the set of all subsets of M. On the alphabet's side, we will talk of set variables, also called predicative variables, though in some examples it will be handier to actually display relational variables. As for notation, we will use the predicative one, Xt, to avoid the membership sign (actually Xt could stand both for the result a silent operator of application, or for $t \in X$, for a relational symbol $\in$ linking terms t and X). Set variables are also called second order variables, the language is called second order language, and so will be called the logic (whatever it may be). The terminology gives some justification to the name of first order logic used for predicative logic, although, historically, second order and higher logics came first, tentatively, at least in modern times.

The definitions of satisfaction and truth are wholly analogous to the first order case, except for the following type condition to be satisfied by assignments: $\sigma(x) \in M$ for individual variables, and $\sigma(X) \in U$ for set variables; we denote the semantic relation of satisfaction by |= as before.

The first problem is which sets should be in U. Structures for which $U = \wp(M)$ are said to be *full*. One wonders whether U must satisfy some closure condition in order that **M** can be accepted as an interpretation. Some sets (and relations and functions) must certainly be in U, for example those which are singled out as denotation of symbols in the alphabet; each R^M and each F^M must belong to U, just as each c^M belongs to M; accordingly, M contained also all the elements t^M denoted by closed terms t. In first order languages, terms had the same logical type as variables: both terms and variables served the purpose of denoting individuals; variables had a generic denotation, so that terms could be substituted to variables. The problem of defining the right counterpart of terms in the new setting is relevant to the purpose of extending certain logical rules that in the first order case only involved variables. For example, the rule of elimination of the universal quantifier will allow the step from $\forall X\ A(X)$ to $A[X/Y]$ whenever Y is a variable free for X in A, and will certainly allow to proceed to $A[X/P]$, for every predicative constant P of the language; what else? Fixing the exact scope of this rule will critically depend on what we want or deem reasonable to obtain from it.

If, for example, we wish to be able to conclude that $\exists Y\ Yx$ follows from $\forall x\ (Xx \vee \neg Xx)$, that is, *tertium non datur* implies that the man without qualities does not exist, then we need among other things the step from $\neg Xx$ to $\exists Y\, Yx$, thus viewing $\neg X$ as a possible substitute of Y. In semantic terms, it is required that in any interpretation U be closed with respect to complement. In a similar way, we can easily convince ourselves that we want U to be closed with respect to other logical operations.

In this sense, formulae are seen to provide the appropriate second order counterpart of first order terms. On the other hand, substitution of second order variables with second order variables or constants, as in $A[X/Y]$ or $A[X/P]$, transforms atomic subformulae of A in different atomic subformulae; a natural generalisation is that of allowing replacement of atomic subformulae with non-atomic formulae. Accordingly, the definition of substitution must be so phrased as to make it legitimate to replace a variable X with a formula B in A. The result will be written as $A[X/B]$; details are omitted. The intended effect, which inspires the inductive definition of substitution, is that every subformula, of the context given by A, in which X occurs, say for instance Xt, must be transformed into the subformula $B[x/t]$. There are problems if X occurs in a term t which in turn occurs in a subformula like Xt; these problems will be taken care later on, when considering the comprehension principle. For the moment, we may limit ourselves to conceiving a descriptive operator transforming formulae into terms denoting sets.

Granted the technical feasibility of the above definition, we must decide which formulae to allow for instantiation in the elimination rule, and consider the impact of our decision on the closure properties of the family of subsets U in the interpretations. Any restricted choice, not involving the totality of formulae, is *a priori* admissible, though unnatural and almost always a bit twisted. For instance, one might quite reasonably decide to allow only first order formulae; after all, connectives and first order quantifiers correspond to such (historically [?]) natural set-theoretical operations as the Boolean operations and projection. On the other hand, the (re)proposition of first order machinery in the new context seems far-fetched, and probably suggested only by acquired habits; in the new context, individuals and sets should be treated on a par, and the first order criterion looks an unstable discriminant; it would probably be more reasonable to put a stronger restriction, such as limiting ourselves to quantifier free formulae; we will not pursue this endless dispute, and only consider substitutivity of all formulae.

The only real difference in the formulation of the deduction rules for second order language is to be found in the rule for the elimination of the universal quantifier (and in the dual one of the introduction of the existential quantifier). Instead of labouring on painful definitions of the substitution rule and descriptive operators, we shall formulate the rule in a simplified version, allowing only the step from $\forall X\, A(X)$ to $A[X/Y]$, with Y a variable free for X in A, and adding a logical axiom called the *comprehension principle* (or axiom) for set existence, as follows: for all formulae $A(x)$,

$$\exists X \forall x\, (Xx \leftrightarrow A(x)).$$

The calculus thus obtained from a first order complete calculus will be called second order calculus, and the notion of derivability will be denoted by $\vdash_2$.

The comprehension axiom is actually an axiom schema, one instance for each formula A. The above formula is not wholly precise, in that A might contain parameters for individuals or sets. We could write $A(x,...)$ or alternatively write it $A(x, \alpha)$ and write the axiom as

$$\forall\alpha \exists X \forall x\, (Xx \leftrightarrow A(x, \alpha))$$

where α denotes a finite list of individual and set-theoretical variables (different from X, for technical reasons).

From the semantic point of view, the effects of the comprehension principle amount to require that in every second order structure, U must contain all the definable subsets of M. To explain what a definable subset is, let $\sigma(x/a)$ denote, for an assignment σ and an element $a \in M$, the assignment that coincides with σ but for the assignment of a to x.. Then a subset X of M is definable is there exists a formula A with a free individual variable x and an assignment σ such that $X = \{ a : \mathbf{M}, \sigma(x/a) \models A \}$. We will not go into details about the effect of possible parameters (from $M \cup U$) in A, since this apparently innocuous question is rather complicated (as are all the inductive definitions involved above). Parameters will play an important role for our later discussion, but they will not belong to $M \cup U$, and will pose another kind of problems. Another way of stating the closure condition for U of course is to say that in the structure all instances of the comprehension principle must be true.

Completeness holds for the second order logical calculus with respect to this notion of interpretation and of satisfaction. The proof follows, *mutatis mutandis*, the same lines as for first order logic. For any restriction of (the formulae allowed in) the comprehension axiom and of the corresponding closure properties of interpretations, there is a similar completeness theorem. Where is the problem, then?

The problem is with the full structures; let us write $\models_2 A$ to say that sentence A is true in all the full second order structures. Let us call *full second order logic* the logic semantically defined by the class of full second order structures.

The difficulties are usually ascribed to the fact that the idea of *all* the subsets of M (hence, the idea of a full structure) is a vague one; Cantor's theorem tells us that the power set of a set has cardinality greater than the set, hence it is an higher infinity if the set is denumerable; there is no hope to cover it with definable subsets, even with parametrically definable sets, so long as the language is denumerable (the conclusion is perhaps a bit too hasty, though right, in that - as noted above - the effect of set-theoretical parameters is not so easy to unravel).

The amount of set theory required for first order semantics is rather weak in its existential assumptions; one needs no strong axiom such as the power set or replacement axiom. The resulting truncated theory Z^- (Zermelo's theory without the power set axiom) is more or less equivalent to an arithmetic theory (whence the option of the arithmetic completeness). For general second order structures, where U is the set of definable subsets of M, a similar possibility holds; the definable subsets are given by the axiom of separation, whatever the kind of (non-)restriction on formulae. By contrast, for the treatment of the notion of all subsets on an infinite set, one needs a full and mature set-theoretical intuition whose reliability is only apparent, and whose associated theory shows worrying relativity phenomena.

To convince oneself of the insurmountable difficulties, recall that the set of all subsets of an infinite set is not effectively generable. This is however just an intuition, not a conclusive proof of the incompleteness, nor a hint of where to look for the proof. Structures for first order logic are not necessarily effectively generated. Effectively generated structures however suffice, in the sense that the completeness theorem is proved using only Skolem-Herbrand interpretations. Couldn't there be anything similar in the new case? After all - inevitably - any completeness theorem for the second order calculus is proved only with reference to structures with definable subsets. But the situation is different for full second order logic in that definable (nonfull) structures are disjoint from full structures. Still, the remark is not conclusive.

To spot the stumbling block one must first investigate some properties of full second order logic. A first important property is the existence of a sentence I without non-logical constant symbols such that for all M, $\mathbf{M} \models_2 \mathrm{I}$ if and only if M is infinite (here **M** is the universe $M \cup \wp(M)$, without additional structure). The sentence I is essentially the same as before, except for replacing the relational symbol R by a variable, and putting an existential quantifier in front:

$$\exists R\, (\forall x\, \neg Rxx \wedge \forall x\, \exists y\, Rxy \wedge \forall x,y,z\, (Rxy \wedge Ryz \rightarrow Rxz)).$$

As a second property of full second order logic, one easily sees that any two models of full second order arithmetic are isomorphic. A third property, which can be obtained from the first, is that the satisfiability problem for first order logic is reducible to the validity problem for full second order logic: as a matter of fact, for any first order sentence A, denote by $\exists^2 A'$ the "second order transform" of A. $\exists^2 A'$ is the second order sentence without constant symbols, given by the following effective construction: let A' be obtained by replacing all relational and functional symbols by variables of the same type and number of places; now pre-pend an existential quantifier for each of these variables.

We now claim that

A has an infinite model if and only if $\models_2 \mathrm{I} \rightarrow \exists^2 A'$.

Proof (sketch): if $\models_2 \mathrm{I} \rightarrow \exists^2 A'$, then given any infinite set M, $\exists^2 A'$ is true in $\mathbf{M} = \langle M \cup \wp(M) \rangle$, hence there are relations and functions on M as dictated by the second order existential quantifiers; adding these to M one gets an infinite first order structure which is a model of A. Conversely, given an infinite model of A, by LS there is a model of A in any infinite cardinality; given now a model of I - that is, an infinite set - transfer on it isomorphically the structure of a model of A of the same cardinality.

If now $\models_2$ were semidecidable, then also first order satisfiability for infinite structures would be so, whence first order satisfiability would be decidable (because, as we have seen, first order satisfiability for finite structures is semidecidable - or by LS). Now the complement of first order satisfiability is semidecidable, since a sentence is unsatisfiable is and only if its negation is valid, and validity is semidecidable. We then conclude that first order satisfiability is decidable, thus contradicting the Turing-Church theorem on the undecidability of first order logic.

The above mentioned properties of full second order logic are well established, but they are so only in a meta-theory talking of full second order structures under the hypothesis that they exist (and that they are different from the definable ones). Saying that $\mathbf{M} \models_2 \mathrm{I}$ if and only if M is infinite implies that, if there exists an infinite set M, then also $\wp(\mathrm{M})$ exists. However, this is not sufficient for our logical conclusions; we must also make sure that

(i) $\wp(\mathrm{M})$ is different from the set of definable subsets of M, and this follows from Cantor's theorem and other cardinality arguments,

(ii) an infinite $\mathbf{M}$ need not satisfy I, if $\mathbf{M}$ is not full.

If M is infinite, there exists an injection r of M in itself, which is not surjective. I essentially establishes the existence of such a function r, or of the binary relation R^M representing its graph. If r is not in U, then I is not true; in general, we cannot be sure that r actually is in U, unless $U = \wp(\mathrm{M})$, because no uniform definition D is able to isolate an infinite proper subset in an arbitrary infinite set. Given any proposed candidate D, by the methods of non-standard analysis (which, in this case, essentially reduce to compactness) one can build a structure in which the subset defined by D is finite.

If one is ready to accept the undefinability of the infinite, the above argument clearly points out the impossibility of applying the comprehension axiom to single out an element such as R^M, in view of the possible undefinability of this element. This is what prevents us from using the tools of $\vdash_2$ to carry out the proof. The distinguishing feature of the proof of the categoricity of full second order arithmetic is more subtle. Here one has to apply the induction axiom - a universal statement - and one needs a particular instantiation of it. Given two models of arithmetic **N'** and **N"**, assume we are trying to define an isomorphism between them. This is done with the help of a certain recursively defined function g from **N'** to **N"** - roughly, $g(S'x) = S''(g(x))$, where S' and S" are the successor functions in the two structures. To define g, and to prove that it is an isomorphism, induction is applied to subsets of **N'** and **N"**. Whenever a subset X' of **N'** is considered, we must also consider X" corresponding to X' in **N"**, and conversely. The construction is intuitively clear, and X" would thus appear to be definable in terms of X' and g. It turns out that, although definable, X" is not definable *within* the structure **N"**, but only *outside* it, i.e., with parameters not belonging to **N"**. Accordingly, when in the definition of full structures one mentions "the totality" of subsets, a crucial specification to be added is "included those definable using also external parameters". This specification however is not available in the formulation of the comprehension axiom.

Since parameters play an essential role in the above proofs, they obviously need to be mentioned and linguistically characterised, otherwise we couldn't formulate our arguments. Their existence, however, is proved in a wider theory, namely in set theory. To prove that a certain subset X of **N** exists, one first has to state the desiderata for X, describe X in the language of set theory, check that the theory indeed guarantees the existence of X. Finally one can asks whether X is also definable within **N**. The set-theoretic machinery needed in the cases under consideration here is not exceedingly strong (weak replacement suffices to get the Cartesian product, then one works with separation). However, the essential fact is that one goes outside the original structure **N**, and these (sets depending on external) parameters may not be definable in the structure. Undefinability is ascertained only indirectly, by way of contradiction, upon noting that definability would entail contradiction with some rather cherished theory or result.

The choice of full second order structures thus appears to be not so much a restriction of the class of admissible interpretations, but rather, the assumption of a different notion of interpretation. When saying that in the (full) structure there positively must be included sets which are only externally definable, we are using - already in the definition of the structure - a notion of satisfiability that depends on the universe of sets, or at any rate, depends on a fragment of it which is substantially wider than the universe of the structure itself. When dealing with first order semantics, it is enough to give the universe of discourse W, equipped with some extra structure (i.e., some specified relations and functions in W). We then automatically obtain an internal and compositional notion of truth in the structure. In the case of second order semantics, we have a weird example of an impredicative circularity of the worst kind: in order to define the structure, we must have and use a truth notion which does not coincide with any built-in notion in the structure itself (internal notions are only able to produce definable subsets). We rather hopelessly try to capture a notion of "truth in itself", or "truth in the universe".

