STRUCTURAL PROOF THEORY:
Uncovering capacities of the mathematical mind

Wilfried Sieg
Carnegie Mellon University

Introduction. In one of his last published notes, Gödel claimed that Turing had committed a "philosophical error" in his paper "On computable numbers" when arguing that *mental procedures cannot go beyond mechanical ones*. Gödel was in error, as Turing did not make such a claim there or anywhere else; however, his argument (intended to refute the alleged claim) is of deep interest. Gödel points to a crucial feature of mind: it is *dynamic* in the sense that "we understand abstract terms more and more precisely, as we go on using them, and that more and more abstract terms enter the sphere of our understanding." Gödel illustrates this dynamic aspect of mind by considering the concept of set and the introduction of increasingly stronger axioms of infinity.

I am not pursuing the goal of discovering the next step of what Gödel envisions as an *effective non-mechanical procedure*, but rather of finding a *strategic mechanical one* that helps to uncover the mind's capacities as they are exhibited in mathematical practice. Such a procedure will have in its background abstract concepts, in particular, of set (or system) and function introduced in the second half of the 19th century. They were used by Dedekind when investigating newly defined abstract concepts in algebraic number theory and, deeply connected, in his foundational essays *1872* and *1888*. The generality of his concepts and methods had a deep impact on the development of modern mathematics (Hilbert, Emmy Noether, and Bourbaki's math. structuralism). However, the very same feature alienated many of his contemporaries, for example, Kronecker.

Hilbert played a central "mediating" role in subsequent developments, having been influenced by Dedekind in the abstract ways just indicated, but also by Kronecker's insistence on the constructive aspects of mathematical experience. He forms the bridge between these two extraordinary mathematicians of the 19th century and two equally remarkable logicians of the 20th, Gödel and Turing. The character of that connection is determined by Hilbert's focus on the *axiomatic method,* the associated *consistency problem* and the fundamental idea that proofs should be objects of mathematical study. In his talk *Axiomatisches Denken* (Zürich, late 1917), he suggested:

> … we must - that is my conviction - take the concept of the specifically mathematical proof as an object of investigation, just as the astronomer has to consider the movement of his position, the physicist must study the theory of his apparatus, and the philosopher criticizes reason itself.

Four years later, after remarkable logical developments in Göttingen, *proof theory* began to tackle the consistency problem for formal theories inspired by Dedekind, but using solely Kroneckerian finitist means.

Some of Hilbert's broader considerations have not yet been integrated into proof theoretic investigations; I am thinking of proof theory's *cognitive* side that was expressed in a talk Hilbert gave in 1927. He emphasized the philosophical significance of the "formula game", claiming that it is being pursued with rules "in which the technique of our thinking is expressed"; he continued:

> The fundamental idea of my proof theory is none other than to describe the activity of our understanding, to make a protocol of the rules according to which our thinking actually proceeds.

It was clear to Hilbert, as it is to us, that mathematical thinking does not proceed in the strictly regimented ways of an austere *formal theory*. But he hoped that the investigation of such theories might suffice for the solution of the consistency problem. He had stated the consistency issue for the arithmetic of real numbers as the second problem of his Paris address of 1900 – in a dramatically different intellectual context.

Part 1. Existential axiomatics. Hilbert's second problem raises the challenge

> …*To prove that they* [the axioms of arithmetic] *are not contradictory, that is, that a finite number of logical steps based upon them can never lead to contradictory results.*

Those axioms had been stated in Hilbert's paper *Über den Zahlbegriff*, written in 1899. Twelve years earlier, Kronecker had published a well-known paper with the same title and had sketched a treatment of irrational numbers without accepting the general notion. It is to the general concept that Hilbert wants to give a proper foundation – using Dedekind's *axiomatic method*, but also respecting Kronecker. Let me describe central aspects of this methodologically striking way of formulating a mathematical theory or defining a structure type; Hilbert and Bernays called this way *existential axiomatics*.

