Reflections on Incompleteness*

Giuseppe Longo

Laboratoire d'Informatique
CNRS et École Normale Supérieure
45, Rue D'Ulm
75005 Paris (France)
http://www.di.ens.fr/users/longo

Introduction: some history, some philosophy

At the age of 7 or 8, Gauss was asked to produce the result of the sum of the first n integers (or, perhaps, the question was slightly less general ...). He then proved a theorem, by the following method:

$$\begin{array}{cccc} 1 & 2 & ... & n \\ n & (n\text{-}1) & ... & 1 \\ \hline (n+1) & (n+1) & ... & (n+1) \end{array}$$

which gives $\Sigma_1^n i = n(n+1)/2$.

Clearly, the proof is not by induction. Given n, a uniform argument is proposed, which works for any integer n. Following Herbrand, we will call **prototype** this kind of proof. Of course, once the formula is known, it is very easy to prove it by induction, as well. But, one must know the formula, or, more generally, the "induction load". A non-obvious issue in automatic theorem proving, as we all know.

Let's now speculate on the possible "cognitive" path which "brings to" (and gives certainty!) to this proof. The reader can surely see, in his mental spaces, the "number line", that is the well-ordered sequence of integer numbers. They are there, one after the other, in increasing order: you may see it on a straight line, it may oscillate, but it should be, for you, from left to right (isn't it? please check ... and give up doing mathematics, if you do not see the number line; see [Dehaene98] for some data about it). It seems as if little Gauss has had the mental courage to put it on paper and then ... reverse it. No induction, just the order and its inverse, and the proof works for any n, a perfectly rigorous proof.

Consider now a non-empty subset in your number line. You can surely see that this set has a least element (look and see: if there is an element, there is a least one among the finitely many preceding it, even if you may not know which one). The "observation" imposes itself, by the (well-)ordering of the line,

* Dedicated to J.-Y. Girard, in his 60th birthday, September 9, 2007.
A preliminary version of this paper was an Invited Lecture at **Types for Proofs and Programs**, Durham, (GB), Dec. 2000; **Lecture Notes in Computer Science**, vol 2277 (Callaghan et al. eds), pp. 160 - 180, Springer, 2002.

as geometric evidence, a very robust one. Moreover, one does not need to know if and how the subset eventually goes to infinity: if it has one point somewhere (the set is not empty), this is at some finite point and, then, there is a smaller one which is the least of the "given" subset. In the conclusion, we will call this, a "geometric judgement".

In the few lines above, we hinted to an understanding of the ordering of numbers with reference to a mental construction, in space (or time). Frege would have called this approach "psychologism" (Herbart's style, according to his 1884 book). Poincaré instead could be a reference for this view on certainty and meaning of induction as grounded on intuition, possibly of space. In Brouwer's foundational proposal as well, the mathematician's intuition of the sequence of natural numbers, which founds Mathematics, relies on a phenomenal experience; however, this experience should be grounded on the "discrete falling apart of time", as "twoness" ("the falling apart of a life moment into two distinct things, one which gives way to the other, but is retained by memory", [Brouwer48]). Thus, "Brouwer's number line" originates from (a discrete form of) phenomenal time and induction derives meaning and certainty from it.

Intuition of ordering in space or time, actually of both, contributes to establish the number line, as an invariant of these active experiences: formal induction follows from, it doesn't found this intuition. This is my understanding of Poincaré's and Brouwer's philosophy. By recent scientific evidence (see [Dehaene98]), we seem to use extensively, in reasoning and computations, the "intuitive" number line; these neuropsychological investigations are remarkable facts, since they take us beyond the "introspection" that the founding fathers used as the only way to ground mathematics on intuition. We are probably along the lines of transforming the analysis of intuition from naive introspection to a scientific, objective, investigation of our cognitive performances.

Let's now go back to ... the sum of the first n integers. About eighty years later, Peano and Dedekind suggested that a proof, such as little Gauss', was certainly a remarkable achievement (in particular for such a young man), but that one had to prove theorems, in Number Theory, by some sort of "formal and uniform method", defined as a "potentially mechanisable" one, insisted Peano and Padoa. Then, they definitely specified "formal induction" as THE proof principle for Arithmetic (Peano Arithmetic, PA).

Frege set induction at the basis of his logical approach to mathematics; he considered it a key logical principle, and gave by this to PA the founding status that it still has. Of course, Frege thought that logical induction (or PA) was "categorical" (to put it in modern terms), that is that induction captured exactly the theory of numbers, or that everything was said within PA: this logical theory simply coincided, in his view, with the structure and properties of numbers (Frege didn't even make the distinction "theory vs. model" and never accepted it: the logic was exactly the mathematics).

We all know how the story continues. In his 1899 book (The Foundation of Geometry), Hilbert set geometry on formal grounds, as a solution of the incredible situation where many claimed that the rigid bodies could be not so

rigid, that light rays could go along (possibly curved) geodetics .... Riemann's habilitation (under Gauss' supervision), in 1854 ([Riemann54]), had started this "delirium", as Frege had called the intuitive–spatial meaning of the axiom for geometry, [Frege84], p.20. Helmholtz, Clifford, Poincaré had insisted on this idea of Riemann's and on its possible relevance for the understanding physical action at distance (gravitation, in particular): "in the physical world nothing else takes place, but (continuous) variations of curvature of space" W.Clifford (1882 (!!)). For these mathematicians, meaning, as reference to phenomenal space, and its mathematical structuring preceded rigor and provided "foundation", see [Boi95], [Bottazzini95]: by mathematics, geometry in particular, they wanted to make the physical world intelligible, more then just deriving theorems by rigorous tools as formal/mechanical games of symbols (more on the connections between proof principles in Mathematics and in Physics is in [BaillyLongo06]). Hilbert had a very different foundational attitude: for the purposes of foundations (but only for these purposes), forget the meaning in physical spaces of the axioms of non-Euclidean geometries and interpret their purely formal presentation in PA. And his 1899 book contains one of the earliest and most remarkable achievements in "programming": he fully formalized a unified approach to geometry, by closely analysing several relative consistency issues, and "compiled" it in PA, by analytic tools. Formal rigor and effective-finitistic reduction are at the core of it.

Thus, on one hand, the geometrization of physics, from Riemann to Einstein and Weyl (via Helmholtz, Clifford and Poincaré), brought to a revolution in that discipline, originating by breathtaking physico-mathematical theories (and theorems); on the other, the attention to formal, potentially mechanisable rigor, independent of meaning and intuition, gave us fantastic formal machines, from Peano and Hilbert to Turing and our digital computers.

The following year, at the 1900 Paris conference, Hilbert definitely contributed to give to PA (and to formal induction) their central status in foundation, by suggesting to prove (formally) the consistency of PA: then the consistency of the geometric axiomatizations would have followed from that of formal Number Theory (with no need of reference to meaning, in time, in space or whatever). Moreover, a few years later, he proposed a further conjecture, the "final" solution to all foundational problems, a jump into perfect rigor: prove the completeness of the formal axioms for Arithmetic. Independently of the heuristics of a proof, its certainty had to be ultimately given by formal induction.