The moral is that the only interpretations under our control are those which do not depend on a preliminary notion of truth - but only depend on (what else?) syntax. This was clear from our positive treatment of first order completeness. When the notion of truth is a prerequisite for the very definition of full structures, our attempts to master these structures are doomed to failure: failure of completeness is a confirmation that anchorage to something more steady than language is an illusion.

On the other hand, the plot thickens, because in the set-theoretical universe, no notion of truth is available, as we know from the Gödel-Tarski undefinability theorem. It we really relied on such notion, then we should reach much more than the non-effectiveness of second order logic, we should reach a contradiction. And yet we do talk about such topics as truth, full structures; we even prove(d) meta-theorems on them. This is so because we have been talking and arguing in a first order logic with the set-theoretical language. Actually, when defining sets, we do not use the notion of truth in the set-theoretical universe; it is only a way of talking (as usual with truth), a way of saying that the existence of the sets we are interest in is demonstrated in the (first order) theory of sets. Luckily, our meta-theoretical arguments are still subjected to effective generation rules. So in the end the picture is as follows: there is a consequence relation $\models_2$, which is perfectly and correctly defined in set theory; we would like to find generation rules appropriate to it, by formulating these rules in a language restricted to special sets only: sets as universe of a structure, their subsets, and nothing else. Thus viewed, incompleteness does not come as a surprise; we are familiar with many other examples of incompleteness arising from restricting rules from larger to smaller languages - the simplest being perhaps given by the equational calculus, which allows us to define in the structure of numbers only the recursive functions, while much larger is the class of functions which are definable in arithmetic, already in the first order language.

A lingering doubt could be represented by the fact that in the end, notwithstanding all the many and varied arguments, the key to the proof of the incompleteness of full second order logic lies in the contradiction one would get with the (albeit indisputable) undecidability of first order logic. But further reflection shows some common elements of the two phenomena. Recall that the completeness proof has shown the necessity of (cases of) non-termination, and related disturbing consequences, like the undefinability of finiteness. The latter depended on the fact that if a sentence has arbitrary large finite models then it has an infinite model. But this very fact is at the basis of the proof of König's lemma. The latter is an essential ingredient of - in fact, it is essentially equivalent to - the completeness proof, when we finally single out an infinite subtable from a nonterminating table. We conclude that the completeness of first order logic is equivalent to the undefinability of finiteness.

If full second order logic were semidecidable, and the first order satisfiability problem could be reduced (not to logical truth but) to full second order derivability, say in the form: A is satisfiable if and only if $\vdash_2 I \to \exists^2 A'$, then first order logic would be decidable. As a matter of fact, according to the machinery of tableaux, A is satisfiable if and only if the table initialised with A in **T** does not close; but the developed table can be, and in some case is, an infinite one; under our standing (*absurdum*) hypothesis, this non-termination would be matched by the termination of the semidecision algorithm searching for a $\vdash_2$-derivation. The infinity of the first nonterminating process would be revealed by a finite message. This argument is far not conclusive, since the infinite allows a finite definition, otherwise we could not even to talk of it. True, the definition of "infinite" in a finite set-theoretic formula - but the verification that a set is infinite is another matter.

Our definitions are finite in space, but they do not afford a clear cut definition of finite (hence by negation of infinite); they are deceptive in their apparent precision, because they have always a tail which is hard to define and which shows in their applications. This tail tends to infinity, though it is always finite. The right measure might be given by some dynamical measure, including not only space, but also time of verification.

When the table is infinite, we would like to discover it from the inside of the process, hence at a finite stage, but there is not always a finitely recognisable loop responsible for non-termination. Upon transferring these considerations to second order structures, it would follow that the infinity of M should be revealed by the inner satisfiability in **M** of some sentence. It is plausible therefore that the injection of M into itself should not be definable in **M**.

These may look wild speculations, without head or tail. However, analogies may be useful to go into the deeper reasons, and to explore the uncertain borders between finite and infinite, and the no man land of the arbitrary large in between. Indeed, the crucial notion is precisely that of “arbitrary large”, or “eventually”. One would altogether say that precisely this very fuzziness allows one to prove that a logic is complete: the finiteness of the syntax can express the intuitive infinite of the semantics. Our leading analogy, between the infinite process of tableaux and definability in structures, is more akin to the distinction between inner and outer than it is with the finite/infinite dichotomy. What looks finite from within (a process, a definition) may look infinite from without, and conversely; this is also the lesson of non-standard methods. What is syntax from within is the semantics from without.

The completeness theorem has a history; such is the destiny of the important theorems, those for which for a long time one does not know (whether there is anything to prove and) what to prove. In its history, one can distinguish at least two main paths; the first one covers the slow and difficult comprehension of the problem in (what historians consider) the traditional development of mathematical logic canon, up to Gödel's proof in 1930; the second path follows the Löwenheim-Skolem theorem. Although at certain points the two paths crossed each other, they started and continued with their own aims and problems. A classical *topos* of the history of mathematical logic concerns the how and the why Löwenheim, Skolem and Herbrand did not discover the completeness theorem, though they proved it, or whether they really proved, or perhaps they actually discovered, completeness. In following these two paths, we will not always respect strict chronology, keeping the two stories quite separate, until the crossing becomes decisive.

In modern pre-mathematical logic, the notion of completeness does not appear. There are some interesting speculations in Kant which, by some stretching, could be realized as bearing some relation with the problem; Kant's remarks, however, are probably more related with incompleteness, in connection with his thoughts on the derivability of transcendental ideas (or concepts of reason) from categories (the intellect's concepts) through a passage to the limit; thus, for instance, the *causa prima*, or the idea of causality, is the limit of implication, or God is the limit of disjunction, viz., the category of "comunance". The striving for completeness leads to antinomies. Also Kant's transcendental deduction of categories could be compared to what is now called functional completeness of sets of connectives. Altogether, such a reading of old works is always a bit forced.

A first inkling of the completeness problem is often said to be retraceable in Bernhard Bolzano. In his 1837 *Wissenschaftlehre*[1] Bolzano gave the first neat formulation of the notion of logical consequence in substitutional version. He had clearly seen that all non-logical parts of the statements, constants and predicates, must be thought of as variables in the logical treatment. Some scholars have reservations on his achievement, since Bolzano did not consider a variable domain; but it is only fair to say that he is in good

[1]B. Bolzano, *Wissenschaftslehre*, 4 voll., Seidel, Sulzbach, 1837.

company here, because also Tarski in his first enquires on semantics did not let the universe vary; the universe of logic, in the twenties, tended to coincide with that of the natural numbers.

Owing to their peculiarity, Bolzano's problems are difficult to translate in modern terminology. His main idea was that of a proposition *in seed* , also called are also propositions in thought, or statement of judgements. We could probably say that Bolzano's propositions are the sense of sentences. A proposition *in se* has a reality, different from that of any written sentence. A proposition *in se* is made of representations and concepts that might happen to have a reality only in our minds; hence, a proposition can be without an object; nevertheless, we can talk of propositions and consider relations among them such as, for example, inclusion. Among these relations, Bolzano also considers those that arise from the fact that propositions contain variable elements. Hence the definition of logical consequence: "In a *more restricted* sense [with respect to the case in which there is compatibility between antecedent and consequent] – and in this sense I will use from now on this expression – a proposition M is deducible from propositions $A, B, C, \ldots$ with respect to variables $i, j, \ldots$, when any particular representation substituted to $i, j, \ldots$ which makes true all the propositions $A, B, C, \ldots$ makes true also proposition M"[2]. In this definition there is no reference to a restricted domain for variables (while in other places this restrinction is imposed to avoid infinite possibilities), nor to a compatibility condition among $A, B, C, \ldots$ and M, a condition to which the *proviso* on "restricted sense" alludes. Variable parts are essential to make sense of representations without an object, such as the representation of "nihil", or of $\sqrt{-1}$; these interact with the others only in so much as "we conceive of certain elements $i, j, \ldots$ as *variable* and we compare the infinitely many *new* representations that emerge by substituting to $i, j, \ldots$ different representations, whenever these representations have an object".

A completeness problem could only arise from the comparison of the consequence relation with another relation; in fact Bolzano considers a new relation among propositions, based on a notion of structure, and having some resemblance with the much later patterns of derivation of natural deduction. In his *Wissenschaftslehre*, he introduces a relation of *Abfolge*, also called

[2]B. Bolzano, *Von der mathematischen Lehrart* (1840-42).

relation of consequence-foundation; in contrast to plain consequence, this new relation concerns truth in itself and not knowledge. Bolzano distinguishes between objective foundations and epistemological foundations, and between objective proof and subjective proof. The former goes backwards from a consequence to its foundations; premises are called subsidiary truths of the conclusion; proceeding in this way, one eventually finds out the basic truths. The underlying structure is a tree, and the subsidiary truth for a proposition P is the set of truths occurring in the tree. The actual description of the tree does not amount to a full-fledged inductive definition, of the sort: the set of truths grows at each level, and in the nodes there are only concepts included in the conclusion. Some people have seen here an analogy with normal Gentzen derivations and cut elimination. While Bolzano's two notions of consequence do ask for comparison, one cannot find here any formulation of the completeness problem as comparison between syntax and semantics.

Let us consider now the proper tradition of mathematical logic. Peano was only interested in showing that all logical identities *normally used* in mathematical arguments could be retrieved from a very small set of identities. Boole felt at least the need for completeness, though in the context of a very traditional conception of the axiomatic organization of a discipline[3]: "All sciences consist of general truths, but of those truths some only are primary and fundamental, others are secondary and derived. The laws of elliptic motion, discovered by Kepler, are general truths in astronomy, but they are not its fundamental truths. And it is so also in the purely mathematical sciences . . . Let us define as fundamental those laws and principles from which all other general truths of science may be deduced, and into which they may all be again resolved[4]. Shall we then err in regarding that as the true science of Logic which, laying down certain elementary laws, confirmed by the very testimony of the mind, permits us thence to deduce, by uniform processes, the entire chain of its secondary consequences, and furnishes, for its practical applications, methods of perfect generality? Let it be considered whether in any science, viewed either as a system of truth or as the foundation of

[3]G. Boole, *An Investigation of the Laws of Thought* (1854), Dover , New York, p. 5.

[4]"It may be added, that they are truths which to an intelligence sufficently refined would shine forth in their own unborrowed light, without the need of those connecting links of thoughts, those steps of wearisome and often painful deduction, by which the knowledge of them is actually acquired", *ibidem*.

a practical art, there can properly be any other test of the completeness and the fundamental character of its laws, then the completeness of its system of derived truths, and the generality of the methods which it serves to establish".

As for Russell, in *Principia Mathematica* he did not go beyond the statement of an empirical completeness of the logical system. Actually, empirical completeness is realized in PM, despite completeness is not explicitly formulated as an hypothesis – Hilbert would do in 1928. Neither our distinctions nor our terminology can be found in PM. Discussing the content of the first part of PM, where deductions of logical truth are expounded, the authors explain that such content can be seen in two ways, "either as a deductive chain based on primitive ideas, or as a formal calculus". By formal calculus they mean a kind of algebraic treatment, alternative to the logical format, but patently equivalent – the choice among them is a matter of taste. In the absence the necessary distinctions, completeness could be conjectured only *à la* Peano. The authors also say that the subject to be treated "is not quite properly described as the theory of propositions. It is in fact the theory of how one proposition can be inferred from another. Now in order that one proposition may be inferred from another, it is necessary that the two should have that relation which makes the one a consequence of the other. When a proposition q is a consequence of a proposition p, we say that p *implies* q. Thus deduction depends upon the relation of *implication*, and every deductive system must contain among its premises as many of the properties of implication as are necessary to legitimate the ordinary procedure of deduction. In the present section, certain propositions will be stated as premises, and it will be shown that they are sufficient for all common forms of inference. It will not be shown that they are all necessary, and it is possible that the number of them might be diminished. All that is affirmed concerning the premises is (1) that they are true, (2) that they are sufficient for the theory of deduction, (3) that we do not know how to diminish their number. But with regard to (2), there must always be some element of doubt, since it is hard to be sure that one never uses some principle unconsciously. The habit of being rigidly guided by formal symbolic rules is a safeguard against unconscious assumptions; but even this safeguard is not always adequate"[5].

[5]A. N. Whitehead, B. Russell, *Principia Mathematica*, Cambridge University Press, Cambridge, 1910-13, Part I, "Mathematical Logic", Section A, The Theory of Deduction; quotation from A. N. Whitehead, B. Russell, *PM to *56*, Cambridge University Press,

Russell had his reasons not to be too sure of completeness, considering the complications and *ad hoc* restrictions introduced in his theory of types; he had made experiments with alternatives and variants, always in the setting of higher order logics, and he was well aware of the many decisions and choices to be made in building a logical system, and of their non-equivalence. He probably believed that, to avoid contradictions, what was excluded in his system would have to be left out in any other system; but he might well have some doubts, at least if we read with attention: "It should be observed that the whole effect of the doctrine of types is negative: it forbids certain inferences which would otherwise be valid, but does not permit any which would otherwise be invalid. Hence we may reasonably expect that the inferences which the doctrine of types permit would remain valid even if the doctrine should be found to be invalid"[6].

Frege had no such doubts, to him logic was one, necessarily ideographic. Frege's formalization was not a tool to abstract from meaning; on the contrary, only when expressed in his *Begriffsschrift* sentences could be properly understood and analysed. Frege's worries were of another type: the *inhalthich* reasoning could be maimed and suffer an *Inhaltsentleerung* in such systems as Boole's calculistic logic, Jevons' machine or formal arithmetic: "Boole's formulistic language reproduces only a part of our thought; in its entirety it could never be done by a machine or substituted by a purely mechanical procedure. Syllogism can certainly be presented under the form of a calculus, which of course will never be executed by a machine, but thanks to the few schematic and intuitive forms in which such a calculus is organized it guarantees great certainty. But the true gain is got when the content is not only shown, but constructed from its component parts through the very logical signs used to calculate"[7].