1.1. Abstract concepts. E. A. is different from the axiomatic method as it evolved in the 20th century. Dedekind and Hilbert both present the axiomatic core under the heading "Erklärung", standardly translated as "definition", but meaning *explanation*, *explication* or also *declaration*. They intend to provide a *frame for discourse*, here the discourse for irrational numbers, and it is provided by a *structural definition* that concerns systems and imposes relations between their elements. This approach shaped Dedekind's mathematical and foundational work; in Hilbert's *1899*, the structural definition of real number systems starts out with, *We think a system of things; we call these things numbers and denote them by a, b, c … We think these numbers in certain mutual relations, the precise and complete description of which is given by the following axioms.* Then the conditions are listed for a complete ordered field as in Dedekind's *1872*, except that completeness is postulated differently. Dedekind praised the introduction of concepts "rendered necessary by the frequent recurrence of complex phenomena, which could be controlled only with difficulty by the old ones" as the engine of progress in mathematics and other sciences. Second important concept Dedekind introduced: *simply infinite system*; Peano axioms.

The crucial methodological problem concerning such axiomatically characterized concepts is articulated forcefully in Dedekind's 1890-letter to Keferstein. In it he asks, whether *simply infinite systems* "exist at all in the realm of our thoughts". His affirmative answer is given by a *logical existence proof* without which, he explains, "it would remain doubtful, whether the concept of such a system does not perhaps contain internal contradictions." Dedekind's *realm of thoughts*, "the totality S of all things that can be object of my [Dedekind's] thinking", was crucial for obtaining an infinite system. In 1897, Cantor wrote to Hilbert that this totality is actually inconsistent. Thus, when Hilbert discussed the arithmetic axioms, he replaced the existence issue by a quasi-syntactic problem: no contradiction is provable from the axiomatic conditions in a *finite number of logical steps*!

1.2. Proofs. In the first sentence of the Preface to his *1888*, Dedekind emphasizes programmatically, "in science nothing capable of proof should be accepted without proof"; at the same time he claims that only common sense is needed to understand his essay. That comes at a cost: readers are asked to prove seemingly obvious truths by "the long sequence of simple inferences that corresponds to the nature of our step-by-step understanding". Though Hilbert's exposition of geometry in his *Grundlagen der Geometrie* is not a modern axiomatic presentation, it is novel: it is conceptually structured in the sense that the defining conditions for Euclidean space are grouped by concepts so that it is easy to investigate by mathematical proof "what rests on what".

But what is Dedekind and Hilbert's conception of proof? They explicate *arithmetic* and *analysis* in similar ways as starting from the defining conditions for simply infinite systems, resp. complete ordered fields. For Dedekind the object of arithmetic is formed by

> …the relations or laws, which are ***derived exclusively*** from the conditions [for a simply infinite system] and are therefore always the same in all ordered simply infinite systems, …

Neither Dedekind nor Hilbert formulate what steps are allowed in proofs, but they have an absolutely clear sense of their *formal, subject-independent character*. Hilbert stated in 1891 during a stop at a Berlin railway station that in a proper axiomatization of geometry "one must always be able to say 'tables, chairs, beer mugs' instead of 'points, straight lines, planes'." This remark is best understood if it is put side by side with a remark of Dedekind's written 15 years earlier: "All technical expressions [can be] replaced by arbitrary, newly invented (up to now meaningless) words; the edifice must not collapse if it is correctly constructed, and I claim, for example, that my theory of real numbers withstands this test." Thus, proofs leading from principles to theorems do not depend on aspects of the meaning of the technical expressions that have not been explicitly formulated or, to put it differently, they cannot be severed by a re-interpretation of the technical expressions (no counterexamples). The question, what **are** Dedekind's *simple inferences* and Hilbert's *logical steps*, remains unanswered, however.

Part 2. Formal proofs. In the Introduction to *Grundgesetze der Arithmetik*, Frege distinguishes his systematic development from Euclid's by pointing to the list of explicit inference principles for obtaining *gapless* proofs. The inferences have to satisfy normative demands, as broad insights are aimed for when tracing the gapless proof of a theorem to the principles ultimately appealed to: the latter determine the "epistemological nature of the theorem"! Frege insists that the inferences are to be conducted "like a calculation" and explains:

> I do not mean that in a narrow sense, as if it were subject to an algorithm the same as … ordinary addition or multiplication, but only in the sense that there is an algorithm at all, i.e., a totality of

rules which governs the transition from one sentence or from two sentences to a new one in such a way that nothing happens except in conformity with these rules.

Hilbert moved only slowly toward a presentation of proofs in logical calculi.