However, there was more than this in the attitudes of many at the time. That is, besides foundation as "a-posteriori formalization", the "potential mechanization" of mathematics was truly dreamed, not only as a locus for certainty, but also as a "complete" method for proving theorems (as mentioned above, the Italian logic school firmly insisted on this, with their "pasigraphy", a universal formal language, a mechanisable algebra for all aspects of human reasoning). Or, the "sausage machine" for mathematics (and thought), as Poincaré ironically called it, could be put at work: provide pigs (or axioms) as input, produce sausages (or

theorems) as output[1]. We know how the story of complete a-posteriori formalization and, a fortiori, of potential mechanization ended ... Hilbert's conjectures on the formally provable consistency, decidability and completeness of PA turned out to be all wrong, and the 1931 proof of this fact originated (incomplete, but) fantastic formal machines (by an early rigorous definition of "computable function"). More generally, Gödel's negative result started a major deepening of Logic: besides Recursion Theory, also Model Theory (the fact that not all models of PA are elementary equivalent strongly motivates futher investigations) and Proof Theory (Gentzen) had a new start (negative results matter immensely in Science, see [Longo06]). The later lead to the results by Girard and Friedman we analyze below.

As for Number Theory, the main consequence is that formal induction is incomplete and that one cannot avoid infinitary machinery in proofs (in the rigorous sense of Friedman, see [Friedman97], for example). In some cases, this can be described in terms of the structure of "prototype proofs", as it will be proposed below. Moreover, even the problem of the induction load, or of the prototype proof in the inductive step, is a non-trivial issue in actual mechanization. Clearly, a posteriori, the induction load may be generally described within the formalism, but its "choice", out of infinitely many possible ones, may require some external euristics (typically: analogies).

The aim of this paper is to focus on some specific limits of formalization, by a close analysis of some "concrete" but formally unprovable number-theoretic statements. This is also done in order to encourage a broadening of the tools for proofs, and stress the role of "interaction" man/machine in proof-assistants and proof-checking: singling out the fully un-formalizable fragments is a crucial component of the work in these areas. It may help to pass over to the machine exactly the fully formalizable parts. Beyond the myth, and in full awareness of the incompleteness of formalisms, we may further develop these remarkable application of Proof Theory and Type Theory.

1 Herbrand's prototype proofs.

"...when we say that a theorem is true for all x, we mean that for each x individually it is possible to iterate its proof, which may just be considered a prototype of each individual proof." Herbrand (1930), see [Goldfarb87], pp.288-9. Little Gauss' theorem above is an example of such a proof. But any proof of a universally quantified statement, over a structure that does not realize induction, is a "prototype" (e.g., for any Euclidean triangle, the sum of the internal angles is 180°: take a generic triangle, draw the parallel line to one side etc.). Similarly, if you want to prove a property for any element of a (non-trivial sub-)set of

[1] At that time, the very beginning of the century, Poincaré, in correspondence with Zermelo, hinted at the possible independence of the Continuum Hypothesis, CH, from Zermelo's formal axioms: he conjectured that there could be no sausage machine for Set Theory either. This was rigorously proved by Gödel and Cohen some 30 and 60 years later, respectively.

reals, of complex numbers .... But, in Number Theory, one has an extra and very strong proof-principle: induction. Clearly, in 1930, Herbrand thought that, in the special case of integer numbers, their universally quantified properties could always be proved by induction: completeness of PA was the current belief (with a few remarkable exceptions, such as H. Weyl, who – though hesitantly – conjectured the incompleteness of PA in Das Kontinuum, 1918 (!), and also stressed that the dream of potential mechanization was a form of trivialization of mathematics).

But what is the difference between prototype proofs and induction?

In a prototype proof, you must provide a reasoning which uniformly holds for all arguments, and this uniformity allows (and it is guarantied by) the use of a "generic" argument (see below). Induction provides an extra tool: the intended property doesn't need to hold for the same reasons for all arguments. Actually, it may hold for different reasons for each of them. One only has to give a proof for 0, and then provide a uniform argument to go from x to $x + 1$. That is, uniformity of the reasoning is required only in the inductive step: this is where a prototype proof steps in again, the argument from x to x +1. Yet, the situation may be more complicated: in case of nested induction also this inductive step, a universally quantified formula, may be given by induction on x. However, after a finite number of nesting, one has to get to a prototype proof going from x to $x + 1$ (induction is logically well-founded).

Thus, induction provides a rigorous proof principle, which, over well-orderings, holds in addition to uniform (prototype) proofs, modulo the fact that, sooner or later, a prototype proof steps in. Note though that the prototype/uniform argument in an inductive proof allows to derive, from the assumption of the thesis for x, its validity for $x + 1$, in any possible model. Moreover, as we shall see, by induction one may inherit properties from x to $x + 1$ (e.g., totality of a function of x, see below).

Yet, one more point should be mentioned. In an inductive proof, one must know in advance the formula (the statement) to be proved: little Gauss did not know it, for example. A non minor problem in automatic theorem proving .... Indeed, (straight) induction (i.e. induction with no problem in the choice of the inductive statement or load) is closer to proof-checking than to "mathematical theorem proving": proving a theorem, in mathematics, in general, is answering a question, not necessarily, not always, checking an already given formula.

Type Theory may help us to give a more rigorous description of what is a prototype proof. Propositions are types and proofs are terms, for us. Then a prototype proof, with a generic argument, is a term which may be uniformly instantiated by that argument: a schema for each individual proof. Let's see now a very informal definition, just a suggestion of what should be formalized:

Definition (Very Informal Type Theory, VITT). *Given a type A, a (closed) term P is* ***generic*** *and a (proof-)term $N : [P/x]A$ is a* ***prototype****, if there exists a term $M : A$ such that*

$$[P/x]M = N : [P/x]A.$$

This is surely VITT, as it is not mentioned: what are types and variables, exactly (1st order dependent types? Variable types? more?); what is equality? which kind of restriction is assumed in the substitution operation $[P/x]M$? A rigorous definition of prototype proofs as lambda-terms, in Girard's System F, and a few results (coherence, decidability), are given in [Longo00]. As for now, consider the informal definition above for what it is worth, with a first order understanding of variables and dependent types (if there is provably no way to make it rigorous, I will return the generous support for this invited lecture ...). The point is that many different answers are possible for each of the three questions above, and each would lead to different results.

Note that in a prototype proof, the generic input P must be typed, possibly with reference to a semantics. It is clearly so in little Gauss' proof in the introduction: as it is, n must be an integer (it must have the type of the - standard - integers). Yet, the proof may be extended to a non-standard model: the extension requires some technicalities related to the peculiar order-structure of the non-standard part. In short, one may generalize Gauss' proof by looking at a two steps proof: if n is standard, go on as Gauss, otherwise adjust it "ad hoc" to the pathologies of non-standard models of PA (try to write this down in full, as an exercise: the required symmetry may be reconstructed, with some work; note that, in any case, even a "unified proof", working on all models at once, should inevitably use the semantic information on the order-structure of models.) A proof by induction, instead, in passing from x to $x + 1$, uses a formal variable, which is typed just as a first order entity and may be interpreted in any model, with no reference to their structure (but, of course, one must know the formula in advance!). Thus, a prototype proof may use a strong information on types or models, along the proof: the input is a standard integer or it is in the non-standard part, in our example. As we know, this model-theoretic information is not formalizable in PA (by the Overspill Lemma: no predicate which holds for infinitely many values, may be valid only on the standard initial fragment of a model). This is why, in general, a prototype *proof* does not need to yield a formal proof. Below, we will point this out precisely, in an example.