According to Poincaré, complete formalization, as achieved by Hilbert for geometry, was a step toward mechanical theorem proving. He was pushing in this direction for polemical reasons, a kind of practical *reductio ad absurdum*. Couturat had remarked[8] that Peano, in translating all mathematical

Cambridge, 1964, p. 90.

[6]Ivi, Preface of 1910, p. vii.

[7]G. Frege, "Booles rechnende Logik und die Begriffschrift", in G. Frege, *Nachgelassene Schriften*, Felix Meiner Verlag, 1969.

[8]L. Couturat, "For Logistics", *Monist* 22 (1912), pp. 481-523.

expressions in symbols, did not aim at excluding thought from logistics, first of all because there was on the background the operation of translation, and moreover because the task of building derivations was not of the nature of blind mechanism. The discussion on formalization was really a discussion on axiomatic method: on the one hand, formalization applied it to axiomatic theories; on the other hand, as was fully understood only later, by the very nature of axiomatization, to axiomatize was virtually the same thing as to formalize – whether or not resorting to logistics. for the very nature of axiomatization (as was later fully understood). Frege was, consistently, a strong opponent of the axiomatic method. The debate on axiomatization involved completeness, specifically, the completeness problem for theories. As we shall see, the difficulties to separate the notions of completeness for theories and completeness for logic, were a source of long lived misunderstandings.

The modern axiomatic method begins with the idea to deprive primitive concepts of their meaning. For our purposes we do not need a detailed analysis of the evolution of this idea. One quotation will be enough, to recall the background. According to Pasch, "in order that geometry becomes a truly deductive science it is necessary that the entailing of consequences be everywhere independent from the *sense* of the geometrical concepts, as it must be from the pictures; one must take in consideration only the relations between geometrical concepts, as determined by the propositions and definitions employed. In the course of a deduction, it is allowed, and useful, to think of the meaning of intervening geometrical concepts, but it is not necessary; when it is necessary, this is a sign of the defective character of deductions and ... of the inadequacy of the propositions invoked to support the proof. When one deduces ... a theorem from a group of propositions – which we'll call generators – the value of the deduction goes beyond the initial aim. In fact, if one derives from generators correct propositions, then by changing the geometrical concepts with others... one gets without reproducing the proof a proposition which is a consequence of the thus modified generators"[9]. In the reasoning process, sensible images are not allowed, nor mental representations of images; the words themselves used to frame geometrical concepts exert an influence from which science must get free. The right synthesis, according to Pasch, was to be found in Peano's school, where not only the axiomatization of theories was pursued but a sterilized language

[9]M. Pasch, *Vorlesungen über neuere Geometrie*, Teubner, Leipzig, 1882, p. 98.

was also made ready for the purpose.

Frege's opposition the axiomatic method was motivated by the lack of determination of the objects described by a set of axioms. Axioms were seen by mathematicians as a new kind of definition; the peculiar nature of this kind of definition was recognized, and accepted. The terminology varied: descriptive definitions *vs* the traditional real and nominal ones, implicit definitions[10] *vs* the explicit ones, definitions by postulates *vs* direct definitions. Poincaré didn't see any problem in accepting the axioms as masquered definitions[11]. Following Couturat's terminology, Poincaré talked of definitions by postulates: "Stuart Mill claimed that any definition implies an axiom, that which asserts the existence of the defined object ... The definition would thus be an axiom in disguise ... Stuart Mill construed existence in a material and empirical sense; he meant that in defining the circle one says that in nature there are round things. In this form, his idea in not admissible. Mathematics is independent from the existence of material things, here existence can have only one meaning, namely freedom from contradictions. Thus modified, Stuart Mill's idea is right: in defining an object, one claims that the definition implies no contradiction. If we have thus a system of postulates and we can show that these postulates do not imply any contradiction we'll be entitled to consider them as the definition of one of the notions occurring in them"[12].

To Frege, axioms are not definitions, since they realize neither of the necessary conditions of existence and uniqueness of the defined object. According to Frege[13] definitions (conceived as assignments of meaning to signs) are to be kept distinct from all other sentences of a science, such as axioms, principles, theorems, explanations. Frege objects to Hilbert that his axioms for the "between" relation do not define a relation at all; the definition should be framed as "'between' is a relation with the following characteristic notes

[10]At the beginning, talk of implicit definitions was inspired by the mathematical terminology of implicit functions, and it referred to single terms included in propositions together with other already known terms; later on, "the implicit definition of systems of concepts by means of propositions has become essential for mathematical logic", F. Enriques, *Per la storia della logica*, Zanichelli, Bologna, 1922, 1987^2, p. 134.

[11]H. Poincaré, *La Science et l'hypothèse*, Flammarion, Paris, 1902, p. 67.

[12]H. Poincaré, *Science et méthode*, Flammarion, Paris, 1908, p. 161.

[13]G. Frege, *Wissenschaftlicher Briefwechsel*, Felix Meiner, Hamburg, 1976, letters to and from Hilbert.

...”; Hilbert is willing to change terminology, but “I have to be entirely free also in the statements of the characteristic notes. In fact, as soon as I have posed an axiom, it exists and is ‘true’ ... You write: ‘The fact that the axioms are true assures us that they do contradict one another’. I have been intrigued in reading this statement in your letter because, since when I began to think and write and lecture on this subject, I have always said the very opposite: if axioms arbitrarily posed are not in contradiction, together with all their consequences, then they are true, then there exist the entities defined by means of those axioms. This is for me the criterion of truth and existence”[14].

Notice that also a few formalists, Couturat among them, were opposed to the idea that a mere consistency proof could guarantee existence. The same opposition is to be found in early intuitionism, with its absolute mistrust for the “linguistic buildings, sequences of sentences connected by logic rules ... If also it were apparent that such constructions can never exhibit the linguistic pattern of a contradiction, they are mathematical only in so far as they are linguistic constructions and they have nothing to do with mathematics, which is outside the building”[15]. Brouwer rejects the existence of models for consistent sets of sentences: “The question is in the following terms: let us suppose we have somehow proved, without thinking of a mathematical interpretation, that a system logically built on the basis of a few linguistic axioms is non-contradictory, i.e., that at no stage of the development of the system we can meet with two contradictory theorems; if then we should also find a mathematical interpretation of the axioms..., does it follow then that such mathematical construction exists? Nothing of the kind has ever been proved by the axiomatizers”[16].

Brouwer’s negative position stems from his conception of mathematical construction, which has nothing to gain by a consistency proof. But Frege does not dismiss the significance of consistency proofs; rather, he questions their feasibility: “Which means do we have to prove that certain properties, or certain requirements (or whatever else they are called) are not mutually

[14]D. Hilbert in G. Frege, *Wissenschafticher Briefwechsel*, cit.

[15]L.E.J. Brouwer, *Over de grondslagen der wiskunde*, Dissertation, Maas & van Suchtelen, Amsterdam, 1907, pp. 183, p. 132, in L.E.J. Brouwer, *Collected Works I*, North Holland, Amsterdam, 1975, pp. 11-101.

[16]ivi, p. 141.

contradictory? The only way I know is that of presenting an object which possesses all those properties, or of mentioning a case in which all those requirements are satisfied. It should not be possible to prove consistency in any other way"[17]. Also according to Poincaré "Usually to prove that a definition does not imply a contradiction, one proceeds *with an example*; one looks for an example of an object which satisfies the definition ... But such a proof is not always possible"[18]. Frege's statement is of December 1899; soon Hilbert would have – if he didn't already have – another idea. An alternative to examples is that of considering all the propositions one can deduce from the postulates and showing that there are not among them two which are mutually contradictory. This idea pleased Poincaré, because starting from it he was able to argue for the impossibility of a consistency proof for arithmetic; such a proof as envisaged by Hilbert would need, circularily, induction; hence the synthetic character of induction was established. Also Mario Pieri, in 1906, had an inkling that for consistency proofs "one will have to reason by induction on the infinite sequence of derived propositions"; he doubted the feasibility of the procedure for a queer reason, because "we could not decide whether [the set of derivable propositions] will be a denumerable series, to which such a principle could be applied"[19]. Pieri was not sure of this because probably he had not grasped the meaning of the (albeit hypothetical) completeness of a set of mechanical logical rules.

Besides the existence problem for definitions (also called the solvability problem, as Frege preferred to say by analogy with equations), there is the uniqueness problem. Non-uniqueness is, to the axiomatizers, a positive feature of definitions by postulates: "It might happen that there are several interpretations pf the undefined symbols which satisfy teh system of unroved propositions. The system of undefined symbols can then be considered as an abstraction obtained from all interpretations"[20]. It is not by chance that there are several interpretations: "The most important property of the primitives of a hypothetical-deductive system is that of being capable of

[17] G. Frege, *Wissenschafticher Briefwechsel*, cit.

[18] H. Poincaré, *Science et méthode*, cit., pp. 161-63.

[19] M. Pieri, "Sur la compatibilité des axiomes de l'arithmétique", *Revue de Métaphysique et de Morale*, 13 (1906), pp. 196-207, p. 199.

[20] A. Padoa, "Essai d'une théorie algébrique des nombres entieres, prédédeé d'une introduction logique à une théorie déductive quelconque", in *Bibl. du Congrès Intern. de Philos.*, Paris, 1900, A. Colin, Paris, 1900, vol. 3, pp. 309-65.

arbitrary interpretations, within the boundaries fixed by the primitive propositions ... or, in different terms, the meaning of words and symbols denoting a primitive object whatsoever is not uniquely determined by the primitive propositions concerning it, and the reader has the faculty to attribute a meaning to these words and signs as he likes, with the only condition that it be compatible with the general properties imposed on objects by the primitive propositions"[21]. If meaning can be attributed in several different ways, then none of the attributions is in a sense complete. As Beppo Levi remarked, with reference to the indeterminacy of primitive ideas: "it is true that a given system of postulates can give a primitive idea (in relations to the other ones) a determination less strong than that associated to the name in the common speech; but the true and complete determination of a primitive idea in not possible, however complex the system of marks (contrassegni) used for it; we will never be able to identify ideas, but only to state some relations for them[22].

Hilbert was clearly well aware of the possibility of multiple interpretations; in some of his first pronouncements, he seems to construe the indetermination only up to isomorphism; at the same time, he claims that the very indetermination is a great gain. He wrote to Frege: "You observe that my concepts, e.g., 'point' and 'between' are not uniquely determined ... But it is obvious that any theory is just a framework, a schema of concepts with their mutual relations, while the basic elements can be thought of in an arbitrary manner. If with my points I want to refer to any system whatsoever, say the system composed of love, law, chimney-sweepers ..., then it will be enough that I construe my axioms as relations among these entities, and all my propositions, e.g., the Phytagorean theorem, will hold for them. In other words: any theory can be applied to infinitely many systems of basic elements. It is enough to apply a one to one transformation and to convene that the axioms for the transformed objects be the same as for the corresponding ones. Such a possibility can never be imputed as a defect of a theory, on the contrary, it is a very great advantage, and at any rate it is inevitable"[23]. It is quite possible that this restriction to one-to-one correspondences had to do with

[21]M. Pieri, "I principii della geometria di posizione composti in sistema logico deduttivo", *Mem. R. Accad. Sci. Torino*, (2a) 48 (1899), pp. 1-62, p. 60.

[22]B. Levi, "Antinomie logiche?", *Annali di Matematica*, (3) 15 (1908), pp. 187-216, footnote (*), p. 188.

[23]D. Hilbert, in G. Frege, *Wissenschaftlicher Briefwechsel*, cit.

his ideas on completeness, which at that time were not so clear, as we shall see.

At the beginning of the modern axiomatic method, mathematicians used to give a somewhat greater emphasis than it is done today, to the fact that isomorphic interpretations are nonetheless different; for example Fraenkel in 1928 observed that when "for a particular concrete (*inhalthichen*) meaning of the primitive concepts, e. g. intuitive 'point' and 'line', a proposition is *richtig*, that is deducible from the axioms, then the proposition cannot be false with respect to another meaning, compatible with the axioms (for example 'point' as 'pair of numbers'), or we would have a contradiction with the proved isomorphism. But this does not mean that the sense, the essential content of primitive concepts, can ever be determined by the axioms, because to any interpretation there is another one, isomorphic but with a different sense"[24]. This can be read as an indirect answer to Levi: sense is not invariant under isomorphism.

Poincaré had no difficulty in accepting the necessity of an axiomatic presentation of geometrical systems; and he was aware that the eventual goal of axiomatization was that of a formal presentation of the theory. "One is struck in the new mathematics by its formal character, that is so explained by Hilbert: 'let us think of three species of *things* we'll call points, lines and planes, and let us convene that a line will be determined by two points; instead of saying that this line is determined, let us say that it passes for the two points, or that these two points lie on the line'. What these *things* actually are we not only do not know, but we must not try to know. We do not need it, and a person who had never seen points, lines and planes could do geometry as well as ourselves. Words like *to pass*, or *to lie* must not generate in us any image, they are simply synonyms for, respectively, *to be determined* and *to determine*. So it is clear that to prove a theorem is not necessary nor useful to know what one means. One could replace the geometer with the *piano à raisonner* imagined by Stanley Jevons; or if you prefer one could think of a machine where from an entry one could introduce the axioms and on the other end to collect theorems, as the famous Chicago machine where pigs enter alive and come out as ham and sausages. No more than these machines, the mathematician needs not to understand what he is

[24] A. A. Fraenkel, *Einleitung in die Mengenlhere*, Springer, Berlin, 1928^3, p. 353.

doing. This formal character of geometry I do not impute to Hilbert, this was necessarily his goal, given the problem he had set to himself. He wanted to reduce to a minimum the fundamental axioms and make a complete enumeration of them; now in reasonings where our spirit is alive, in which intuition plays a role, in live reasonings, so to speak, it is difficult not to introduce some unobserved axiom or postulate. Only after reducing all geometrical reasonings to a mechanical form he could be sure to have succeeded in his project and reached his goal"[25].