2.1. Natural reasoning. Hilbert and his students started only in 1913 to learn modern logic … from *Principia Mathematica*. In 1917-18, he gave the first course in *mathematical logic* and sketched, at the end, how analysis can be developed in ramified type theory with the axiom of reducibility. But he aimed for a framework allowing a natural and direct formalization of mathematics. The calculus of *PM* did not lend itself to that task. In 1921-22 he presented a calculus that is especially interesting for sentential logic.[i] He points to the parallelism with his axiomatization of geometry: groups of axioms are introduced for each concept *there* and for each logical connective *here*. Let me formulate the axioms for disjunction; the sole rule is *modus ponens*:

$(A \rightarrow C) \rightarrow ((B \rightarrow C) \rightarrow ((A \vee B) \rightarrow C))$

$A \rightarrow (A \vee B)$ and $B \rightarrow (A \vee B)$

In all the proof theoretic investigations of the 1920s, linear derivations from these axioms were transformed for good reasons into proof trees.

Using Hilbert's axiomatic calculus as starting point, Gentzen formulated his rule-based *calculi of natural deduction*; making and discharging assumptions were viewed as their distinctive features. His 1936-paper presents detailed, extensive mathematical developments; it has the explicit goal of showing the "naturalness" of nd calculi. Here are the E- and I-rules for & and v; the configurations that are derived with their help are sequents of the form $\Gamma \supset \psi$ with Γ containing all the assumptions on which the proof of ψ depends:

$$\frac{\Gamma \supset A \vee B \qquad \Gamma, A \supset C \qquad \Gamma, B \supset C}{\Gamma \supset C}$$

$$\frac{\Gamma \supset A}{\Gamma \supset A \vee B} \qquad\qquad \frac{\Gamma \supset B}{\Gamma \supset A \vee B}$$

I reformulated nd calculi as *intercalation calculi*. One proves *structured* sequents, applying E-rules on the lhs and I-rules on the rhs: (The E-rules take for granted that the major premise is an element of Γ, Δ.)

$$\frac{\Gamma, A; \Delta \supset \mathrm{G} \qquad \Gamma, B; \Delta \supset \mathrm{G}}{\Gamma; \Delta \supset \mathrm{G}}$$

$$\frac{\Gamma;\,\Delta \supset A}{\Gamma;\,\Delta \supset A\mathrm{v}B} \qquad \frac{\Gamma;\,\Delta \supset B}{\Gamma;\,\Delta \supset A\mathrm{v}B}$$

One fact is important for thinking that IC calculi provide the theoretical basis to a systematic search for normal proofs: they are complete in a refined sense: if ψ is a logical consequence of Γ, then there is a *normal* proof of ψ from Γ (subformula property).

Such a search procedure cannot be replaced by a decision procedure. In the 1920s the decision problem was viewed as one of the most important problems in mathematical logic, and some thought that a positive solution was utterly implausible. E.g., von Neumann took in 1924 the position that there is no way of deciding whether a statement is provable or not:

> …the undecidability is even the *conditio sine qua non* for the contemporary practice of mathematics, using as it does heuristic methods, to make any sense. The very day on which the undecidability does not obtain any more, mathematics as we now understand it would cease to exist; it would be replaced by an absolutely mechanical prescription, by means of which anyone could decide the provability or unprovability of any given sentence.

However, he also asserted, "we have no idea how to prove the undecidability." It was only twelve years later that Turing provided the idea by answering the question, what is mechanical procedure!

2.2. Local axiomatics. Turing's question was really, what effective operations on symbolic configurations can humans carry out mechanically, without thinking? As basic steps Turing isolated rule-governed operations on finite strings. The number of finite strings on which to operate is bounded, and the operations are local[ii]; thus, his machines can simulate them. Finally he proved: There is no Turing machine that solves the decision problem.

Our issue can now be paraphrased: How can we structure the mechanical generation of proofs for particular statements? We can do so, by a natural systematization and logical deepening of properly *bidirectional reasoning*. That is precisely what the IC calculi are designed to do. Recall that the E-rules are applied to premises in a goal-directed forward way, and that the I-rules are applied to a goal in a backward way. This approach has been implemented in the proof search system AProS with suitable strategies. The IC-approach has been extended to elementary set theory, where it involves now also the use of definitions and lemmata.