Intermezzo: completing incompleteness (Part I and II)

Inter-Part I: On the incompletability of PA

Hilbert's concern, in proposing his famous wrong conjectures was twofold. First, in view of the semantic delirium of geometry, since the fall of the 2300 year old Euclidean empire, one had to retrieve certainty in logico-formal reasoning only, with no reference to meaning in space, time or whatever, as recalled above. Second, in his great mathematical rigor, he was aware of the general mess of the extraordinary mathematics of the XIX century: an extraordinary but turbulent growth, where proofs (of valid theorems) were often grounded on hand-waving (Cauchy's work provides good examples for this ...) and theorems were not

always true. Thus, Hilbert proposed to give a frame where one could decide "what is a proof". As a matter of fact, in hilbertian systems one can give a decidable notion of proof: this is a key aspect of Hilbert's notion of formal systems. Then, as recalled above, he further conjectured that any proposition had to be decided, beginning with the describable propositions of the language of PA, of course.

As we all know too well, Gödel, by the I Incompleteness Theorem, proved that

If PA is (ω-)consistent, then it is incompletable

That is, no consistent formal extension of PA is complete, or that no consistent extension, with a decidable notion of proof, allows to decide all aritmetic propositions. Moreover, given such an extension, the formalized statement asserting its own consistency is one of the undecidable propositions (II Incompleteness Theorem).

It is easy though to give an example of **non** formal system, extending PA. Consider Arithmetic with

$$(\omega - rule)\ \ \frac{A[n]}{\forall x A[x]}\ for\ all\ n \in \mathbb{N}$$

This system is complete, but it has a non decidable notion of proofs (yet, proofs are well-founded trees).

Note that, the ω-rule derives $\forall x A[x]$ from the assumption of the infinitely many instances of $A[n]$. This is different from any use of prototype proofs, where $\forall x A[x]$, over N, is obtained from a schematic (prototype) proof of $A[n]$, w.r.t. a generic (replaceable) $n \in N$.

Inter-Part II: Formal Proofs of Consistency

The firm formalists often insistingly remark that, after all, any unprovable statement can be proved in a "suitable" formal frame. Now, Gödel's theorem implies that the order of quantifiers cannot be reversed: for any formalized statement, there exists an extension of PA which proves it ... (and not conversely.) As stated, this is trivially true, since, given a formal statement in PA, there exists for sure an extension of PA which proves it: add that very statement as a new axiom .... Yet, even in the case of the (trivial) extension (by the very statement, but by more interesting formal principles as well), there is a non minor problem: one has to prove the consistency of the intended extension! Often, in philosophical discussions, this fact is forgotten.

Consider, say, the formalized statement of the consistency of PA, $Cons_{PA} \equiv \neg Theor_{PA}(0 = 1)$ (i.e. "$0 = 1$" is not a theorem), as a typical unprovable proposition, given by Gödel's second theorem (yet the following argument applies a fortiori to the many formally unprovable propositions, that imply Cons_{PA}, such as those analysed below). One can surely formally derive $\neg Theor_{PA}(0 = 1)$ from a suitable, and consistent, formal frame: ZF, for example (Zermelo-Fraenkel formalized Theory of Sets). Even the axiom of infinity in ZF can be formally

stated in a finitistic fashion; so, a Turing Machine, or our firm formalist, can mechanically derive $\neg Theor_{PA}(0 = 1)$ from the encoded version, in PA, of the axioms of ZF. Call the conjunction of these (encoded) axioms Ax_{ZF} and observe that:

$$PA \vdash (Ax_{ZF} \rightarrow \neg Theor_{PA}(0 = 1)) \tag{1}$$

since, by the various equivalence theorems and by Gödel's representation lemma, any Turing computable function can be fully represented in PA.

Can one then say that the consistency of PA has been formally proved, by finitistic-formal tools, as many claim? Well, PA is consistent if it generates no contradictions, that is if $\ldots Cons_{PA} \equiv \neg Theor_{PA}(0 = 1)$ **holds**. Does (1) prove this fact?

No, it only proves what is written, i.e. the formal implication $(Ax_{ZF} \rightarrow \neg Theor_{PA}(0 = 1))$. This statement implies the consistency of PA, as validity of $\neg Theor_{PA}(0 = 1)$, provided that ZF is ... consistent (otherwise, from Ax_{ZF}, one could derive everything, including false statements). Now, the consistency of ZF can be shown in either of the following ways:
A - Ax_{ZF} formally generates no contradiction
B - ZF, i.e. Ax_{ZF}, has a model.

A and B are well known to be equivalent, but the formalist who rejects to give meaning to formulae, in particular to the axiom of infinity, may insist about proving A formally. Easy, give a formal Set Theory with a stronger formal axiom of infinity .... and so on so forth towards a never ending regressing chain (in this approach the model construction is hidden or just implicit).

Consider then B. If one explicitly constructs or assumes to have a model of ZF, including of the axiom of infinity, then (1) does prove the consistency of PA. This is so, because this construction/assumption implies that Ax_{ZF} generates no false theorems, and because, if Ax_{ZF} has a model, then also $Cons_{PA} \equiv \neg Theor_{PA}(0 = 1)$ holds (and PA is consistent), by (1).

In summary, a purely formal derivation of $Cons_{PA}$, from whatever formal axiom system, *does not show* the consistency of PA, unless one involves the **meaning** of the required axiom of infinity, for example by giving a model of ZF (by the way, ZF axiom of infinity essentially says: "PA has a model").

I am here saying a triviality that everybody should be aware of. Yet, too often, in (automated) theorem proving in particular, many people claim that they can prove formally the consistency of PA, or whatever formally unprovable property. Yes, of course, one can derive the implication in (1) or, more generally, given any formalized statement, one can propose some strong enough formal axioms, which mechanically imply it. But, this implication proves the statement, in a mathematical sense, if one assumes the consistency of ZF or of the intended stronger theory. In other words, just "writing the axioms" and computing is not sufficient: one also has to assume/prove that the derivation is sound, i.e. that axioms and rules are consistent (meaningful). This is one of the general reasons for the need of interaction man/machine in theorem proving and it shows up when one has to bootstrap the machine with a suitable formal frame (which

may depend on the result one aims at), but also along the proof, as it will be argued below.

Finally, recall that Gödel's first incompleteness theorem is "just" an undecidability theorem and it says nothing about the truth of the undecidable sentence, call it G. Yet, the second theorem *derives* G *from* (formalized) consistency, $Cons_{PA}(\equiv \neg Theor_{PA}(0 = 1))$, and this within PA. Observe then that those who refer to the "truth" of the undecidable sentence G as to an ontological miracle, actually, if asked, *derive* it, by handwaving, from its unprovability. And, thus, they just mimic the (short, but subtle) formal proof of G in the second theorem, as derived from $Cons_{PA}$. As a matter of fact, one has to *assume* consistency in order to prove the undecidability (thus the unprovability) of G (first theorem), yet formalized consistency *implies* G *within* PA (second theorem). Most commentators miss this fantastic "calembour" in the interplay between the two theorems and the double role of consistency.

There is no miracle here, as too many claim, since there is no other access to the "truth of a proposition", in Mathematics, but by proof, including of G. (In short, ask the platonist: why is the undecidable sentence true? he/she would prove it by handwaving, that is, if he/she is good enough and makes the assumptions explicit, roughly as follows:
"The consistency of PA implies the unprovability of G. But PA proves the equivalence of G and $\neg Theor_{PA}(code(G))$ (usual handwaving: ...and, hence, G "means" or says that G is unprovable - use exclamation mark here!). Thus (further handwaving), by the truth of $Cons_{PA}$, the consistency of PA, one has the truth of G.")

This, or any similar argument, is just a bad imitation of one implication in the proof of the second incompleteness theorem. Yet, a "proof.