Hilbert's aim in fact, stated at the beginning of his 1899 *Grundlagen der Geometrie*, was that of obtaining for geometry an axiom system "complete and as simple as possible". What he meant by "complete" can be recovered from his letters to Frege. Hilbert explains to Frege that a complete definition is given only by the whole set of axioms; every axiom contributes to the definition, hence every new axiom modifies the concept. "Adding any new axiom once a concept has been determined in a unique and complete way is absolutely forbidden and non-logical"[26]. In the case of real numbers and geometry, however, Hilbert also discusses the notion of a structure being full – being impossible to further extend it. This notion and that of deductive closure do not exactly coincide. The *Vollständigkeitsaxiom* for geometry states the impossibility of extending the universe of points, lines and planes and still keeping the validity of all the axioms (for real numbers, he replaces the continuity axiom by the completeness and Archimedean axioms). At the same time he looks for a system of axioms so closed that no other axioms can be consistently added, and no proposition can be considered correct unless it is derivable from the axioms. When comparing the genetic and the axiomatic method, he observes that with the latter there is no need to think of the generation laws for reals, but only of "a system of things whose relations are given by means of the finite and closed system of axions I.IV, about which new assertions hold only if they can be derived from those xions by means of a finite number of logical steps"'[27].

Here Hilbert seems to consider categoricity a consequence of completeness, though in 1901 he apparently was no more convinced of the identity of the two notions. According to Edmund Husserl, after a lecture in which this

[25] H. Poincaré, *Science et Méthode*, cit., pp. 156-8.
[26] D. Hilbert, in G. Frege, *Wissenschaftlicher Briefwechsel*, cit.
[27] D. Hilbert, "Über den Zahlbegriff", *Jahresbericht DMV*, 8 (1900), pp. 180-4.

very subject had been discussed, Hilbert remarked that one should carefully consider the logic by which the consequences of the axioms are derived; in Husserl's words: "When we suppose that a proposition be decided on the basis of the axioms of a domain, what can we use besides the axioms? *Alles Logische. Was ist das?* All the propositions that are free from any particularity of a knowledge domain, what is independent from all particular axioms, from all matter of knowledge". But here one has a spectrum of possibilities: "the algorithmic logic domain, that on numbers, of combinatorics, of the general theory of ordinals. And at last the most general set theory is not in itself pure logic?[28]". Combinatory logic for example is sufficient to derive the *Schnittpunktsatz* from Pascal theorem (Hilbert had proved that there was no need of continuity for this derivation); the logic of numbers comes in when the Archimedean axiom is used, and to use the *Vollständigkeitsaxiom* one has to resort to the logic of sets, the *allgemeinste Mannigfaltigkeitslehre*[29]. In 1909 Hilbert again remarks that "in certain investigations of modern mathematics the problem is not that of establishing a specific fact or the validity of a proposition, it is that of carrying on the proof with the restriction to given methods or to prove the impossibility of such a demonstrative process"[30].

In the same years, completeness and categoricity where being discussed also in USA; one of the merits of the discussion was that of forging the definite terminology. Initially, Huntington[31] calls "complete" a supposedly categorical theory; in 1902 he proposes a set of six postulates for continuous magni-

---

[28]E. Husserl, *Philosophie der Arithmetik*, Martinus Nijhoff, The Hague, 1969, p. 445, quoted by J. C. Webb, *Mechanism, Mentalism and Metamathematics*, Reidel, Dordrecht, 1980, p. 85.

[29]According to Otto Blumenthal, already in 1904 Hilbert was convinced that "without a complete and perspicuous formalization of logical inferences no progress could be done in the direction he had indicated", O. Blumenthal, "Lebensgeschichte", in D. Hilbert, *Gesammelte Abhandlungen*, Dritter Band, Springer, Berlin, 1935, pp. 388-429, p. 422. The 1904 paper to which Blumenthal refers is D. Hilbert, "Über die Grundlagen der Logik und der Arithmetik", *Verhandlungen des Dritten Internationale Mathematiker Kongresses in Heidelberg vom 8. bis 17. August 1904*, Teubner, Leipzig, 1905, pp. 174-85.

[30]D. Hilbert, "Wesen und Ziele einer Analysis der unendlichvielen unabhängigen variablen", 1909, in D. Hilbert, *Gesammelte Abhandlungen*, Dritten Band, cit., pp. 56-72, p. 72.

[31]E. V. Huntington, "A Complete Set of Postulates for the Theory of Absolute Continuous Magnitude", *Trans. AMS*, 3 (1902), pp. 264-79. Huntington uses assemblage for "set", mentioning also "Menge" and "ensemble"; while it is curious that he uses set for the set of postulates.

tudes and he claims that it is complete, meaning that the axioms are non-contradictory, sufficient and mutually independent; here "non-contradictory" means that "there is at least one *assemblage* in which the chosen rule of combination satisfies all six requirements"; the meaning of "sufficient" is that "there is essentially *only one* such assemblage possible" (modulo correspondences which we now call isomorphisms). This use of "complete" will remain for a while also after the spreading of "categorical"; according to Enriques, in 1922, "a system of postulates is said to be complete when two systems of entities forced to satisfy the system can be put in a one to one correspondence, in such a way that the properties of the one translate in perfectly homologous properties of the other, so that they appear abstractly the same, as far as the ideas in question are concerned"[32].

Also Veblen is interested in categoricity[33]: since his terms *point* and *order* are undefined, he claims he has the right to apply those terms to whatever class of objects for which the axioms are valid. "It is part of our purpose hoeìwever to show that there is *essentially only one* class in which the twelve axioms are valid"[34]. It will follow that any proposition expressed in terms of point and order either is in contradiction with the axioms, or is equally true of all classes which verify the axioms. This means that the validity of every possible proposition expressed in these terms is completely determined by the axioms. Such a system is called by Veblen categorical, while a system to which it is possible to add some new axiom (possibly in many different ways) is called disjunctive. In a footnote, he ascribes these terms to John Dewey; he recalls also Hilbert's terminology of *Axiom der Vollständigkeit*, translated as *Axiom of Completeness*, and Huntington's use of the adjective *complete* for the corresponding definition of the structure (where Veblen prefers to speak of determination, instead of definition).

After 1905, Huntington dismisses the term "sufficient" and following Veblen adopts "categorical"; however, initially, categoricity is understood by him in the sense that every proposition expressed with the primitive terms either is deducible from the postulates or else it is in contradiction with

[32]F. Enriques, *Per la storia della logica*, cit.

[33]O. Veblen, "A System of Axioms for Geometry", *Trans. AMS*, 5 (1904), pp. 343-84.

[34]Again, modulo correspondences we call isomorphisms, though Veblen does not use this term.

them[35]. Actually, in Veblen's terminology, one does not require that every consistent proposition is derivable, but only that it was valid in all models. Following a suggestion of H. N. Davis, Huntington soon corrects himself: "in the case of any categorical set of postulates one is tempted to state the theorem that if a proposition can be stated in terms of the fundamental concepts, either it is itself deducible from the postulates or else its contradictory is so deducible; it must be admitted however that our mastery of the processes of logical deduction is not yet, and possibly never can be, sufficiently complete to justify this assertion"[36].

A thorough discussion of the problem is finally given by Edwin Wilson[37], in a contribution on the axiom of choice. He starts with the remark that it is not always desirable to have a categorical system, because the cardinality of different models can be an interesting feature, for example in group theory; he concedes that from categoricity it follows that any proposition built from the primitive terms (thus, with the exclusion of propositions dealing with such non-mathematical features as, e.g., colour) is either compatible or incompatible, and he asks whether it can also be said that it must be either deducible or in contradiction with the axioms. "This question, this suggestion that compatibility and deducibility may not be the same when applied to categorically determined systems is vital in logic and requires careful discussion ... What, however, does the word deducible mean? The meaning is entirely relative to the system of logic which is available for drawing conclusions from the set of primitive propositions. Some may consider that the human mind has instinctively at its disposal all valid methods of deduction. This is a tremendous postulate, and one entirely devoid of other than sentimental value. In fact, it leads to the abandoning of the research for valid methods of deduction, it is dangerous and worse than useless. It is essential of the modern attitude in logic that the deducer should state distinctly his form of inference". The efforts to relate and equate compatibility and deducibility can be useful and produce advancement in logic. "It appears to me, however, that it may be a distinct gain in precision and hence a considerable advantage to admit the following orienting propositions, namely:

[35] E. V. Huntington, "A Set of Postulates for Real Algebra", *Trans AMS*, 6 (1905), pp. 17-41, footnote §, p. 17.

[36] E. V. Huntington, "A Set of Postulates for Ordinary Complex Algebra", *Trans. AMS*, 6 (1905), pp. 209-29, footnote †, p. 210.

[37] E. B. Wilson, "Logic and the Continuum", *Bull. AMS*, 14 (1908), pp. 432-43.

so long as there is an unsolved problem of pure mathematics the solution may be lacking 1° because the class of objects to which the problem belongs is not sufficiently determined or 2° because the available logical methods of deduction are insufficient; but in case the class of objects is categorically determined, 2° alone applies". In a subsequent discussion on the contrast between Peano and Zermelo, with reference to the axiom of choice, Wilson observes that for Peano and Zermelo "deducible" has different meanings. As for the well-ordering theorem, it seems to Wilson that the relevant definitions can be given without introducing new concepts with respect to Huntington's axiomatisation, hence Zermelo's axiom cannot be a new postulate on the continuum, given the completeness of latter – a remark which was often heard in those times. "As a matter of fact, the postulate that he does add concerns classes in general and the usage he makes of it and his own statements show that what he has done is to demand a new postulate or principle of logic. In view of the fact that it may be doubted whether our logic is yet complete and that Zermelo's postulate is apparently not in contradiction with the other logical postulates, it is difficult to see how any one can deny him the right to proceed as he does". Wilson is so appalled by the complications arising in the theory of well-orderings, that he doubts whether Zermelo's decision is a sensible one; but Zermelo has the right to do it: "to deny him that privilege would be to put an embargo on the development of logic and to assume a completeness of our logical system wuìhich is quite unwarranted in view of past developments and future possibilities".

In the following years, the problem of undertanding the meaning of a complete axiomatisation and of giving one, as well as the related problem of determining the *Alle logisches* were dealt with in a theory which was perhaps too difficult to handle in a simple-minded way, namely set theory. An additional difficulty here was represented by questions of inner definability, starting from Zermelo's notion of *Definitheit*. A lot of new knowledge and logical wisdom is acquired in the process, including the formulation of a restriction axiom – a symmetric counterpart of completeness axioms on the fullness of the universe – and the discussion between Skolem and Zermelo on finite methods and impredicativity[38]; but when the set-theorist Abraham Fraenkel, in 1928[39], summarizes the state of the art of the axiomatisation

[38]See G. Lolli, "Da Zermelo a Zermelo", in *Le ragioni fisiche e le dimostrazioni matematiche*, il Mulino, Bologna, 1985, cap. VII, pp. 175-239.

[39]A. A. Fraenkel, *Einleitung in die Mengenlhere*, Springer, Berlin, 1928^3, in particular

problem, there are not many novelties with respect to the first years of the century. Fraenkel acknowledges that there are fewer results on completeness than on independence (on which he himself laboured) and, moreover,one still has the impression that the completeness notion is not always intended in a unique way. Fraenkel's own discussion is a complete mess, though we'll try to be fair in the following summary.

According to a first version, completeness of a set of axioms means that any question stated in primitive terms can have an answer, one way or the other, by means of deductive inferences from the axioms. Such a property would entail that no new axiom can be added without altering the primitive concepts; every proposition P that is not contradictory with the axioms would actually be a provable consequence, hence P is not independent. This property is to Fraenkel clearly different from Hilbert's axiom of completeness: the latter requires the non-extendibility of the domain, not of the axioms – though he admits that not all is clear. Completeness has to do with the decidability of all mathematical questions[40], and since in the last years the faith in a positive solution has been shattered, or, to say the least, a positive solution was no longer considered obvious, Fraenkel is not surprised that the completeness (e.g., of set theory) is so difficult to achieve.

The second version of completeness is rather difficult to distinguish from the first, but the way Fraenkel tries to clarify the issue – first of all to himself – is quite instructive: according to this version, we talk of incompleteness when several mutually contradictory assumptions are not derivable from a given axiom system, but are all compatible with it. The difference with the first notion seems to emerge in a remark in which Fraenkel says that an incomplete – in this second sense – axiom system does not allow to decide which of the assumptions to accept, not because of some feebleness of present deducibility methods, but rather in an absolute sense, relative to all present or future methods. As an example of a stubborn problem of this kind he considers the continuum problem. In a further, rather confusing, remark he says that in order to have completeness one does not require that the decidability of any proposition be determined, but only assured by a kind of inner determination of the domain, so that only the *Richtigkeit* or the

§18.4, pp. 347-54.

[40]We will come back later to Hilbert' 1900 axiom of the decidability of all mathematical problems.

Falschigkeit of a proposition be compatible, if not deducible from the axioms. Geometry without the parallel axiom is incomplete in the second sense, while arithmetic is not.

The third sense of completeness discussed by Fraenkel is (to him) a wider one, and he calls it categoricity, according to Veblen, or monomorphicity, in a sense attributed to Carnap and Feigl: between any two realizations it must be possible to establish a one-to-one isomorphic correspondence. The term "isomorphism" has at last found its way, according to Fraenkel, as a generalisation of that used for algebraic structures such as groups and fields – a generalisation in that it applies to any relation.