Hilbert's grouping of the axioms for geometry had the express purpose of organizing proofs and the whole subject in a conceptual way: his development contains marvelous instances of *local axiomatics*, analyzing which notions and principles are needed for which theorems. This technique is invoked in later metamathematical investigations, famously in the proof of Gödel's 2nd theorem. To illustrate this point, I consider Gödel's incompleteness theorems and Löb's theorem. The proofs of these theorems make use of the connection between the metamathematics in which a formal theory is presented and the mathematics that can be formally developed in the theory. Three components are crucial for their abstract, i.e., locally axiomatic proofs:

1. *Local axioms*: representability of the core syntactic notions, the diagonal lemma, and Hilbert & Bernays's derivability conditions.
2. *Proof-specific definitions*: instances of existential claims, for example, the Gödel sentence for the first incompleteness theorem.
3. *Leading idea*: moving between object- and meta-theory, expressed by appropriate E- and I-rules for the theorem predicate.[iii]

The above components go beyond the logical strategies and have been used to expand them: AProS finds the proofs directly. A second, human point of interest: von Neumann discovered a proof of Gödel's 2nd theorem in 1931, independently, and used essentially the local axioms.

Part 3. Conceptually structured search. In his report on *Intelligent Machinery*, Turing suggested having machines search for proofs of theorems in formal systems. It was clear to him that one cannot just specify axioms and logical rules, then state a theorem and expect a machine to demonstrate it. To exhibit such intelligence, he claims, a machine must "acquire both discipline and initiative". Discipline can be acquired by becoming a universal computer; but "discipline is certainly not enough in itself to produce intelligence" and he continues:

> That which is required in addition we call initiative. This statement will have to serve as a definition. Our task is to discover the nature of this residue as it occurs in man, and try and copy it in machines.

The dynamic character of logical strategies constitutes but a partial, very limited copy of human initiative in the mathematical context or, in Gödel's terms, of the dynamic nature of mind; the introduction of suitable concepts associated with local axioms and leading ideas is crucial. I want to sketch two examples; we can discuss details after my presentation.

3.1. Axiomatics, modern. Local axiomatic developments can be integrated into a broader framework via a hierarchical organization; that has been part and parcel of mathematical practice. With a group of students I have been working on obtaining an automated proof of the Cantor-Bernstein theorem. The theorem claims, as you know, that there is a bijection between two sets, in case there are injections from the first to the second and from the second to the first. We are finishing the formal development following AProS's logical strategies and starting from Zermelo's axioms. Three layers are used for the organization of the proof: i) Construction of sets, e.g. empty set, power set, union, and pairs; ii) Definition of functions as sets; facts on compositions, restrictions etc.; iii) The abstract proof.

The abstract proof is completely independent of the set theoretic definition of functions. It is divided in roughly the same way as that of Gödel's theorems. The *local axioms* here are lemmata for injective, surjective, and bijective functions; the crucial *proof-specific definition* is that of the bijection claimed to exist in the theorem; the *leading idea* is to obtain a 2-partition of domain and co-domain in such a way that the parts are related by bijections.

Shaping a field and its proofs by concepts is classical; so is the deepening of its foundations. That can be illustrated by the developments in the first two books of Euclid's *Elements*. Proposition I.47, the Pythagorean theorem, is at their center. The mathematical context is given by the *quadrature problem*, i.e., the issue of determining the size of geometric figures in terms of squares. That problem is discussed in Book II for polygons. Polygons can be partitioned into triangles that can be transformed individually first into equal rectangles and then into equal squares. The question is, how can we join these squares to obtain one single square that is equal to the polygon we started out with?

3.2. Axiomatics, classical. It is precisely here that the Pythagorean theorem comes in and provides the most direct way of determining the larger square. Byrne's colorful diagram captures the construction and depicts the memorable idea of *one* proof of the theorem, Euclid's favorite proof. Euclid constructs the squares on the triangle's sides and observes that the extensions of the sides of the smaller squares by the contiguous sides of the original triangle are straight lines. In the next step a crucial auxiliary line is drawn, namely, the line that is perpendicular to the hypotenuse and passes through the vertex opposite the hypotenuse. This line partitions the big square into the blue and yellow rectangle. Two claims are considered: i) the blue rectangle is equal to the black square, and ii) the yellow rectangle is equal to the red square.

Euclid uses three facts that are readily obtained from earlier propositions: (α) Triangles are equal when they have two equal sides and when the enclosed angles are equal; (β) Triangles are equal when they have the same base and when their third vertex lies on the same parallel to that base; (γ) A diagonal divides a rectangle into two equal triangles. With (α) through (γ), there is an easy proof that the red square and the yellow rectangle are the same and, analogously, that the black square and the blue rectangle are equal.