The concrete incompleteness results below do not allow to cheat about their "truth" nor to speculate on insights over God's shoulders or Quantum Mechanics in the brain, since they require a (non-obvious) proof of the truth, in the standard model, of the unprovable sentence in PA. The point is to see where exactly, along the proof, the formal unprovability shows up. This is our aim below.

2 Concrete incompleteness I: Normalisation.

In 1958, Gödel gave a proof of the first combinatorial statement, unprovable in PA: normalisation for a typed extension of lambda-calculus, system T. Lambda-calculus and its effective extensions may be (easily) encoded in PA and, thus, the encoded Π^0_2 statement, "for each term, there is a normal form" (in system T), may be shown to hold, even though it is not formally provable, since it implies the consistency of PA (proving consistency of PA was Gödel's aim, in 1958). Gödel's proof was extended and improved by J.-Y. Girard, by the theorem we discuss here.

Some claim that this theorem cannot be called an independent "concrete" or combinatorial statement of PA, as it is "too much" related to consistency. However, lambda-calculus has also a mathematical-combinatorial interest per

se, not just for proving consistency of PA. Thus, we do not see the reasons for depriving the Thoery of Types of the first achievement in this sense: a provably true, but formally unprovable mathematical statement of PA. Paris-Harrington theorem, a remarkable result of 1977, is usually given this honour; however, Ramsey theorem, which underlies this finitary variation, is not less related to consistency, via Set Theory.

But how normalisation can ever be proved, if it is unreachable within PA? Easy (oh, no, very difficult), by a prototype proof, besides induction.

2.1 The unprovability

I will briefly analyse now the "internal reasons" for the formal (PA) unprovability of normalisation. By this I mean an informal insight into the parts of the proof where non-encodable arguments are used. Clearly, the unprovability is rigorously shown by the formal implication: from normalisation derive consistency. Yet, one may try to spell out explicitly the places where the incompleteness phenomena shows up, along the proof of normalisation.

This exercise is analogue, by duality, and may serve as a guideline, to the everyday task in interactive theorem proving (proof-assistants, proof-checking). As a matter of fact, in order to feed a computer with parts of a proof (e.g. a very difficult combinatorial lemma, lots of very long computations ...) one must be able to isolate the fully formalizable parts in the intended logical frame, and have the computer develop or check them. That is, one must be able to point out or distinguish the non-computable from the computable, the essentially higher order from the first order, the use of axioms of infinity and their models from arguments in PA and so on so forth.

Of course, every theorem may allow many different proofs. In our case, this means that the non-encodable passages may be different, as their very nature may depend on the kind of assumptions and proof adopted. I will then focus on Girard's argument by "candidates of reducibility". This approach to normalisation applies both to Gödel's system T and to Girard's System F (see [Girard90]). Indeed, I will mostly refer to the presentation in [Girard90] for the discussion.

Girard's proof uses a very heavy induction load. That is, in order to prove that every term has a normal form, by induction, it adds to an inductive assumption on the type of a term (see below), two extra assumptions. Why is this done (and needed)?

The point is that in no way induction can be straigthfowardly applied to terms (e.g. by an induction on the complexity or lenght of terms or whatever). This is due to several features of typed calculi. First, the arrow (in the implicative types) is "contravariant": in any formula such as "$\forall x(\rho \to \sigma)$", the properties expressed by ρ are "negated", and this increases the complexity of the type as formula and of the terms living in it. Second, in second order types (i.e. in the types of System F, where universal quantification is over type variables), the type variables may be instantiated by any type, including the one under consideration; this strong impredicative feature forbids any (inductive) stratification of types and terms living in them. In particular, terms may contain types and depend

on them: this is one further reason which does not allow induction on "pure" terms. Yet, and surprisingly enough, the dependence of terms on types is very uniform and this is crucial to the proof (see [Girard90]; in a sense, the specific value of a variable type in a term may be disregarded - or all its values affect the computation in the same way: this fact requires some technicalities and it is fully spelled out by the Genericity Theorem in [Longo93]).

Then, the induction on types (not on terms) goes on by using a set of terms in the given types (the "candidates of reducibility"). The terms in such sets are supposed to be normalisable (first clause of the induction), but also to satisfy two further properties, not obviously related to normalisability (see [Girard90], p. 43, 116, 117); a "fine tuning" of these properties in extensions is a common and relevant practice, both in Logic and in Computing (see [Coquand88] for a classic; a non-obvious extension to "subtyping" may be found in [Castagna95]). The further key observation is that the properties in the induction load are not written at the theoretical level (System F or whatever second order system, see 2.2), but belong to the metalanguage. Then, along the proof, one collects the metalinguistic sets of candidates of reducibility into a type and performs some computations, within system F (in a sense, one "brings down" the metatheoretic notion to a theoretic one).

Finally, the very handling of second order collections may be understood, in set-theoretic terms, as the use of a proper Second Order Impredicative Comprehension Axiom. This axiom requires an essential blend of syntax and semantics (this is why the competent formalists reject it), in view of the semantic convention on variables (capital variables, say, must range on (sub-)sets of any model, small ones on elements of these sets). Thus, the proof, as given, uses a blend of meta-theory, theory and semantics and, by this, it lies outside PA or of any coding of System F into PA (and much beyond it, see below).

Remark: Unprovability and Hilbert's "organisation of the mathematical discourse". The blend of meta-theory, theory and semantics is very common in mathematics and in every day language: the dream of an unique, definitive formal universe, where all of mathematics could be formalised, relies exactly on this three level distinction and on the conjecture that a well isolated theoretical level could completely describe mathematics. Now, Hilbert(-Tarski) organisation of the mathematical discourse into meta-theory, theory and semantics has been one of the remarkable ideas of the century, in Logic, but it does not describe an absolute objectivity. In [Longo01], it is compared to Euclid's organisation of space, by rigid figures and their homotheties, as for relevance. Clearly, the later is an extraordinary approach to physical space, a non-arbitrary, well motivated proposal, but a rather artificial one, as there is no such a thing as a rigid body and physical space is not closed under homotheties (according to Relativity Theory and current microphysics, since only the group of automorphisms of Euclidean geometry contains the homotheties).

Similarly, Hilbert's approach is so artificial that it has been instrumental to the invention of fantastic artefacts, Turing Machines and, then, our digital

computers. These machines work just at one level, the formal-theoretical one. The incompleteness theorems proved for us that this organisation of the proof is not an absolute: first, Gödel Representation Lemma showed that the metatheory of PA could be fully encoded into PA itself and, thus, that it is "part of it"; later, the above proof of normalisation essentially used a blend of the various levels, as just pointed out, and blurred this artificial difference. Again, it was extremely useful to invent such a way to analyse the proof (in a sense, I have been using it above), but it is just a technical tool and a temporary one: it lives nowhere, as, for us, humans, there is no such a thing as a "metalanguage" ("I may play chess according to certain rules. But I may also invent a game where I play with the rules themselves. The pieces of the game are then the rules of chess and the rules of the game are, say, the rules of logic. In this case, I have *yet another game*, not a *metagame*" [Wittgenstein68] p.319).

However, there may be a difference in method (even within the same "cage of language", as Wittgenstein would put it): proofs should be analysed *also* by non-mathematical *arguments*, e.g. by non-mathematized insights into the general structure and dynamics of thought, not only by logico-mathematical proofs. Mathematical Logic, yet the main tool for the foundational analysis, is still part of mathematics, by its method and its proofs, thus it cannot *completely* found it. It is somewhat surprising that many leading colleagues, who carefully avoid some consistent and expressive mathematical circularities (or vicious/virtuous circles, impredicativity or non well-foundeness, say), accept this conceptual and philosophical, severe circularity and only develop a metamathematical analysis of foundation. Now, no *proof* can found the notion of *proof*. Or, as suggested by Wittgenstein ... "Hilbert's metamathematics will turn out to be a disguised Mathematics" (quoted in [Waismann31], see also [Heinzmann90], [Floyd98] for more on Wittgenstein and incompleteness).