To illustrate all three cases, Fraenkel uses the same example, namely Fermat's last theorem; it is a way of stressing that the notions are tentative and not yet well-defined. If one should prove that Fermat's theorem cannot be solved with the methods of number theory (first type of incompleteness), than there would be two possibilities: either by its own nature the problem transcends the capabilities of the human mind, or else it transcends only our present capabilities. In the first case, the way should be open to the addition of new axioms, a situation similar to that of geometry; one such axiom could be given by Fermat's proposition itself. But when discussing the second type of incompleteness, Fraenkel denies that such situation can obtain for arithmetic; in the case of Fermat's theorem, the hypothesis that its negation is compatible is expressed by the possibility to think [*Denkbarkeit*] that a certain relation holds for four numbers; Fraenkel seems to be saying that the very nature of this *Denkbarkeit* lends compatibility to the negation of Fermet's theorem, whence it contradicts the *Richtigkeit* of Fermat's hypothesis; in a footnote he mentions the categoricity of the number system, as if it had some bearing on the issue, though he confesses that all is dark. Since the number system is monomorphic, Fermat's theorem should be decided, and for arithmetic and all its open problems there will never occur the situation of geometry. In another footnote, he mentions the possible difference between provability and truth, but he adds that talking of "unprovable but true" for a proposition like Fermat's has so little mathematical sense that it is impossible to discuss it.

To sum up, in Fraenkel's words, the first sense of completeness is akin to that of *Entscheidungsdefinitheit* used in Hilbert's school; Fraenkel quotes

Husserl as one of the first to conceive of a *Mannigfaltigkeit* such that any proposition stated in terms of its concepts, either is a formal consequence of the axioms or else is a formal contradiction with the axioms. In this case, for Fraenkel, truth and formal consequence are the same. The second sense seems to have to do with the distinction between what is actually provable and the possibility of a decision in principle. A system could be such that for no proposition, both the proposition itself and its negation are compatible with the axioms, but a decision method could not exist. The third sense – categoricity – is realized for natural numbers, real numbers and geometry; while categoricity implies completeness in the second sense, Fraenkel does not know whether the converse implication is true[41]. As a conclusion, Fraenkel admits that in order to make a progress in these matters, where almost nothing is really proved, a lot of work is still to be done to delimit both the proof methods, and the concept itself of completeness; symbolic logic is the most promising tool; some light could be thrown by the investigation of the completeness of the very logical system, along the lines of Emile Post's pioneering work, which was known to Fraenkel.

In his attempt to clarify various notions of completeness Fraenkel is not alone; in Felix Kaufmann's 1930 book on the infinite, where Husserl's influence is strongly felt[42], there is a chapter on the categoricity of arithmetic; significantly, its title refers to the "complete decidability" of arithmetic questions. Here Fraenkel's alternatives are presented as follows: the first notion is that of categoricity; the second one is non-ramification of a theory, in the sense that no proposition P we have compatibility pf P together with its negation; the third one is decidability, in the sense that "any question which refers to [the theory] can be decided". Kaufmann does not venture to saying whether the three concepts are different or not, but he says that they point to the same criterion, that is, to (the necessity) of a determination of a theory requiring no further specifications. This is the case of arithmetic, according to Kaufmann, since "the definition we have given of the numerical

[41]The underlying hypothesis of categoricity is perhaps responsible for the frequent identification of domains and theories, as in "Mengenbereich" and "Mengenlehre" which are synonyms in the twenties – a curious linguistic phenomenon that should be better investigated. Notice also that one talked of the real *numbers*, but of Euclidean *geometry* as domains where one got categoricity.

[42]F. Kaufmann, *Das Unendliche in der Mathematik und seine Ausschaltung*, Franz Deuticke, Wien, 1930.

series, as Peano's logical system, describes the cognitive object number and determines it as a logical singularity; this means that nothing logical concerning the natural numbers is left open"; the opposite is true for geometry. To Kaufmann, the system of arithmetic axioms is thus monomorphic. But with Fraenkel we have reached the year 1928, when Hilbert officially posed the problem of completeness; and Fraenkel mentions researches began with Post's 1921 paper, the only paper (on completeness) then known and quoted in the logical literature. So it is time to see what was happening in this field.

Meta-mathematical research had had a slow start, at first with reference to *Principia Mathematica*; the choice of a meta-mathematical level of analysis is found for the first time in Post, who however acknowledges Lewis' influence. Lewis had stressed the purely formal definition of mathematical systems, and the content-free character of the logical rules[43]; Post views PM as a purely formal construction, and considers himself entitled to use any logical and mathematical instrument to study the system. He restricts to propositional logic, where the existence of truth tables suggests the possibility of a comparison between two different methods. Post's 1921 paper[44] is his 1920 dissertation; here Post introduces the definitions of completeness and consistency that bear his name[45]. Post-consistency requires that at least one propositional letter is not derivable, in the system; Post-completeness is the property that if one adds one non-derivable formula one gets a Post-inconsistent system (the definitions are so phrased in order to apply to systems without negation). Let us remark that this is not Post's own terminology; he uses "completeness" for the functional completeness of a set of connectives, and "closed system" for Post-complete systems.

Post considers rules for the assertion of propositions, denoted by $\vdash$, as well as truth tables. The completeness of propositional logic is expressed by the "Fundamental theorem: A necessary and sufficient condition that a function of F be asserted as a result of the postulates II, III, IV is that all its truth values be $+$". (F is the set of connectives.) Post pays his debt

[43] C. I. Lewis, *A Survey of Formal Logic*, Univ. of California Press, Berkeley, 1918.

[44] E. Post, "Introduction to a general theory of elementary propositions", *Amer. J. Math.*, 43 (1921), 163-85, also in J. van Heijenoort, *From Frege to Gödel*, Harvard Univ. Press, Cambridg Mass., 1967, pp. 264-83, see introduction of J. van Heijenoort, ivi, pp. 264-65.

[45] See A. Church, *Introduction to Mathematical Logic*, Princeton Univ. Press, Princeton, 1956, § 17 and 18.

to Schröder, who had in part anticipated some of his considerations, but observes with Lewis that "formal and informal logic are inextricably bound together in Schröder's development to an extent that prevents the system as a whole to be completely determined". In particular, in the investigation of completeness one finds "all the theoretical difficulties met with in passing from the theory of classes to that of propositions when the development is not strictly formal". Again, the relevant distinction here is between formal and informal, rather than between syntax and semantics; the latter can barely be recognized beneath the distinction between a proposition and its interpretation as a class. The theorem proved by Post "gives us an actual *method for immediately writing down a formal derivation of its assertion* [of a proposition always true] *by means of the postulates of* Principia". This is what Post is interested in, namely a finite method of decision, which he will later try without success to extend to all PM.

In the meantime, first order logic was being singled out of the whole logical system of PM. Hermann Weyl seems to have been the first to define first order languages, in his dissertation written in 1910 under Hilbert's supervision and dedicated to an analysis of Zermelo's notion of *definit*. Weyl considered of "the utmost importance that the logic had furnished the list of defining principles (if we are not wrong in considering it *complete*)"; actually, he himself had given such a list. Here "completeness" seems to refer to the expressive power of first order constructs. Weyl had been much impressed by a lecture Poincaré had given in Göttingen in 1909, upon Hilbert's invitation; he became convinced of the importance of linguistic antinomies, hence of the necessity of a mathematical treatment of definability. Later, the importance of first order restrictions became clear in the formulation of the separation and replacement axioms in set theory.

In Hilbert and Ackermann's 1928 book[46], first order logic is called restricted functional calculus. One can find there a completeness proof for propositional logic along the following lines: (i) every proposition has a logically equivalent conjunctive normal form, (ii) the transformations to get normal forms can be done deductively, and (iii) a final step consisting in the remark that a clause is a tautology if and only if it contains a complementary pair (a literal and its complement). The original semantic argument is

[46] D. Hilbert, W. Ackermann, *Grundzüge der theoretischen Logik*, Springer, Berlin, 1928.

repeated in deductive form and gives completeness. The proof is due to Bernays[47], in his 1918 dissertation; Bernays didn't publish it until 1926, because (he said that) logic research was at the time held in low esteem.

As for predicate languages, Hilbert and Ackermann first discuss a strong completeness property, requiring that all statements true in some structure be provable, and dismiss it; then they consider the weaker requirement, that the logical system be complete at least in the sense that all logical formulae which are *richtig* for all domains be derivable; they admit that the problem is still open. "One can only state on an empirical basis that in all applications this system of axioms has always succeeded"[48]. Next they consider the *Entscheidungsproblem*, the decision problem for any theory, which through formalization is reduced to the decision problem for logic. They quote positive results by Löwenheim and Behmann on the monadic predicate calculus, which they obtain by means of finite structures – already used by Ackermann in his researches on decidability. Finite models allow quantifier elimination, and it is clear the authors' aim of extending with this technique the results known for propositional logic. But the right strategy was already used in another context, to which now Hilbert looks with interest, and to which we will also turn in a moment. Hilbert and Ackermann observe that the most general solution to the decision problem would be given by a procedure to determine, for every formula, for which domains – meaning for which cardinalities of the domains – the formula is satisfiable, and for which domains the formula is not; they again quote Löwenheim-Skolem theorem as an example of a restricted class, namely the class of denumerable structures, which is sufficient to check logical validity[49].

[47]P. Bernays, "Axiomatische Untersuchungen des Aussage-Kalkuls der 'Principia Mathematica' ", *Mathematische Zeitschrift*, 25 (1926), pp. 305-20. The paper contains the results of 1918 *Habilitationsschrift*.

[48]B. Dreben and J. van Heijenoort (in K. Gödel, *Collected Works*, Vol. I, Oxford Univ. Press, New York, 1986, p. 48) have remarked that there is here a circularity, in that logical formulae had been immediately before defined as the derivable ones; the flaw will be corrected in the 1938 edition; however, it is probably a slip of the pen, since the notion of *allgemeingültig* was already available; other parts of Hilbert and Ackermann's exposition are unsatisfactory: for example there is the following confusing statement, that the *Allgemeingültigkeit* of formulae, except for those that are always valid, depends on the cardinality of the domain.

[49]D. Hilbert, W. Ackermann, *Grundzüge der theoretischen Logik*, cit., p. 68.

In the same 1928, at the Bologna Congress, Hilbert calls attention to completeness as an open and urgent problem; but he is still wavering between two formulations, one more syntactic, *à la* Post, for mathematical theories, and another one for the consequence relation. The completeness problem for logic is seen as a special case of that for arithmetic; the latter is more urgent, due to the necessity of giving a finitistically satisfactory version of categoricity. "It is true that usually one states the completeness of the axiom system for number theory as well as that for analysis; but the usual argument by which one proves that any two realizations of the axiom system for number theory (resp. for analysis) must be isomorphic does not satisfy the requirements of finitary rigour. What has to be done – first of all for number theory, whose domain can be delimited with precision – is to transform the usual isomorphism proof in a finitary one, so that one can prove the following: If for a proposition S of number theory can be proved the compatibility with the axioms of number theory, then the compatibility with the axioms cannot be proved also for $\neg S$ (the opposite of S). And strictly related also: if a sentence is compatible, then it is also provable"[50].

This is problem III of Hilbert's address. Problem IV follows: "The assertion of the completeness of number theory can also be expressed in the following way: if to the axioms of number theory a formula is added which belongs to [the language of] number theory but is not provable, then from the extended system a contradiction can be derived[51]. Since here, in proof theory, we always deal with formalized proofs, in the statement on the completeness of number theory it is at the same time implicit the claim that the formalized rules of reasoning are anyway sufficient in the domain of number theory. The question of the completeness of the logical rules, in a general form, is a problem of theoretical logic". To logic one arrives at starting from

[50]D. Hilbert, "Probleme der Grundlagen der Mathematik", in *Atti Congresso internazionale dei matematici, Bologna, 3-10 ottobre 1928*, Zanichelli, Bologna, 1929, vol. I; pp. 135-41, with additions and corrections in *Mathematicshe Annalen*, 102 (1929), pp. 1-9.

[51]G. Kreisel has interpreted this statement as a first inkling of the logical completeness in this form: either F is derivable, or else a suitable arithmetical translation F_ω is contradictory with the arithmetic axioms (Gödel would change this second case in the alternative that $\neg F_\omega$ be true in the natural numbers). But this rendering seems a bit forced, since Hilbert talks here always of the formula, not of an interpretation; moreover, the completeness alluded to is explicitly that of number theory. This version of the logical completeness is to be found only later in 1939 in Hilbert and Bernays' *Grundlagen*, as we'll see in due time.

arithmetic and substituting the arithmetical predicates with arbitrary letters, and dismissing the proper axioms; one forgets about the ordered domain of numbers and one considers an arbitrary system made up of objects and associated predicates [a substitutional completeness in reverse]. Here Hilbert gives for the first time the semantic definition of logically valid formulae, "those which are not refutable by any determination of the predicates one chooses. There arises now the question whether all these formulae are provable from the rules of logical reasoning by the addition of the above mentioned equality axioms or, in other words, whether the system of usual logical rules is complete. Up to now, by trials and efforts, we have made the firm belief that these rules are sufficient. A true proof is available only for pure propositional logic. For the logic of monadic predicates a completeness proof can be obtained by the method of solution of the decision problem (Schröder elimination problem), as has been shown first by Löwenheim capitalising on first efforts of Schröder and in a definitive way by Behmann".

We have to turn now, as Hilbert did, to the other trend of the algebra of logic. Following the lead of Schröder's algebra of logic[52], researchers had first tackled the solution of logical equations, looking for necessary and sufficient solvability conditions on the coefficients; then they had focused attention to the satisfiability of logical equations, starting from finite domains[53]. While working on decidability questions, Löwenheim in 1915[54] borrows some techniques from this tradition, such as (what today we call) Skolem functions. People working in the context of the algebra of logic had discovered first order languages owing to the possibility of treating the first order existential quantifier as a disjunction over the elements of the domain: this enabled them to deal with the existential quantifier in an algebraic setting. Actually, their logic was an infinitary first order logic, also allowing infinite lists of quantifiers. Algebraic terminology was still different from ours; our first

[52] E. Schröder, *Vorlesungen über die Algebra der Logik*, 3 voll., 1890-1895, Leipzig.

[53] In early algebraic logic, the accepted laws of quantifiers were those that could be proved for a finite, though arbitrary and indeterminate, number of elements, see C. I. Lewis, *A Survey of Formal Logic*, Univ. of California Press, Berkeley, 1918, cap. IV.

