STUDIA SEMIOTYCZNE (SEMIOTIC STUDIES), t. XXXIV, nr 1 (2020), s. 57–80
ISSN 0137-6608, e-ISSN 2544-073X
DOI: 10.26333/sts.xxxiv1.03

Essay

WILFRIED SIEG *

GÖDEL'S PHILOSOPHICAL CHALLENGE (TO TURING)[1]

SUMMARY: The incompleteness theorems constitute the mathematical core of Gödel's philosophical challenge. They are given in their "most satisfactory form", as Gödel saw it, when the formality of theories to which they apply is characterized via Turing machines. These machines codify human mechanical procedures that can be carried out without appealing to higher cognitive capacities. The question naturally arises, whether the theorems justify the claim that the human mind has mathematical abilities that are not shared by any machine. Turing admits that non-mechanical steps of intuition are needed to transcend particular formal theories. Thus, there is a substantive point in comparing Turing's views with Gödel's that is expressed by the assertion, "The human mind infinitely surpasses any finite machine". The parallelisms and tensions between their views are taken as an inspiration for beginning to explore, computationally, the capacities of the human mathematical mind.[2]

KEYWORDS: computability, Church's Thesis, Turing's Thesis, incompleteness, undecidability, Post production systems, computable dynamical systems.

* Carnegie Mellon University. E-mail: sieg@cmu.edu. ORCID: 0000-0002-7130-0524.

[1] This essay by Wilfried Sieg, "Gödel's Philosophical Challenge (to Turing)" was previously published in Copeland, B. Jack, Carl J. Posy, and Oron Shagrir, eds., *Computability—Turing, Gödel, Church, and Beyond*, © 2013 Massachusetts Institute of Technology; it is reprinted here by permission of The MIT Press. The author has added a new Postscriptum in June 2020.

[2] In Milan Kundera's *Ignorance* (2002) one finds on page 124, "We won't understand a thing about human life if we persist in avoiding the most obvious fact: that a reality no longer is what it was when it was; it cannot be reconstructed". These remarks of Kundera, born in Gödel's hometown Brno, apply even to attempts of understanding and reconstructing a limited aspect of past intellectual life.

Introduction

"To Turing" is flanked by parentheses in the title, as the philosophical challenge issued by Gödel's mathematical results, the incompleteness theorems, was not only a challenge to Turing but also to Gödel himself; it certainly should be taken up by us. At issue is the question whether there is a rigorous argument from these results to the claim that machines can never replace mathematicians or, more generally, that the human mind infinitely surpasses any finite machine. Gödel made the former claim already in 1939; the latter assertion was central in his Gibbs Lecture of 1951. In his note of 1972, Gödel tried to argue for that assertion with greater emphasis on subtle aspects of mathematical experience in set theory. He explored, in particular, the possibility of a humanly effective, but non-mechanical process for presenting a sequence of ever-stronger axioms of infinity.

To understand the claims in their broad intellectual context, one is almost forced to review the emergence of a rigorous notion of computability in the early part of the twentieth century. Gödel's role in that emergence is "dichotomous", as John Dawson noted in his lecture (2006). There are crucial impulses, like the definition of general recursive functions in the 1934 Princeton Lectures. This definition was the starting point for Kleene's work in recursion theory and served as the rigorous mathematical notion in Church's first published formulation of his "thesis" in (1935). However, there is neither a systematic body of recursion theoretic work nor an isolated central theorem that is associated with Gödel's name. The reason for that is clear: Gödel was not interested in developing the theory, but rather in securing its conceptual foundation. He needed such a foundation for two central and related purposes, namely, (i) to formulate the incompleteness theorems in mathematical generality for all formal theories (containing arithmetic) and (ii) to articulate and sharpen philosophical consequences of the undecidability and incompleteness results.

The philosophical consequences, as I indicated, are concerned with the claimed superiority of the human mind over machines in mathematics. This takes for granted that a convincing solution to the issue indicated under (i) has been found and that such a solution involves suitably characterized machines. The first two parts of this essay, entitled *Primitive & General Recursions* and *Finite Machines & Computors*, present the general foundational context. It is only then that the central philosophical issue is addressed in the third part, *Beyond Mechanisms & Discipline*. Gödel's and Turing's views on mind are usually seen in sharp opposition to each other. Indeed, Gödel himself claimed to have found a "philosophical error in Turing's work"; his argument for such an error rests on the (incorrect) assumption that Turing tried to establish in (1936) that mental procedures do not go beyond mechanical ones. If one focuses on the real challenge presented by the incompleteness theorems, then one finds that Gödel and Turing pursue parallel approaches with complementary programmatic goals, but dramatically different methodological perspectives. Concrete work to elucidate the situation is suggested in the last part of the essay, *Finding Proofs (With Ingenuity)*.