2.2 Berardi-Altenkirch normalisation of System F, in LEGO

As already mentioned, given any formally unprovable statement, one can surely invent a formal frame to prove it. As a matter of fact, this frame may even reconstruct the very path of the given proof. The unprovability will be then described by the "proof-theoretic strength" of the formalisation and the validity of the statement will rely on the consistency of the proposed formal theory.

This kind of analysis is one of the major contributions given by Mathematical Logic to the foundation of Mathematics. One of its main applications is the invention of systems to handle automatically as much mathematics as possible. In our case, the analysis amounts largely in displaying the exact "formal order" of the different constructions, understood above in terms of language, metalanguage, meta-metalanguage etc. It differs by this from the set-theoretic analysis, mentioned in 3.1 below.

S. Berardi started an analysis of Girard proof by higher order Arithmetic ([Berardi91]). T. Altenkirch completed it and fully encoded it into LEGO, a very interesting and effective type-theoretic proof-assistant (see [Altenkirch92]).

Let us first clarify what is meant by "higher order logical system". In the case of Arithmetic, sometimes people refer, say, to Second Order Arithmetic, as THE categorical theory of numbers. That is, to the non-formal theory where the interplay between induction and a full second order comprehension axiom for sets, allows to say formally: "any non-empty subset of N has a least element". This is not a formal system, in the sense we attributed to Hilbert, since the notion of proof is not recursive: as a matter of fact, the system is categorical (and thus complete) and, by this, the set of "theorems" (which coincides with the valid propositions over N) cannot be recursively enumerable.

One may instead define a formal system handling higher order variables (PA_2 formalises set variables, PA_3 variables of sets of sets), and leave induction restricted at each level (number theoretic induction, induction over formulae containing set variables ... all treated differently, see below). In these systems, the notion of proof is decidable, or deductions are effective (even fully mechanisable by LEGO!), and, of course, Gödel's theorems apply to them: thus, they are incomplete. Now, PA_3 allows a universal quantification over (the sort of) *sets of sets of integers*, i.e. a universal quantification over $P(P(N))$ (the powerset of the powerset of N). By this, as we shall see, it provides the right (and minimal) formal frame for the specific theorem we are currently interested in, normalisation for System F (and a fortiori for Gödel's T).

Note finally that many prefer to call PA_2, PA_3 ..., first order *multisorted* systems (and "apparent" higher order ones, as extended first order systems with just different labels for variables). This is a matter of taste, as "proper" higher order systems are non-formal and yield non-effective deductions: the point is to be clear as for what one is talking about and to have clearly in mind that the core idea of Hilbert's formal systems relies on the decidability of the notion of proof. This is so in PA_2, PA_3 ..., while it is lost in "proper" (categorical or complete) higher order systems.

As already mentioned, Girard's proof uses an inductive definitions over "candidates of reducibility", as *subsets of types* (in the sense that types are "sets of terms" and they have the kind of the sets of integers). This induction may be described as a recursive definition of a function over sets, or a "third order induction". Now, first order induction works on integers, second order induction takes the *minimum fixed point* of monotone operators over sets of integers. The non-obvious concept of third order induction amounts to say that one has to take the minimum fixed point of monotone operators over sets: clearly, such an operator, a function, is a (single-valued) set of (pairs of) sets, similarly as a function over numbers is a (single-valued) set of (pairs of) integers. It is then a "third order" operator and it requires PA_3 to be formalized. Note that no less can be used, as normalisation for System F implies consistency of PA_2, not just PA ($= PA_1$).

Let's see more closely what happens along the proof, by hinting to the insightful formalization by Berardi. As mentioned in 2.1, the key point is that the proof does not go by induction on terms (as first order entities), but by higher

order induction. That is, the proof uses a combined induction on types (second order induction) and on operators on types (third order induction). In particular, the recursive definition of candidates of reducibility uses a map on types, $[.]_\rho$, from a candidate assignments ρ to candidate $[\sigma]_\rho$, for each type σ. This map is a (single-valued) set of (pairs of) sets. Or, more precisely, one considers the graph of the operator:

$$(\sigma, \rho)| \rightarrow [\sigma]_\rho$$

and this has the type of *sets of sets of integers*. The key step in the formalization of Girard's proof is based on taking the minimum fixed point for this operator. Its construction (existence) implies the convergence of the normalization algorithm on System F.

Call now $\forall x \exists y Norm(x, y)$ the (first order) formal statement of PA_1 that "for any (coded) term of F, there exists a (code for) its normal form" (more precisely, in $Norm(x, y)$, y is the code for the redution sequence ending with a normal form.) Then $Norm(x, y)$ is a decidable predicate in x and y. Thus we have just pointed out that

$$\mathrm{PA}_3 \vdash \forall \mathrm{x} \exists \mathrm{y} \mathrm{Norm}(\mathrm{x}, \mathrm{y}) \tag{2}$$

Thus, *under the assumption that* PA_3 is consistent (more precisely: that it is 1-consistent), the proposition holds in the standard model (where number-codes of terms refer to actual terms), or

$$\mathrm{N} \models \forall \mathrm{x} \exists \mathrm{y} \mathrm{Norm}(\mathrm{x}, \mathrm{y}). \tag{3}$$

Now, fix $n \in N$ and consider $\exists y Norm(n, y)$. This is a Σ_1^0 predicate and any *valid* Σ_1^0 predicate over N is provable in PA_1 (easy: scan the integers till you find one satisfying the predicate). Then

$$\textit{given } n \in N, \mathrm{generic}, \mathrm{PA}_1 \vdash \exists \mathrm{y} \mathrm{Norm}(\mathrm{n}, \mathrm{y}) \tag{4}$$

As already observed, no induction on n (as code of the term to be normalised) could be used in the proof: induction is entirely transferred at the level of types and functions on types. Of course, the reader must appreciate the difference between (2) and (4), a subtle but crucial difference: the PA_1 unprovability of $\forall x \exists y Norm(x, y)$ says that there is no way to prove it by first order induction on x. Thus, the proof in (4) is essentially a prototype proof, with generic input n, a standard integer or a true code for a term. And a detour must be taken, via PA_3 and its higher order forms of induction, over sets of types.

Observe also that (4) proves only that there is a computable function that, taken an integer n as code of a term, gives (the code) of its normal form, y. Yet, as stated, (4) does not prove that this "normalising function" is total (and this is where lies part of the logical complexity of the problem). An inspection of the proof of $\forall x \exists y Norm(x, y)$ within PA_3, a proof needed in order to assert its truth, shows that this function is indeed total[2]. Finally, PA_2 may suffice to "isolate"

[2] As already mentioned, induction, in contrast to prototype proofs, proves totality of Π_2^0 predicates "for free". Consider the following theorem: (D) $\forall x \exists y (2x < y)$. Of

the standard integers (PA_2 contains a predicate for them, exactly, i.e. it defines the type of integers), but this is not sufficient to formalise (4), as the proof that the normalising function is total entirely relies on the third order structure of the normalisation proof.

3 Concrete incompleteness II: Kruskal-Friedman theorem.

Everybody knows what is a tree. Trees in Mathematics grow downwards. They are partial orders with a root on top (the largest element), and, for each node a in a tree T, $\{x/x > a\}$ is totally ordered. A tree-embedding h, form T to T' (notation: $T < T'$), is an injective map, which preserves upper bounds (i.e. h(sup{a,b}) = sup'{h(a),h(b)}).