[54] L. Löwenheim, "Über Möglichkeiten im Relativkalkül", *Math. Ann.*, 76 (1915), 447-70, in J. van Heijenoort, *From Frege to Gödel*, cit., pp. 228-51, see introduction of J. van Heijenoort, ivi, pp. 228-32. Löwenheim considers also the calculus of relatives – of the title – without variables but with algebraic operators for sum and product of relations, and he proves that it is not equivalent to the first order calculus.

order formulae were called numerical expressions [*Zählausdrücke*], and valid formulae were called identical equations.

In his 1915 paper Löwenheim proved the decidability of the monadic calculus by showing that if a sentence of this restricted language is true in all finite structures then it is logically true; he also showed that the result does not hold for extended languages; finally, he reduces the decision problem for predicate logic to that for sentences with only a binary relation. His results were improved by Skolem, Behmann, Herbrand, Kalmár, and were appreciated and developed also by Hilbert's school, as contributions to the *Entscheidgungsproblem*. The theorem named after him is stated by Löwenheim as follows: if a formula is valid in all finite domains but is not valid, then it is not the case that the formula is valid in all denumerable domain.

Löwenheim's theorem will become Löwenheim-Skolem theorem after Skolem's contributions and refinements. There will be two versions of the theorem: in the first version one merely talks of a denumerable model for a sentence having a model at all; in the second version one notices that the denumerable model is a substructure of the given one. This latter version is more interesting for model theory: from it the notion of elementary substructure will follow; insofar as the completeness theorem is concerned, however, there is no particular added value. The substructure version is due to Skolem in 1920, or at least it can be said that the 1920 proof implies it; the explicit statement will be given by Skolem only in 1929; later he will mistakenly attribute it to Löwenheim; only in 1938 Skolem will make a neat distinction between the following two statements:

(i) if F is satisfiable, then it is satisfiable in the natural numbers, with a suitable assignment of arithmetical predicates to the predicative letters,

and

(ii) if F is satisfiable in a domain $\mathfrak{D}$ then it is also satisfiable in a denumerable subdomain of $\mathfrak{D}$, with the same interpretation of the predicates.

While in 1920 Skolem actually proves (ii), in 1922 he proves (i). The 1922 proof is more similar to Löwenheim's old one. Moreover, Skolem sometimes uses his normal forms, other times Löwenheim's forms; notwithstanding these fluctuations, it is a widespread opinion that he was able to put Löwenheim's proof on its feet (possibly by taking the union of all his contributions). In 1920 Skolem also remarks that his result holds for denumerable sets of sentences (a remark he will correctly prove only in 1929); from 1922 on, he uses this generalized version of (i) for the applications, e.g., to set theory; it has been remarked that, had he used (ii), he could have anticipated Gödel's later discovery of submodels containing all ordinals. Version (ii) needs the axiom of choice, while (i) does not; Löwenheim had used choice, and Skolem corrected him; but all this refers to the *minutiae* of the history of Löwenheim-Skolem theorem, and it does not have much bearing on the completeness story.

In 1915 Löwenheim had used the equivalence between a first order sentence A and a second order one of the form $\exists f \forall x M(x, f(x))$, where f and x are vectors; actually, he used infinite lists of existential quantifiers, which we can translate as quantified functions; Löwenheim almost achieved (what today is known as) the normal form for satisfiability, though its form depends the finite domain, and also needs an infinite list of quantifiers in the infinite case. After eliminating the second order quantifier, Löwenheim proceeds to substitute all natural numbers in the matrix $M(x, f(x))$, following a precise strategy: first replace the variables by a finite number of constants, then enlarge the resulting finite domain so as to have values for $f(x)$, and repeat the process, each time appending the newly obtained closed matrix to the conjunction of the previous ones. At each stage Löwenheim has a proposition P_n, and he considers all ways of assigning truth values to the atomic propositions occurring in P_n. Finally, he claims that if for some n all these assignments give false value to P_n, then also the given sentence A has value false; if on the contrary for every n there is a propositional valuation under which P_n is true, then also A is true. Most commentators agree[55] that the argument is only roughly sketched, but in essence it is there, including the intuition of the necessity of some form of König's lemma, needed to obtain a single valuation pasting together the valuations of the different P_n's. Others are more doubtful, among them van Heijenoort and Quine[56] (who feels that Löwenheim argument needs a law of infinite conjunction – more or less equivalent in its effect to König's lemma); they find the correct reasoning only in Skolem; the latter uses two different tools for his arguments, Dedekind's chains in 1920, while the techniques used by him in 1922 are more reminiscent of Quine's desiderata.

[55] R. L. Vaught, "Model Theory before 1945", in L. Henkin et al. (eds.),*Proceedings of the Tarski Symposium*, AMS, Providence, R. I., 1974, pp. 153-72.

[56] W. O. Quine, "A proof procedure for quantification theory", *Journal Symbolic Logic*, 20 (1955), pp. 141-9.

In the 1920 paper, Skolem[57] addresses himself to decidability questions for Boolean algebras, elementary geometry and dense sets. For every sentence A he introduces the normal form $\forall x \exists y B$ of A for satisfiability (the Skolem normal form of A): Skolem proves that this $\forall\exists$ sentence is satisfiable if and only if A is satisfiable, in the same domain. The denumerable domain is carved out with the help of the axiom of choice. Skolem is convinced that his method is more in accord with the usual ways of mathematical logic than Löwenheim's method: once A is reduced to $\forall x \exists y B$, one chooses for every x a y such that the quantifier-free formula B is satisfied, then for such y one chooses another element witnessing the existential quantifier, and so on, thus building a chain. Then one takes the intersection of all sets closed under this operation, in a sense that will be made clearer in the next lines. Skolem's terminology is always purely semantic.

Given, in our notation, a normal form

$$\forall x_1 \ldots \forall x_m \exists y_1 \ldots \exists y_n B(x_1, \ldots, x_m, y_1, \ldots, y_n),$$

Skolem 1922[58] systematically substitutes numbers to variables: he first replaces all x's by ones, then "it must be possible to choose $y_1, \ldots, y_n$ among the numbers 1, 2, …, n+1 in such a way that $B[1, \ldots, 1, y_1, \ldots, y_n]$ is satisfied. Thus we obtain one or more solutions of the first step, that is, assignments determining the classes and relations in such a way that $B[1, \ldots, 1, y_1, \ldots, y_n]$ is satisfied. The second step consists in choosing, for $x_1, \ldots, x_m$, every permutation with repetitions of the n+ 1 numbers 1, 2, …, n+ 1 taken m at the time, with the exception of the permutation 1, 1,…, 1, already considered in the first step. For at least one of the solutions obtained in the first step, it must then be possible, for each of these $(n+1)^m - 1$ permutations, to choose $y_1, \ldots, y_n$ among the numbers $1, 2, \ldots, n+1+n((n+1)^m - 1)$ in such a way that, for each permutation

[57] T. Skolem, "Logisch-kombinatorische Untersuchungen über die Erfüllbarkeit oder Beweisbarkeit mathematischer Sätze nebst einem Theoreme über dichte Mengen", *Skr. Vid. Krist. I. Math.-Naturvid. Kl.*,(1920), No. 4, 36 pp., in J. van Heijenoort, *From Frege to Gödel*, Harvard Univ. Press, Cambridg Mass., 1967, pp. 252-63, see introduction of J. van Heijenoort, ivi, pp. 252-54.

[58] T. Skolem,"Einige Bemerkungen zur axiomatischen Begründung der Mengenlehre", *Wissenschaftliche Vorträge gehalten auf dem Fünften kongress der skandinavischen Matematiker in Helsingfors vom 4. bis 7. Juli 1922*, pp. 217-32, in J. van Heijenoort, *From Frege to Gödel*, Harvard Univ. Press, Cambridg Mass., 1967, pp. 290-301, see introduction of J. van Heijenoort, ivi, pp. 290-91.

$x_1, \ldots, x_m$ taken within the segment $1, 2, \ldots, n+1$ of the number sequence, the proposition $B(x_1, \ldots, x_m, y_1, \ldots, y_n)$ holds for a corresponding choice of $y_1, \ldots, y_n$ taken within the segment $1, 2, \ldots, n+1+n((n+1)^m-1)$. Thus from certain solutions gained in the first step we now obtain certain continuations, which constitute solutions of the second step. It must be possible to continue the process in this way indefinitely if the given first-order proposition is consistent". In order to obtain a "uniquely determined solution for the entire number sequence, we must be able to choose a single solution from among all those obtained in a given step". To past together the solutions, Skolem considers a kind of lexicographic order, in such a way that different solutions coincide on the common part; by systematicaly choosing the first element in this order, he constructs a sequence of solutions, whose limit yields the desired proof that the proposition is satisfied in the natural numbers. Skolem is interested here only in the fact that if a sentence is consistent then it has a model in the natural numbers; still, again denoting by P_n the quantifier-free propositions obtained at step n, by inspection of Skolem's proof we see that the original sentence has no model if and only if some P_n has no propositional model; today many textbooks present the above result as a version of the compactness theorem, or of the Skolem-Herbrand theorem. Only in 1928 Skolem will notice the interest of this version.

In 1928[59] Thoralf Skolem the Great writes another epoch making paper, after the 1922 one. The opening recalls and rejects Kant's opinion, that logic is the only science that has made no progress from the ancient times: in fact, after giving some examples of how equations can be solved, in Schröder's calculus, by reducing logical to combinatorial problems, Skolem turns to first-order logic. The terminology is still that of Russell, formulae are called propositional functions of the variables, taking values in a domain; for each substituted value the function becomes a proposition, which can be true or false. "Now we can not only give a precise formulations to the mathematical propositions but also to represent mathematical proofs as transformations of such logical expressions according to certain rules ...", of which he mentions – in modern notation – the equivalences $\neg\exists x A(x) \leftrightarrow$

[59]T. Skolem, "Über die mathematiche Logik", *Norsk. Mat. Tidsk.*, 10 (1928), pp. 125-42, in J. van Heijenoort, *From Frege to Gödel*, cit., pp. 508-24; see the introduction of B. Dreben and J. van Heijenoort, ivi, pp. 508-12. See also T. Skolem, "Über einige Grundlagenfragen der Mathematik", *Skr. Norsk. Akad. Oslo I. Mat.-Natur. Kl.*, (1929), No. 6, 38 pp.

$\forall x \neg A(x), \forall x(A \wedge B(x)) \leftrightarrow A \wedge \forall x B(x), \forall x(A \vee B(x)) \leftrightarrow A \vee \forall x B(x)$, and similar ones for the existential quantifier. These are the equivalences that allow to transform a sentence in prenex form. "But I do not go into this more deeply, especially since I believe that it is possible to deal with the deduction principle in another, more expedient way, to which I shall return in a moment". After a discussion of first and second order logic, which will be resumed again the following year in a discussion with Zermelo, he gives his method. "I shall not go into these difficult questions more deeply; instead, I shall indicate how the deduction problem for first-order propositions can be reduced to a problem of combinatorial arithmetic [*eine arithmetisch-kombinatorische Frage*]. If U and V are first-order propositions and if we pose the question whether V follows from U, this is equivalent to asking whether $U \wedge \neg V$ is a contradiction or not. It is therefore clear that everything depends upon our being able to decide whether a given first-order proposition is contradictory or not".

The method is the same presented today as Herbrand method (and theorem)[60] in logic textbooks for computer science. Skolem first of all recalls the normal forms for satisfiability, then by introducing new functional symbols he transforms a sentence

$$\forall x_1 \forall x_2 \ldots \exists y_1 \exists y_2 \ldots \forall z_1 \forall z_2 \ldots \exists u_1 \exists u_2 \ldots \\ U(x_1, x_2, \ldots, y_1, y_2, \ldots, z_1, z_2, \ldots, u_1, u_2, \ldots)$$

in a quantifier-free formula of the form

$$U(x_1, x_2, \ldots, f_1(x_1, x_2, \ldots), f_2(x_1, x_2, \ldots), \ldots, \\ z_1, z_2, \ldots, g_1(x_1, \ldots, z_1, \ldots), g_2(x_1, \ldots, z_1, \ldots), \ldots)$$

in such a way that "the given sentence states that the latter is true for all values of the variable in a certain domain". If the universally quantified variables are arbitrarily given "it is possible to introduce $y_1, \ldots, u_1, \ldots$, where the y depend only on the x, the u only on the x and the z, and so on, and to determine the truth values of the functions $A, B, C, \ldots$ [atomic components in U] for these arguments in such a way that U turns out to be true. Since it does not matter what notation we use for the symbols", the formalism of

[60]Responsible for the mistaken attribution is Martin Davis, in 1963, at the beginning of automated deduction; he later acknowledged the mistake, but it was late; it is ironic that in the classroom presentations so called Herbrand's theorem is entirely semantical. See M. Davis, "The Prehistory and Early History of Automated Deduction", in J. Siekmann and G. Wrightson (eds.), *Automation of Reasoning*, Springer, Berlin, 1983, pp. 1-28.

functional symbols can be introduced, and instead of y_1 it will be written $f_1(x_1, x_2, \ldots)$.

Next he describes what is now called Herbrand universe, formed by the closed terms generated in stages starting with 0 and the functional symbols. "The symbols of the $(n+1)$th level shall be those that result from the insertion of symbols up to the nth level as arguments in the 'functions' $f_1, f_2, \ldots, g_1, g_2, \ldots$ and that do not already occur among the symbols of the 0th up to the nth level" [the 0th level contains 0]. Now after substituting 0 everywhere, if it is possible to assign to $A, B, C, \ldots$ propositional values in such a way as to render U true, these values are called solutions of the first level; if it is not possible, the given proposition "is not satisfiable; there is a contradiction". Otherwise, one substitutes to the variables the terms of the first level, and then looks for an assignment to the atomic parts which gives U value 1 and which is in agreement with the previous assignment, if there are symbols in common. If there is no such an assignment, again, "we have a contradiction". One goes on indefinitely; "the real problem now is whether there are solutions of an arbitrarily high level or whether for a certain n there exists no solution of the nth level. In the latter case the given first-order proposition contains a contradiction. In the former case, on the other hand, it is consistent [*widerspruchlos*]".