The structure of this argument is "similar" to that of the abstract proofs of the incompleteness theorems and the Cantor-Bernstein theorem: (α) through (γ) are the *local axioms*. What corresponds to the central *proof-specific definition*? Euclid's formulation of the theorem! That is obvious when the underlying problem is recalled and the theorem formulated as: Given two squares A and B find a square C, such that C equals A and B. (Remark: geometric computation; Euclid's proof verifies correctness!) The *leading idea* is the partitioning of squares and establishing that corresponding parts are equal. This idea of dissecting figures is pervasive: here is a diagram "showing" a Chinese proof from the second century BC; Paul Mahlo investigated in his 1905-thesis dissection proofs of the Pythagorean theorem from a topological perspective. Deepening: Euclid's and Zermelo's axioms!

4. Cognitive aspects. In *1888* Dedekind refers to his *Habilitationsrede*, in which he claimed that the need to introduce appropriate notions arises from the fact that human intellectual powers are imperfect. Their limitation leads us to frame the object of a science in different forms, and introducing a concept means, in a certain sense, formulating a hypothesis on the inner nature of the science. How well the concept captures that inner nature is determined by its usefulness for the development of the science, in mathematics mainly by its usefulness for constructing proofs. Thus, Dedekind viewed abstract concepts and general forms of arguments as tools to overcome, at least partially, the imperfection of our intellectual powers. He remarked:

Essentially there would be no more science for a man gifted with an unbounded understanding – a man for whom the final conclusions, which we obtain through a long chain of inferences, would be immediately evident truths; and this would be so even if he stood in exactly the same relation to the objects of science as we do. (*Ewald*, pp. 755-6)

The theme of a particular form of human understanding is sounded in a remark of Bernays in *1950*, where he writes: "Though for differently built beings there might be a different kind of evidence, it is nevertheless our concern to find out what is evident *for us*, not for some *differently built being*." In his later writings Bernays emphasized that mathematical evidence is acquired by *intellectual experience* and through *experimentation*:

In this way we recognize something like intelligence and reason that should not be regarded as a container of [items of] a priori knowledge, but as a mental activity that consists in reacting to given situations with the formation of experimentally applied categories. (Bernays 1946)

Such intellectual experimentation supports in particular the introduction of abstract concepts, and the proofs of central theorems using these concepts are at the heart of such experimentation. Even for Gödel, each step in the process of formulating stronger axioms of infinity is based on further mathematical experience requiring "a substantial advance in our understanding of the basic concepts of mathematics". In a note published in Wang's *1974*, Gödel offers a different reason for not yet having a precise definition of that process: it "would require a substantial deepening of our understanding of the basic operations of the mind".

Explicitly, I wanted to turn attention to capacities of the mind that are central, if we want to connect reasoning and mathematical understanding, i.e., if we want to see the role of leading ideas in guiding proofs and that of concepts in structuring proofs. When we focus on formal methods and carry out proof search experiments, we isolate creative elements in proofs and come to a deeper understanding of the technique of our mathematical thinking. Implicitly, I argued for an *expansion of proof theory*: let's move toward a theory that articulates principles for organizing proofs and for finding them dynamically! A good start is a thorough reconstruction of parts of the rich body of mathematical knowledge that is *systematic*, but is also *structured for intelligibility and discovery*. Such an expanded proof theory can be called *structural* for two reasons: on the one hand it exploits the internal structure of formal proofs and uses, on the other hand, the notions and principles for mathematical structures in a well-integrated way.

Let me end with remarks by a mathematician who completed his thesis with the title "*Abgekürzte Beweise im Logikkalkul*" in Göttingen, in 1933. Two years later when commenting on it, he remarked that proofs are not "mere collections of atomic processes, but rather complex combinations with a highly rational structure"; forty-four years after that, he ended a review of his early logical work by saying: "There remains the real question of the actual structure of mathematical proofs and their strategy. It is a topic long given up by mathematical logicians, but one which still – properly handled – might give us some real insight." The mathematician was a close friend of Gentzen; his name: Saunders MacLane.

---

[i] In 1931b, Hilbert formulates what he calls "definitional rules" for the quantifiers (essentially the standard natural deduction rules); those formulations were of course of no interest in the earlier proof theoretic investigations, as quantifiers were eliminated for the sake of the epsilon-substitution method.

[ii] That is actually the way of Turing's analysis in 1936 and leads essentially to Post production systems; this line is taken up in 1954 when Turing argues that puzzles are reducible to substitution puzzles.

[iii] For example, if a proof of A has been obtained in the object-theory, then one is allowed to introduce the claim 'A is provable' in the meta-theory.