Kruskal, in 1960, proved the following theorem (KT): *For any infinite sequence of finite trees* $\{T_n/n < \omega\}$*, there exist* j *and* k *such that* $j < k < \omega$ *and* $T_j < T_k$.

A non-obvious result. In 1981, Friedman proposed a finitary version of this fact, i.e. a variant that may be stated in PA. Here is a form of it:
For any n, there exists an m such that for any finite sequence $T_1, T_2, ..., T_m$*, where* T_i *has at most n(i+1) nodes, there exist j and k such that* $j < k < m$ *and* $T_j < T_k$.
(FFF or Friedman's Finite Form, see [Harrington85], [Gallier91]).

This Π^0_2 statement of PA has a purely combinatorial nature, since in no apparent way it is related to consistency issues. Moreover, both KT and FFF have several interesting consequences in finitary combinatorics (e.g. in Term Rewriting Systems).

Let's try to sketch the "reasons for unprovability", as we did in the previous case. However, now, two radically different proofs of KT (and FFF) are available, each worth analysing, although briefly. Once more, the logical reasons for unprovability are grounded on Gödel's second theorem, since, surprisingly enough (and this is the remarkable insight of Friedman), FFF implies the consistency of PA (and much more). We will hint to the specific passages of the proofs where unprovability shows up. FFF easily follows from KT, by an application of Koenig's lemma ("any infinite finitely branching tree has a infinite path"). This lemma is conservative over PA, thus the problems are hidden along the proof of KT.

---

course, this has both a prototype proof ("For a generic n, take m = 2n+1") and an inductive one, in PA ("for 0 take 1; assume that, given x, you have y, then, for (x+1), set y' = y + 2"). Difference: the second proof inductively proves also that the map from x to y in (D) is total. The first one, instead, requires a further insight: one has to prove that "." and "+" are total (not too hard, in this case; very complex, as for the totality of the normalising function, since this is shown in PA_3).

3.1 The set-theoretic analysis

The usual, set-theoretic proof of KT goes by a strong non-effective argument. It is non-effective for several reasons.

First, one argues "ad absurdum", i.e. one shows that a certain set of possibly infinite sequences of trees is empty, by deriving an absurd if it were not so. More precisely, one assumes that the set of "bad sequences" (or sequences without ordered pairs of trees, as required in the statement of KT) is not empty and defines a minimal bad sequence from this assumption; then one shows that that minimal sequence cannot exist, as a "smaller" one can be easily defined from it.

Note that this minimal sequence is obtained by using a quantification on a set that is ... going to be proved to be empty, a rather non-effective procedure. Moreover, the empty-to-be set is defined by a Σ^1_1 predicate, well outside PA (a proper, impredicative second order quantification over sets, see the discussion on system F, 2.1).

For the non-intuitionist who accepts a definition of a mathematical object (a sequence in this case) ad absurdum, as well as an impredicatively defined set, the proof poses no problem. It is abstract, but very convincing (and relatively easy). The key non-arithmetizable steps are in the Σ^1_1-definition of a set and in the definition of a sequence by taking, iteratively, the least element of this set. Yet, the readers (and the graduate students to whom I lecture) have no problem in applying our shared mental experience of the "number line" to accept this *formally* non-constructive proof: from the assumption that the intended set is non-empty, one understands ("sees") that it has a least element, without caring of its formal (Σ^1_1-)definition. Or, if the set is assumed to have an element, then the way the rest of the set "goes to infinity" doesn't really matter, in order to understand that it must have a least element: the element supposed to exist (by the non-emptiness of the set) must be somewhere, in the finite, and the least one will be among the finitely many which precede it, even if there is no way to present it explicitly. Finally, the sequence defined ad absurdum, in this highly non-constructive way, will never be used: it will be absurd for it to exist. So its actual "construction" is irrelevant.

Of course, this is far away from PA, but it is convincing to anyone accepting the "geometric judgement" mentioned in the introduction: a non-empty subset of the number line has a least element (see the conclusion as well).

3.2 The constructive version

An intuitionistically acceptable proof of KT has been recently given in [Rathjen93]. This proof of KT is still not arithmetizable, of course, but it is "constructive", at least in the broad sense of infinitary inductive definitions, as widely used in the intuitionist community (see the seminal work by Martin-Löf; a classical introduction is in [Aczel78]).

The idea is to construct an effective "reverse embedding" of the partial order of finite trees (partially ordered by the tree-inclusion above) into a suitable ordinal representation system. This requires an insightful study of the combinatorial properties of tree-embeddings.

In short, let $FinBad_T$ be the set of *finite* bad sequences of trees (see 3.1). Once given a system of ordinals (ORS,<), Rathjen and Weiermann construct a function $f : FinBad_T \to ORS$ such that, if s and t are in $FinBad_T$, and t extends s strictly, then $f(t) < f(s)$ (in ORS). Clearly then, if there exists an infinite bad sequence, and thus an infinite ascending sequence in $FinBad_T$, then ORS would contain an infinite descending sequence.

The function f is actually primitive recursive and everything up till this point can be done in Intuitionist Arithmetic, HA. KT then follows from the proof that (ORS,<) is well-founded and, in particular, that every primitive recursive sequence $p(0) > p(1) > p(2) > ...$ must terminate after finitely many steps. This proof is done in an intuitionistically acceptable formal frame, called ID_1, which extends induction along constructible ordinals and well beyond PA.

This is the non-formalisable part, in PA. As a matter of fact, PA cannot prove the Π^0_2 statement that (ORS,<) is (primitive recursively) well-founded, since this well-ordering is sufficient to prove the consistency of PA (it actually implies induction well beyond ϵ_0, the ordinal of consistency for PA).

As in the previous case, the Π^0_2 statement of FFF cannot be proved by first order or ω-induction, i.e. within PA. In the approach by inductive definitions though, the difficulty is taken care by "pulling induction along the ordinals", well beyond ω, or even ϵ_0.

Conclusion

The proof-theoretic investigation of Mathematics has been one of the major achievements of XX century Science: it gave us mathematical rigor and modern computing. The latter, its major fall-out, is changing our live. Yet, we need to go beyond its techniques and philosophy. First because, in view also of its successful paradigms, Mathematics has now reached a remarkable level of rigor and we are no longer scared of the novel geometric intelligibility of physical space, which originated Frege's "royal way out" from the "delirium" in Geometry (see [Tappenden95] for more on Frege's view). Second, because we may take advantage, also in computing, by an enlargement of our foundational paradigms, beyond the traditional linguistic-finitistic certitudes. In a sense, we should try to bring together the two "foundational ways" that split at the end of XIX century, as I tried to summarize in the introduction. In short, the foundation of Mathematics lies in:

- Logic
- Formalisms
- Regularities of phenomenal space.

We all know what the first two points mean and their relevance. By the third, I mean the reference to a few regularities of phenomenal space, such as connectivity (Riemann) or symmetry (Weyl), but "ordering" as well. By the subsequent and constructed notion of well-ordering, it was meant here the (very strong) geometric judgement: "consider a *generic* non-empty subset of the integer number line, observe that it has a least element" (see [Dehaene98] for a

neuropsychologist's experimental analysis of our mental number line, as a cognitive experience).