The conclusion is perhaps a bit hasty, but it does not justify the criticism that has been moved to it on the basis of a later remark of Gödel. According to Gödel, Skolem is trying here to introduce syntactic considerations, and ends in a mess because of his informal ways. Instead of resorting to his 1922 argument to correctly conclude the lemma, he gives "an entirely inconclusive argument"[61]. Skolem's inconclusive argument is as follows, in the continuation of the above quotation: "The following will make this clear. Every consequence of the [proposition] results from repeated and combined uses of it. Every theorem derived can therefore be formulated as a proposition formed by means of the functions $A, B, C, \ldots$, and in this functions there will occur, on the one hand, indeterminate symbols $a, b, c, \ldots$ and, on the other, further symbols that have been obtained from these by possible repeated substitutions in the functional expressions $f_1, f_2, \ldots, g_1, g_2, \ldots$ Every such proposition must, however, retain its validity when $a, b, c, \ldots$ all are

[61]K. Gödel, letter to H. Wang of December 7, 1967, in H. Wang, *From Mathematics to Philosophy*, Humanities Press, New York, 1974, p. 8.

replaced by 0. Thus, if a contradiction is derivable, a contradiction must be provable in which there occur both 0 and the symbols obtained from 0 by substitution in $f_1, f_2, \ldots, g_1, g_2, \ldots$ up to, say, the nth level. Hence there cannot then exist any solution of the nth level". It is commonly accepted that Skolem is trying here – without success – to formulate a syntactic lemma, the very lemma which, if only it had been formulated, would have anticipated Herbrand's thesis. However, this is not entirely convincing, if we bear in mind the following facts: (i) Skolem terminology as far as *Widerspruchlosigkeit* is concerned, is always semantic; (ii) this is just a lecture, and the 1922 argument need not be incompatible with the present sketched considerations; last but not least, (iii) from the very beginning Skolem had stated that he was not interested in using and refining the syntactic notion of derivation, since he was interested in an alternative method.

Skolem was well conversant with deductive first order methods; he did not accept them as a foundation for mathematics, because he believed that "the attempt to base the notions of logic upon those of arithmetic, or vice versa, seems to me mistaken. The foundations for both must be laid simultaneously and in an interrelated way". The above quoted remark, about consequences being built up from the atomic formulae in the premises, shows a keen attention to finer details of deductive processes. But Skolem was looking for a method to establish logical consequence, and he believed that the combinatorial (nonsemantical) reduction presented in his paper was more expedient. The problem he addresses in 1928 is more general than the refinement of Löwenheim's theorem: it is the quest for an effective procedure to establish logical consequence; owing to this explicitly stated goal, we are entitled to attribute to Skolem the intention to prove (an equivalent of) the completeness theorem. His method starts from semantic notions, and uses combinatorial conditions to solve the deducibility problem. Skolem reaches a satisfactory conclusion, which we call a completeness theorem, though he gives the impression of halting in the middle of the river; the same, however, happens to modern proofs, when one starts from syntactic notions and arrives to something one does not know whether it is syntax or semantics[62]. Moreover, Skolem's approach explains the strong connection, also perceived in Hilbert's school, between completeness and decidability; in both cases the problem was that of finding an effective method for logical consequence, pos-

[62] As, for example, in the proofs of completeness by refutation trees

sibly different from the usual available logic systems, which in themselves have no peculiar privilege.

"To be sure, this procedure is infinite; but there are some cases in which it is possible to make the procedure finite". Skolem solves the decision problem for the prefix $\forall y \exists x_1 \ldots \exists x_n$ (Ackermann had the same result, which Skolem improves in a review, with a better estimate of the number of individuals in the finite domain in which such formulae are satisfiable; Herbrand later will also consider and improve the bounds for the $\exists x_1 \ldots \exists x_n \forall y \exists z_1 \ldots \exists z_m$ case). At the end of his lecture Skolem describes Langford's theorem on the decidability of dense orderings, obtained by means of quantifier elimination[63].

Now in the history of the Löwenheim-Skolem theorem it should be the turn of Herbrand, with the birth of the Skolem-Herbrand theorem. Herbrand used syntactic methods, and more than that, he used *only* syntactic methods; but Herbrand's results did not become known until after Gödel's. Gödel's dissertation was approved in June 1929, and discussed in February 1930; Herbrand's dissertation dates back to April 1929, and was discussed in June 1929; after September 1929, it seems that Herbrand did no work on it until the middle of 1930. His 1931 paper contains some corrections on the thesis, dating from September 1929. Apart from this chronology, Gödel's work explicitly aimed at a proof of the completeness theorem; Herbrand's goal was to give a syntactic, or finitistic, version of Löwenheim theorem. We shall consider first Gödel's proof.

Before plunging into the last act, it may be useful to pause to sum up all uncertainties and ambiguities that Gödel's theorem would finally dispel. It is not true that everybody was convinced of completeness: somebody, like Weyl, feared it. Already in 1917, talking of the completeness of the real numbers, he had confessed that "we do not know if it is so (perhaps we believe it)". Weyl believed in completeness, but was also afraid of it, since he wrote, with reference to elementary geometry: "the firm belief to be able to derive all general true judgements of elementary geometry ... from the geometrical axioms by means of a logical reasoning, is a scientific *faith declaration*: we are not able to have a true *intuition* that it is so, and still less to 'prove' it by an examination of the logical laws. If this should happen some day,

[63] C. H. Langford, "Some theorems on deducibility", *Ann. of Math.*, 28 (1926), pp. 16-40, "Theorems on deducibility", *Ann. of Math.*, 29 (1927), pp. 459-71.

this intuition would open to us the way to decide the truth or falsity of any geometrical judgement . . . by methodically applying a certain deductive technique ('in a finite number of steps'): mathematics would thus become *trivial*, at least in principle[64]". In 1926, while discussing the fundamental concepts of axiomatics, such as independence and completeness of the axioms, Weyl again wrote: "completeness would be assured only by establishing for the development of the proofs such rules as would lead automatically to the solution of any pertinent problem. Mathematics would become an unexciting enterprise. But such a philosophical stone has not been discovered, and it will never be". The root of such a strong belief is simply the subjective feeling that doing mathematics is not a matter of machine applications; in practice "it is not possible to proceed as the scholar Gulliver finds at Barnilarbi, who develops in order all the consequences, to thorw away later the non-interesting one[65]". Weyl was afraid of the procedure that was later to be called, after Alan Turing, the "British Museum procedure". This dreadful procedure – which does not seem to have been considered by Hilbert's school – is not practically feasible, but Weyl correctly notes that it guarantees the decidability of a theory. Still probably unable to distinguish between logical completeness and completeness of theories, Weyl feared however that mathematics would be thus made trivial.

This is reminiscent of the old argument of Poincaré, now reinforced by the (partial) results on decidability. The *Entscheidungsproblem* was one of those that could be rigorously attacked in the frame of Hilbert's formalization and proof theory. It was the problem of the decidability of mathematical problems by means of a finite number of operations. In the general opinion, the problem evoked (the belief and) the axiom of the solvability of all mathematical problems, as it had been expressed by Hilbert at the Paris Congress in 1900. Hilbert has been careful in the precise wording of his conjecture, "that every definite mathematical problem must necessarily be susceptible of an exact settlement, either in the form of an actual answer to the question asked, or by the proof of the impossibility of its solution and therewith the necessary failure of all attempts". Hilbert had drawn his conviction ("which every mathematician shares, but which no one has yet supported by a proof")

[64]H. Weyl, *Das Kontinuum. Kritische Untersuchungen über die Grundlagen der Analysis*, Veit, Leipzig, 1918.

[65]In an entry for the *Handbuch der Philosophie*, 1926, later included in H. Weyl, *Philosophy of Mathematics and Natural Science*, Princeton Univ. Press, Princeton, 1949.

by the very negative cases – such as the fifth postulate, and the solution by radicals of algebraic equations. In all these cases, a proof had been found that the problem was not solvable on the basis of the assumptions explicitly stated in the conditions of the problem. In 1925 he again will repeat that "now, to be sure, my proof theory cannot specify a general method for solving every mathematical problem; that does not exist"[66]. He had no intention of trivializing mathematics. However, the misunderstanding was strong to die; even von Neumann mixed ethical worries with theoretical doubts, as Weyl did; in 1927[67] von Neumann didn't believe in a positive solution of the *Entscheidungsproblem*, because he was morally certain of the undecidability of mathematics: "undecidability is a *sine qua non* condition for the present work with heuristic methods having sense; the day undecidability were eliminated, mathematics as it is intended today would not exist any more, and a purely mechanical process would take its place".

After proving the completeness of the logical rules, Gödel will feel entitled to claim as a well known fact, in 1930, that "the development of mathematics toward greater precision has led, as is well konwn, to the formalization of large tracts of it, so that one can prove any theorem using nothing but a few mechanical rules"[68]. Russell's *Principia Mathematica* and Zermelo and Fraenkel's set theory are "so comprehensive that in them all methods of proofs today used in athematics are formalized, that is, reduced to a few axioms and rules of inference". Actually, not everybody was aware of this possibility, since he had just proved it in his dissertation the previous year; he had spoken of the result only at the Menger's *Mathematische Kolloquium* in May 1930. Gödel is not afraid of formalization and of restricted logic methods; but he only makes this explicit at the same time when he also announces the incompleteness of PM, thus banning every fear of trivialization: "One might therefore conjecture that these axioms and rules of inference are sufficient to decide any mathematical question that can at all be expressed in these systems. It will be shown below that this is not the case, that on the

[66]D. Hilbert, "Über das Unendliche", *Mathematische Annalen*, 95 (1926), pp. 161-90, in J. van Heijenoort, pp. 367-92.

[67]J. von Neumann, "Zur Hilbertschen Beweistheorie", *Mathematische Zeitschrift* , 26 (1927), pp. 1-16, p. 10.

[68]K. Gödel, "Über formal unentscheidbare Sätze der Principia Mathematica und verwandter Systeme I", *Monatshefte für Mathematik und Physik*, 38 (1931), pp. 173-98, in K. Gödel, *Collected Works*, cit., pp. 144-95.

contrary there are in the two systems mentioned relatively simple problems in the theory of integers that cannot be decided on the basis of the axioms". Notice that Gödel doesn't use for his system the term "incomplete", but "*nicht* entscheidungsdefinit"[69].

Before obtaining his incompleteness proof, Gödel, too, had had some doubts. In the introduction to his dissertation, he lingers over the relationship between categoricity, completeness and decidability of all problems: "Here 'completeness' is to mean that every valid formula expressible in the restricted functional calculus (a valid *Zählaussage*, as Löwenheim would say) can be derived from the calculus by means of a finite sequence of formal inferences. This assertion can easily be seen to be equivalent to the following: Every consistent axiom system consisting of only *Zählaussagen* has a realization. (Here 'consistent' means that no contradiction can be derived by means of finitely many formal inferences.) The latter formulation seems also to be of some interest in itself, since the solution of this question represents in a certain sense a theoretical completion of the usual method for proving consistency (only, of course, for the special kind of axiom systems considered here)". The existence of a realization for a consistent theory, now guaranteed by the theorem, offers support and justification to the axiomatic position[70]. But – Gödel goes on brooding – to identify consistency and existence, or to make the latter depend only on consistency constraints, seems to presuppose "the axiom that every mathematical problem is solvable. Or, more precisely, it presupposes that we cannot prove the unsovability of any problem". A problem requires an answer, either positive or negative; if one could prove that a problem is unsolvable, one would prove in particular that neither the positive nor the negative answer are derivable form the axioms; hence both would be consistent with the theory, and there would be at least two non-isomorphic models, contrary to what one can prove for the theory. Clearly, Gödel has in mind categorical theories, such as the theory of real numbers, that is explicitly mentioned by him.

In the same introduction, Gödel's way out reminds us of Hilbert's old

[69]K.Gödel, "Einige metamathematische Resultate über Entscheidungsdefinitheit und Widerspruchsfreiheit", *Anzeiger der Akademie der Wissenschaften in Wien*, 67, pp. 214-5, in K. Gödel, *Collected Works*, cit., pp. 140-43.

[70]A position still as recently as 1928 contrasted by L. E. Brouwer, in a Wien lecture apparently attended to by Gödel, who was strongly impressed.

reflections on proof methods: ultimately, the completeness result does not imply the impossibility of finding unsolvable problems since "what is at issue here in only unsolvability by certain *precisely stated formal* means of inference. For, all the notions that are considered here (provable, consistent, and so on) have an exact meaning only when we have precisely delimited the means of inference that are admitted. These reflections, incidentally, are intended only to properly illuminate the difficulties that would be connected with such a definition of the notion of existence, without any definite assertion being made about its possibility or impossibility". Gödel did not have at this time an example of an unsolvable problem in number theory, but he wanted to leave an open door; the other alternative was contrary to his sensibility more that to his reason, which was still debating without a solution. It is to be noted that he cancelled all these reflections in the 1930 printed version of the dissertation.

Gödel's proof[71] is similar to Skolem's 1922 proof; in a remembrance of the sixties, Gödel will say that he did not know Skolem 1922 when he proved his result; Gödel had some recollection of having read Skolem's proof only later, at the moment of printing his dissertation, and he wondered why he didn't quote it; it probably was an oversight or a misprint, since he was certain he would have made a reference to Skolem 1922, whose proof was more similar to his own than that of the 1920 paper – dutifully quoted for Skolem's normal forms. "The completeness theorem, mathematically, is indeed an almost trivial consequence of Skolem 1922. However, the fact is that, at that time, nobody (including Skolem himself) drew this conclusion (neither from Skolem 1922 nor, as I did, from similar considerations of his own)"[72].