As a matter of fact, in Mathematics, we transform these regularities of space - that we happen to "see" - (as well as our cognitive approach to them by mental re-constructions), into explicit conceptual invariants (as "hypothesis", in Riemann's terminology). This process grounds mathematics, as a conceptual construction, in our "phenomenal lives" (as Weyl would put it): concepts and structures are the *result* of a cognitive/knowledge process. Then they are further extended, by language and logic: from connectivity we go to homotopy theory or to the topological analysis of dimensions, say. Symmetries lead us to duality and adjunctions in Categories. The ordering of numbers is extended into transfinite ordinals.

Of course, these notions may be formalized, each in some "ad hoc" way, as there is no Newtonian or ZF absolute universe. But evidence and foundation are not completely captured by the formalizations, since "The primary evidence should not be interchanged with the evidence of the 'axioms'; as the axioms are mostly the result already of an original formation of meaning and they already have this formation itself always behind them", [Husserl33]. Moreover, incompleteness tells us that the reference to this underlying and constitutive (not independent) meaning cannot be avoided in foundation, as the consistency issue is crucial in all formal derivations (see the Intermezzo, part II).

In this perspective, we need to ground mathematics also on a few "geometric judgements" which are not less solid than the logic ones: "symmetry" for example is at least as fundamental as "modus ponens", or it steps heavily into mathematical constructions (and in proofs, as pointed out by Girard - see [Girard01] for recent advances of his program in Logic and foundations). As already mentioned, physicists argue since long "by symmetry". More generally, modern Physics extended its analysis from the newtonian "causal laws" (the analogue to the logico-formal and absolute "laws of thought", since Boole and Frege) to an understanding of phenomenal world by our active geometric structuring of it: from the conservation laws as symmetries (Noether's theorem), to the geodetics of Relativity Theory (see [Weyl27] for an early mathematical and philosophical insight into this, [BaillyLongo06] for recent reflections). The normative nature of geometric structures is currently providing a further understanding even of recent advances in microphysics ([Connes94]). Our foundational analyses and their applications should also be enriched by this broadening of paradigm in scientific explanation: from laws to geometric intelligibility (grounded on accessibility of space, see [Longo02]). But in Logic as well, we have to move from viewing formal properties and logical laws as a linguistic description of an independent reality, to their appreciation as a result of a praxis: they are the constituted invariants of our practice of reasoning and language, as an open ended "game" between us and a world to be organised also by language.

In Number Theory, well-ordering, as a structuring of phenomenal space and time by integer numbers, is a founding "geometric judgement": its certainty is the consequence of a common conceptual construction, the number line, we all

experience in our (mathematized) mental space (see above and the Introduction). When "formalized", as proper second order impredicative statement (induction plus second order full comprehension), it is highly ineffective. Instead, it is perfectly "robust" (and "effective" - as a mental construction) if seen as a "geometric judgement", related to the constructed order structure of integer numbers. Of course, considering - "seing" - a *generic* non-empty subset is crucial, instead of formally taking *all* non-empty subsets (see also 3.1 and below). And, as a judgement, it provides a reliable geometric argument for consistency of PA: all other proofs use consistency of stronger theories, large ordinals or axioms of infinity. Clearly, it is not an alternative to the fine analysis of consistency provided by Proof-Theory (normalisation, relative consistency results ...) but is complements it, by grounding Mathematics on broader mental experiences.

It should be clear, though, that the mathematical constructions, such as the "number line", are "shared" cognitive performances as they are done in language and intersubjectivity, along history. They are "progressive conceptualizations", as suggested by Enriques ([Enriques01], [Faracovi82]), which originate from regularities of space, time and reasoning, and in no way the grounding of mathematics also on geometric judgements should let us forget the key role of language for our communicating human community. Only by language we can conceive and express never ending, discrete iteration, which we later place back into mental space (the number line). By action *and* language we organize (we "order") space and time. In general, intersubjective exchange, by language, is a core component of the *constructed* invariance and conceptual stability of Mathematics: that is, invariance is also the result of a conscious appreciation of "what we all share" (Poincaré), including a constructed mental image. Thus, in spite of this approach's debt to Brouwer's constructivism (but this debt has been filtered by the remarkable teaching by Dana Scott, J.-Y. Girard and Per Martin-Löf in Logic), we radically depart here from Brouwer's "languageless" mathematics, [vanDalen91]. As well as from the Platonism/formalism debate, the "new scholastic" of the XX century, as Enriques called it ([Enriques36], but Poincaré and Hermann Weyl should be quoted as well).

The reference to these leading geometers' critical (anti-formalist and anti-ontological) attitude w.r.to the main-stream foundational debate may be the occasion for a concluding remark concerning the use of "generic" elements in proofs (and in judgements). This is of course an essential notion in defining prototype proofs (and geometric judgements).

In Mathematical Logic, since Frege and Hilbert, by the prevailing algebraic approach and by the focus on Arithmetic, generic elements are dealt with as variables, formally handled by "for all introduction/elimination rules" and induction. Typically, in algebra, one proves $(x + y)^2 - 2xy = x^2 + y^2$, say, by formally manipulating two variables; then, the full generality of the result is obtained by the "for all introduction" rule. In this frame, "generic" means "variable ranging on all elements of the intended domain": in the proof of a universally quantified statement, just type and formally manipulate your variables and you are done. A crucial point, of course, is that "for all" is interpreted by ... "for all",

along the proof. This proof-theoretic treatement of "for all" (as well as the naive set-theoretic interpretation[3]) is confirmed by the use of arithmetic induction: in order to prove a property for all numbers, prove it for 0, than extend *the proof* to *all* numbers by moving from x to $x + 1$.

However, when one has to prove a property of structures or objects, that do not form a well-ordered set, *all* right triangles, say, or *all* Riemann's manifolds, or *all* algebraically closed fields, or even *all* real numbers ... in no way the proof is done "for all". One considers a *generic* right triangle, or Riemann's manifold, or non-empty subset of bad-sequences (see 3.1) ..., gives the proof and, at the end, one observes, by scanning again the argument: note that I only used the very definition of the intended structure (or mathematical object), no more no less, and this shows that my proof has the right level of generality. The drawing on the sand of a greek beach of a right triangle, the geometric proof by Pythagoras, the observation that the proof *depends only* on the right angle and *not* on the length or ratio of the sides, is the real birth of Mathematics. Note that, in order to give the geometric proof, the sides (their ratio) have to be given, as a specific right triangle (ratio of sides) *has* to been drawn. In no way *all* right triangles, with their different (ratio of) sides, are scanned, as, say, in inductive proofs; on the contrary, a (provably) generic one is used. The methodological difference is crucial and the key role given to Formal Arithmetic in foundation and in Proof Theory has been hiding it. Moreover, this contributed to the new scholastic in the philosophy of Mathematics: on one side, it confirmed the formalist lack of appreciation of the *construed* genericity of individual Mathematical objects and structures (formal variables are the "generic" signs); on the other it contributed to ontological commitments (the reference to *existing* sets of *all* right triangles, Riemann manifolds, real numbers ... instead of the analysis of their conceptual construction).

There is one more reason for going in this further direction, which stresses also the role of "geometric judgements" and generic structures (and should accompany the proof-theoretic analysis, when technically insightful). Mathematics is not only grounded on proofs, the main concern of late XIX and XX century Mathematical Logic, but it also goes by *construction of concepts and structures.* Indeed, new concepts and structures are required even along proofs, as shown by the permanent need of "new axioms" (or, also, when just trying to find the right induction load). The analysis of these constructions should not be left to the magic or the metaphysics of some ontological "intuition", but it should become part of a scientific investigation. This analysis I call "the cognitive component" of the foundation of mathematics and it is an ongoing project (see the research program "Géométrie et Cognition", this author's web page, or [Longo02], [Longo05], [BaillyLongo06]).