Gödel's proof, in effect, differs from Skolem's only in this, that the formal system is well specified – a first order fragment of PM in Hilbert and Ackermann's notation. Given the normal form $\forall x M(x, f(x))$ of a sentence

[71]K. Gödel, "Über die Vollständigkeit des Logikkalküls", Univ. Wien, 1929, in K. Gödel, *Collected Works*, cit., pp. 60-101; K. Gödel, "Die Vollständigkeit der Axiome des logischen Funktionenkalküls", *Monatshefte für Mathematik und Physik*, 37 (1930), pp. 349-60, in K. Gödel, *Collected Works*, cit., pp. 102-23. See here the introduction of B.Dreben and J. van Heijenoort, pp. 44-59. There are no significant differences in the two versions regarding the proof; in the published paper the denumerable compactness theorem is more clearly stressed.

[72]Letter of December 7, 1967, in H. Wang, *From Mathematics to Philosophy*, cit.4, p. 8.

F and the conjunctions $A_n = M_1 \wedge \ldots \wedge M_n$ obtained through systematic substitutions of terms in the matrix M, instead of using more or less conclusive semantical considerations, Gödel proves that for every n the implication $F \rightarrow |existsA_n$ is derivable in the formal system. For the other direction, Gödel gives a new proof of König's lemma (proved by König in 1926) referring to it as "a familiar argument". Löwenheim-Skolem theorem of course becomes now a corollary of the completeness theorem[73]. The only general remark by Gödel is that the equivalence between validity and provability is a kind of reduction, for the decision problem, of the nondenumerable to the denumerable; "validity" refers to the more than denumerable totality of functions, while "provable" presupposes only the denumerable totality of formal proofs.

As for Herbrand's dissertation[74], Hilbert and Bernays will say in 1939 that it is difficult to understand, and the same judgement is valid today, notwithstanding later corrections; Gödel remembered to have found some errors in it in the forties; in 1963 B. Dreben, P. Andrews and S. Aanderaa filled all the gaps[75], but still, the very formulation of Herbrand's goal is hard to grasp. Herbrand talks of true propositions, but he means propositions provable in his first order logical system Q_H; sometimes he calls them identities, and he says he is looking for "the most general properties that are sufficiebt for a proposition to be true". He is looking for a finitistic reduction of the truth problem, that is, of the derivability problem. His problem is different, but it is cast in the same terms as Skolem 1928's; his rules, too, are exactly those presented by Skolem, except for the fact that Herbrand considers

[73]So it is now presented in the textbooks, see e.g. S.C. Kleene, *Introduction to Metamathematics*, Van Nostrand, New York, 1952, Corollary 2, p. 394.

[74]J. Herbrand, "Recherches sur la théorie de la démonstration", *Trav. Soc. Sci. Lett. Varsovie*, Cl. III Sci. Math. Phys. No. 33 (1930), 128 pp.; in J. Herbrand, *Ecrits Logiques*, PUF, Paris, 1968, pp. 35-133; chap. 5 in J. van Heijenoort, *From Frege to Gödel*, Harvard Univ. Press, Cambridge Mass., 1967, pp. 525-81, see notes of B. Dreben and introduction of J. van Heijenoort, ivi, pp. 525-29.

[75]B. Dreben, P. Andrews and S. Aanderaa, "False lemmas in Herbrand", *Bull. AMS*, 69 (1963), pp. 699-706. The relationship between Herbrand's theorem and Gentzen's "verschäfter Haupsatz" is difficult to pinpoint, also for proof-theory specialists: to Gentzen, Herbrand's is a particular case of his own, with empty antecedent and only one prenex formula in the consequent; but Gentzen's *Hauptsatz* holds only for prenex formulae, though it is extendable to intuitionistic logic; both give cut elimination; Herbrand's theorem is perhaps more informative on the *Mittelsequenz*.

also conjunction, and he states the rules in double form, in order to have commutativity.

As we shall see, Herbrand also aims at proving Löwenheim's theorem, in a finitistically acceptable form; among the consequences of his main result, he mentions the completeness problem, in the following way: "if to our rules of reasoning, listed at the beginning of Chapter 2, we were to adjoin other rules that could not be derived from them, then we would be led to regard as true some propositions that are in fact false in some infinite domain. We must acknowledge that such a consequence would be difficult to accept. This fact [in 1931 "theorem" instead of "fact"] corresponds to what the Germans call the *Vollständigkeit* of our system of rules. (If we could prove that these additional rules lead us to regard as true a proposition P that, without them, would not be so, then, as we can readily see, the inconsistency of classical mathematics would follow, because we could construct a denumerable set over which P would be false.)". In 1931[76] Herbrand again writes, in more detail: "If we assume that the decision problem has been solved and if we do not wish R + Infin Ax + Mult Ax [simple theory of types with the axiom of infinity and the multiplicative axiom] to be inconsistent, we must not add any rule of reasoning to those already considered". When he got acquainted with Gödel's result he added an Appendix, date of April 1931, in which he wrote: "If the decision problem is solved for a proposition P, if the solution is formalizable in T + Infin Ax + Mult Ax (as it is for all the particular cases of the decision problem solved up to now) and if P is not an identity, then

(1) No new rule of reasoning that makes P an identity can be added without entailing a contradiction in R + Infin Ax + Mult Ax;

(2) P cannot be true in R + Infin Ax + Mult Ax".

In 1931 Herbrand is still using Post version of the completeness problem. at a variance with Hilbert's Bologna speech, where this formulation was used ony for the completeness of arithmetic, not for logic; but, according to some commentators, his insistence on the formalization of the solution of the decision problem is a hint that perhaps he is thinking of the strong version of the problem, later to be called finitistic rendering of the completeness

[76] J. Herbrand, "Sur le problème fondamentale de la logique mathématique", *Spraw. z pos. Towarsz. Nauk. Warszawskiego, W.*, III, 24 (1931), pp. 12-56, in J. Herbrand, *Ecrits Logiques*, cit., pp. 167-207.

theorem[77].

In his dissertation Herbrand seems to be aiming not so much at new results, but rather, at setting a clear distinction between the scope of (finitary) meta-mathematics, and that of (infinitary) mathematics. He explains how the consequences of his theorem tinclude the following results:

"Theorem 1. If P is an identity, $\neg P$ is not true in any infinite domain;

Theorem 2. If P is not an identity, we can construct an infinite domain in which $\neg P$is true.

Similar results have already been stated by Löwenheim (1915), but his proofs, it seems to us, are totally insufficient for our purposes. First, he gives an intuitive meaning to the notion 'true in an infinite domain', hence his proof of Theorem 2 does not attain the rigour that we deem desirable ... Then – and this is the gravest reproach – because of the intuitive meaning that he gives to this notion, he seems to regard Theorem 1 as obvious. This is absolutely impermissible; such an attitude would lead us, for example, to regard the consistency of arithmetic as obvious. On the contrary, it is precisely the proof of this theorem ... that presented us with the greatest difficulties. One could say that Löwenheim's proof was sufficient in mathematics; but, in the present work, we had to make it 'meta-mathematical' so that it would be of some use to us". Similar reproaches are addressed to Ackermann for his recent use of the satisfiability notion. In the existing mathematical proofs of the above theorems, one was only allowed to obtain a system of values in an infinite domain via a principle of choice [König's lemma]; to avoid this use, one has to finitistically define the construction of an infinite domain; this is why Herbrand introduces his term-models, and deals with them in the most careful way.

"We shall say that we have an *infinite domain* if we have a definite procedure for correlating with every number p: first a domain C' that contains C_1 [given at the beginning and finite]; then, for the functions, a system of values in C' that permits us to obtain in C' the value of any function of height not greater than p whenever the arguments are taken in C_1; and, finally, a system of logical values that are assigned to the atomic propositional functions ..." (in Herbrand's procedure, some troubles are due to the elimination of

[77]So thinks Dreben, in van Heijenoort, cit., who refers to D. Hilbert, P. Bernays, *Grundlagen der Mathematik*, vol. II, Springer, Berlin, 1939, pp. 243-63; see later.

quantifiers *in itinere*, from which some errors in the indexing crop up). "We shall say that P is *true in the infinite domain* if for every number p we have a procedure enabling us to verify that each of the propositions obtained over D_p [system of values at stage p] has 'true' as its logical value (derived from the logical values that its atomic propositions take in the domain) ... It is absolutely necessary to adopt such definitions if we want to give a precise sense to the words 'true in an infinite domain', words that have frequently been used without sufficient explanation, and also if we want to justify a proposition proved by Löwenheim, a proposition to which many refer without clearly seeing that Löwenheim's proof is totally inadequate for our purposes and that, indeed, the proposition has no precise sense until such a definition has been given". As a matter of fact, rather than proofs being defective, the problem here is that the definitions are unsatisfactory; accordingly, Herbrand is not really concerned to strengthen Löwenheim theorem, but to find a suitable re-wording for it.

Differently from all his predecessors (Löwenheim, Skolem and Gödel) and notwithstanding the similarity of the procedure for obtaining the sequence of the sentences P_n, what Herbrand actually proves is the following: when for all n P_n is propositionally satisfiable, then (it is not the case that the original F is satisfiable in a denumerable domain, but) the negation of F is not provable finitistically. Alternatively, if there is an n for which P_n is not propositionally satisfiable, Herbrand shows that the negation of F has a proof, in a cut free formal system with the subformula property[78].

Gödel has explained that Skolem's confusion in 1928 was due to the fact that "non-finitary reasoning in mathematics was widely considered to be meaninglful only to the extent to which it can be "interpreted' or 'justified' in terms of finitary meta-mathematics"[79]. While the key word here is in

[78]Herbrand defines "P has property B of order p" if the proposition obtained by substitution of terms of D_p is an identity, that is derivable in Q_H. Then he proves that if P is false in some infinite domain then it cannot have property B for any p, and conversely if for every p a proposition P does not have property B of order p then we can construct an infinite domain in which P is false. The fundamental theorem states that if for some p P has property B of order p then P is true, and given p one can construct a proof of P; and if P is true and we have a proof of P then we can find a p for which P has property B of order p.

[79]K. Gödel, letter to H. Wang of December 7, 1967, in H. Wang, *From Mathematics to Philosophy*, cit., p. 8.

fact “non-finitary”, this remark is better suited for Herbrand. The same applies to Gödel’s remarks contained in a letter of 1970, to the effect that “in consequence of the philosophical prejudices of our time: 1. nobody was looking for a relative consistency proof because it was considered axiomatic that a consistency proof must be finitary in order to make sense, 2. a concept of mathematical truth as opposed to demonstrability was viewed with general suspicion and widely rejected as meaningless”[80]. It can be worth to recall an episode reported by A. Mostowski, that when Tarski saw notice of Gödel’s theorem in the *Monatshefte*, he was sceptical about it, because Gödel didn’t define validity[81]; however, as regards the completeness theorem this definition was not necessary, the arithmetic interpretation being sufficient.

An improvement of the completeness theorem was obtained by Hilbert and Bernays in 1939[82], through the formalization of Gödel’s proof; they proved thus that every syntactically consistent sentence F has an arithmetic model, in the sense that the sentence becomes true if its predicative symbols are interpreted on arithmetical relations; stated otherwise, a sentence F_ω obtained by replacing the predicative symbols of F with arithmetic formulae, also becomes true; these arithmetic formulae can be taken of complexity Δ_2. Hilbert and Bernays notice that if F is not refutable then not only F_ω is satisfiable in the natural numbers, but it is derivable in an arithmetic system in which to the usual axioms one adds the formalized statement of the non-refutability of F, or $Con(F)$[83]. Therefore, addition to arithmetic of the arithmetical interpretation F_ω of a logically unprovable F would render arithmetic ω-inconsistent[84]. This property is to Hilbert and Bernays *einer Art von deduktiver Abgeschossenheit*, even more so for arithmetic then for logic; in some sense, this is the best possible result, in the light of Gödel’s incompleteness theorem.

[80]H. Wang, *Reflections on Kurt Gödel*, The MIT Press, Boston, 1985, p. 85.

[81]So G. Kreisel, in “Gödel’s excursions into intuitionistic logic”, in P. Weingartner and L. Schmetered (eds.), *Gödel remembered*, Bibliopolis, Napoli, 1987, Appendix I, pp. 131-31.

[82]D. Hilbert, P. Bernays, *Grundlagen der Mathematik*, vol. II, cit., pp. 263-63. See also S. C. Kleene, *Introduction to Metamathematics*, cit., Theorem 35, p. 394.

[83]This is “Bernays’s lemma”, according to Wang’s generalization in H. Wang, “Arithmetic models for formal systems”, *Methodos*, 3 (1951), pp. 217-32.

[84]D. Hilbert, P. Bernays, *Grundlagen der Mathematik*, vol. II, Springer, Berlin, 1939, pp. 252-3, and S.C. Kleene, *Introduction to Metamathematics*, Van Nostrand, New York, 1952, Theorem 36, p. 395.

The history of the completeness theorem has, to say the least, one more episode, with Leon Henkin's proof of 1949-1950[85]. This proof is free from the constraint of the natural numbers, and uses more general techniques by exploiting the equivalence of models and maximally consistent rich extensions. Other methods will be introduced around 1955[86] by Beth, Hintikka, Schütte, through a simplification of Henkin's proof, obtained by looking for minimal closure conditions which are strictly necessary for the construction of the model. Among the by-products of these proofs, one should mention the tableaux method and the method of partial valuations in Hintikka sets. Further, Henkin's proof deserves credit also for remaining valid also for the theory of types: indeed, this proof is a major step in the clarification of several issues concerning the status of higher order logics[87].

[85]L. Henkin, "The completeness of the first-order functional calculus", *Journal Symbolic Logic*, 14 (1949), pp.159-66, "Completeness in the theory of types", *Journal Symbolic Logic*, 15 (1950), pp. 81-91.

[86]E.W. Beth, "Semantic Entailment and Formal Derivability", *Mededelingen van der Koninlijke Nederlandse Akademie van Weteschappen*, 18 (1955), pp. 309-42, in J. Hintikka (ed.), *The Philosophy of Mathematics*, Oxford Univ. Press, Oxford, 1969, pp. 9-41; J. Hintikka, "Distributive Normal Forms in the Calculus of Predicates", *Acta Philosophica Phennica*, Fasc. VI, Helsinki, 1953.

[87]The author wants to thank the (not too) anonymous referees who took on the painful job of trying to improve his English.