[3] cf. the much more structured and informative intepretation of first and second order variables and quantification in categories, [Lawvere76], [Lambek86], [AspertiLongo91].

Acknowledgements

I am greatly indebted to Stefano Berardi and Micheal Rathjen for several very stimulating and helpful discussions and e-mail messages. An anonimous referee and a disclosed one, Gilles Dowek, contributed to the revised version by their relevant and numerous comments. Of course, any mistake and the strong philosophical commitment remain of my own responsibility. This work has been partially supported by the "Action Cognitique" (MENRST), as part of the "Atelier Géométrie et Cognition". Longo's papers are downloadable:
http://www.di.ens.fr/users/longo/ .

References

[Aczel78] Aczel P., "An introduction to inductive definitions", Handbook of Mathematical Logic, Barwise ed., 1978.

[Altenkirch92] Altenkirch T. "A formalization of the Strong Normalization proof for System F in Lego", December '92.

[AspertiLongo91] Asperti A., Longo G. Categories, Types and Structures, M.I.T. Press, 1991 (out of print, downloadable: http://www.di.ens.fr/users/longo/)

[BaillyLongo06] Bailly F., Longo G., Mathématiques et sciences de la nature. La singularité physique du vivant, Hermann, Paris, 2006 (english introduction, downloadable: http://www.di.ens.fr/users/longo/).

[Berardi91] Berardi S. "Girard's normalization in Lego", Univ. Torino, 1991.

[Boi95] Boi L. Le problème mathématique de l'espace, Springer, 1995.

[Bottazzini95] Bottazzini U., Tazzioli R., "Naturphilosophie and its role in Riemann's mathematics", Revue d'Histoire des Mathématiques n. 1, 3-38,1995.

[Brouwer48] Brouwer L. "Consciousness, Philosophy and Mathematics", 1948, in Collected Works vol. 1 (Heyting ed.), North-Holland, 1975

[Castagna95] Castagna G., Ghelli G. and Longo G. "A calculus for overloaded functions with subtyping", Information and Computation, 117(1):115–135, Feb. 1995 (downloadable: http://www.di.ens.fr/users/longo/).

[Connes94] Connes A. Non-commutative Geometry, Academic Press, 1994.

[Coquand88] Coquand T., Huet G. "The calculus of Constructions" Information and Computation, 76, 95 - 120, 1988.

[vanDalen91] van Dalen D. "Brouwer's dogma of languageless mathematics and its role in his writings" Significs, Mathematics and Semiotics (Heijerman ed.), N.H., 1991.

[Dehaene98] Dehaene S. The Number Sense, OxfordUP, 1998. (Review/article downloadable from http://www.di.ens.fr/users/longo/)

[Enriques01] Enriques F. Problemi della Scienza, 1901.

[Enriques36] Enriques F. "Philosophie Scientifique", in Actes du Congrès International de Philosophie Scientifique, Paris, 1935, Hermann, Paris, vol.I, 1936.

[Faracovi82] Faracovi O. "Ragione e progresso nell'opera di Enriques", in Federigo Enriques. Approssimazione e verita', a cura di O.P. Faracovi, Belforte, Livorno, 1982.

[Floyd98] Floyd J. "Wittgenstein on Gödel and Mathematics", Conference on "Wittgenstein et les Fondements des Mathématiques", ENS, Paris, 1998 (to appear).

[Frege84] Frege G. The Foundations of Arithmetic, 1884 (Engl. transl. Evanston, 1980.)

[Friedman97] Friedman H. "Some Historical Perspectives on Certain Incompleteness Phenomena", May 21, 5 pages, draft, 1997.

[Gallier91] Gallier J., "What is so special about Kruskal's theorem and the ordinal Γ_0?" Ann. Pure. Appl.Logic, 53, 1991.

[Girard90] Girard J.Y., Lafont Y., Taylor P. Proofs and Types, Cambridge U. Press, 1989.

[Girard01] Girard J.Y., "Locus Solum", Mathematical Structures in Computer Science, vol.11, n.3, 2001.

[Goldfarb87] Goldfarb H., Jacques Herbrand: logical writings, 1987.

[Harrington85] Harrington L. et al. (eds) H. Friedman's Research on the Foundations of Mathematics, North-Holland, 1985.

[Heinzmann90] Heinzmann G. "Wittgenstein et le théorème de Gödel", Actes du Colloque sur "Wittgenstein et la philosophie aujourd'hui" (Klinschsiech), 1990.

[Husserl33] Husserl E., The origin of Geometry, part of Krysis, 1933.

[Lambek86] Lambek J., Scott P.J., Introduction to higher order Categorical Logic, Cambridge University Press, 1986.

[Lawvere76] Lawvere F.W. "Variable quantities and variable structures in topoi", in Algebra Topology and Category Theory: a collection of papers in honor of Samuel Eilenberg, A. Heller and M. Tierney (eds.), Academic Press, (101-131), 1976.

[Longo00] Longo G. "Prototype proofs in Type Theory", in Mathematical Logic Quaterly (formely: Zeit. f. Math. Logik und Grund. der Math.),vol. 46, n. 3, 2000.

[Longo01] Longo G. "The reasonable effectiveness of Mathematics and its Cognitive roots", in New Interactions of Mathematics with Natural Sciences (L. Boi ed.), World Scientific, pp. 351 - 382, 2005.

[Longo02] Longo G. "Space and Time in the foundation of Mathematics, or some challenges in the interaction with other sciences", invited lecture at the First AMS/SMF meeting, Lyon, July 2001 (published in french, see http://www.di.ens.fr/users/longo/ for an english translation).

[Longo05] Longo G. "The Cognitive Foundations of Mathematics: human gestures in proofs and mathematical incompleteness of formalisms", in Images and Reasoning, (M. Okada et al. eds.), Keio University Press, Tokio, pp. 105-134, 2005.

[Longo06] Longo G. "On the importance of Negative Results". (published in french, see : http://www.di.ens.fr/users/longo/ for an english translation)

[Longo93] Longo G., Milsted K. and Soloviev S., "The genericity theorem and parametricity in the polymorphic Lambda-calculus" Theor. Comp. Sci. vol. 121, 1993.

[Paris78] Paris J., Harrington L., "A mathematical incompleteness in Peano Arithmetic", Handbook of Mathematical Logic, Barwised ed., 1978.

[Rathjen93] Rathjen M., Weiermann A. "Proof-theoretic investigations on Kruskal's theorem." Annals of Pure and Applied Logic,60, 49–88, 1993.

[Riemann54] Riemann B. "On the hypothesis which lie at the basis of geometry", 1854 (English transl. by W. Clifford, Nature, 1873).

[Tappenden95] Tappenden J. "Geometry and generality in Frege's philosophy of Arithmetic" Synthese, n. 3, vol. 102, March 1995.

[Waismann31] Waismann F., Wittgenstein und der Wiener Kreis, Frankfurt a. M., Suhrkamp, 1967 (Waismann's notes: 1929-31). (English translation: Wittgenstein and the Vienna circle : conversations recorded by Friedrich Waismann, Barnes & Noble Books, New York, 1979)

[Weyl27] Weyl H. Philosophy of Mathematics and of Natural Sciences, 1927 (Engl. transl., Princeton University Press, Princeton, New Jersey, 1949).

[Wittgenstein68] Wittgenstein L. Philosophical Remarks. Engl. transl. by G. E. M. Anscombe, Barnes & Noble, New York, 1968.