I. Primitive & General Recursions

It was of course Kronecker who articulated in the 1870s forcefully the requirement that mathematical objects should be finitely presented, that mathematical notions should be decidable, and that the values of functions should be calculable in finitely many steps. And it was of course Dedekind who formulated in his essay *Was sind und was sollen die Zahlen?* the general schema of primitive recursion. At the turn from the nineteenth to the twentieth century, Hilbert transferred Kronecker's normative requirements from mathematics to the frameworks in which mathematical considerations were to be presented, i.e., to axiomatic theories. This shift was accompanied by a methodologically sound call for proofs to establish the theories as consistent.[3] A syntactic and, in Hilbert's view, first "direct" consistency proof was given in his (1905) for a purely equational theory. The approach was criticized fairly by Poincaré and, for a long time, not pursued further by Hilbert. Only in 1921 did Hilbert come back to this particular argument and used it as the starting point of novel proof theoretic investigations, now with a finitist foundation that included recursion equations for all primitive recursive functions as basic principles.[4]

In order to carry out the proof theoretic arguments, functions in formal theories have to be calculable, indeed, calculable from a finitist perspective. That is clear from even a rough outline of the consistency proof Hilbert and Bernays obtained in early 1922. It was presented in (Hilbert, 1923) and concerns the quantifier-free theory we call primitive recursive arithmetic (PRA) and proceeds as follows. The linear proofs are first transformed into tree structures; then all variables are systematically replaced by numerals resulting in a configuration of purely numeric statements that all turn out to be true and, consequently, cannot contain a contradiction. Yet to recognize the truth of the numeric formulae one has to calculate, from a finitist perspective, the value of functions applied to numerals.[5] This was a significant test of the new proof theoretic techniques, but the result had one drawback: a consistency proof for the finitist system PRA was not needed according to the programmatic objectives, but a treatment of quantifiers was required. Following Hilbert's *Ansatz* of eliminating quantifiers in favor of ε-terms, Ackermann carried out the considerations for "transfinite" theories, i.e., for the first-order extension of PRA (correctly, as it turned out, only with just

---

[3] This is in the logicist tradition of Dedekind (cf. Sieg & Schlimm, 2005; Sieg, 2009a).

[4] For the development of Hilbert's foundational investigations, it has to be mentioned that the Göttingen group had in the meantime assimilated Whitehead and Russell's *Principia Mathematica*; that is clear from the carefully worked out lecture notes from the winter term 1917–1918; cf. (Sieg, 1999).

[5] That was done in (Hilbert & Bernays, 1921/2); a summary is found in Section II of (Ackermann, 1925), entitled *The Consistency Proof Before the Addition of the Transfinite Axioms*. Ackermann does not treat the induction rule, but that can easily be incorporated into the argument following Hilbert and Bernays. The presentation of these early proof theoretic results is refined and extended in (Hilbert & Bernays, 1934).

quantifier-free induction). Herbrand obtained in 1931 the result for essentially the same system, but with recursion equations for a larger class of finitistically calculable functions; that is how Herbrand described the relation of his result to that of Ackermann in a letter of 7 April, 1931, to Bernays.

As to the calculability of functions, Hilbert and Bernays had already emphasized in their lectures from 1921–1922, "For every single such definition by recursion it has to be determined that the application of the recursion formula indeed yields a number sign as function value—for each set of arguments". Such a determination was taken for granted for primitive recursive definitions. We find here, in a rough form, Herbrand's way of characterizing broader classes of finitistically calculable functions according to the schema in his 1931 letter to Gödel:

> In arithmetic, we have other functions as well, for example functions defined by recursion, which I will define by means of the following axioms. Let us assume that we want to define all the functions $f_n(x_1, x_2, \ldots, x_{pn})$ of a certain finite or infinite set F. Each $f_n(x_1, \ldots)$ will have certain defining axioms; I will call these axioms (3F). These axioms will satisfy the following conditions:
>
> (i) The defining axioms for f_n contain, besides f_n, only functions of lesser index.
>
> (ii) These axioms contain only constants and free variables.
>
> (iii) We must be able to show, by means of intuitionistic proofs, that with these axioms it is possible to compute the value of the functions univocally for each specified system of values of their arguments. (This letter is found in [Gödel, 2003].)

Having given this schema, Herbrand mentions that the non-primitive recursive Ackermann function falls under it. Recall that Herbrand, as well as Bernays and von Neumann at the time, used "intuitionistic" as synonymous with "finitist".

In two letters from early 1931, Herbrand and Gödel discussed the impact of the incompleteness theorems on Hilbert's Program. Gödel claimed that some finitist arguments might not be formalizable even in the full system of *Principia Mathematica*; in particular, he conjectured that the finitist considerations required for guaranteeing the unicity of the recursion axioms are among them. In late 1933, Gödel gave a lecture in Cambridge (Massachusetts) and surveyed the status of foundational investigations; see (Gödel, 1933). This fascinating lecture describes finitist mathematics and reveals a number of mind changes: (i) when discussing calculable functions, Gödel emphasizes their recursive definability, but no longer the finitist provability requirement, and (ii) when discussing Hilbert's Program, Gödel asserts that a l l finitist considerations can be formalized in elementary number theory. He supports his view by saying that finitist considerations use only the proof and definition principle of complete induction; the class of functions definable in this way includes all those given by Herbrand's schema. I take Gödel's deliberate decision to disregard the provability condition as a first and very significant step toward the next major definition, i.e., that of general recursive functions.

A few months after his lecture in Cambridge, Gödel was presented with Church's proposal of identifying the calculability of number-theoretic functions with their λ-definability. Gödel, according to Church in a letter of 29 November, 1935, to Kleene, viewed the proposal as "thoroughly unsatisfactory" and proposed "to state a set of axioms which would embody the generally accepted properties of this notion [i.e., effective calculability], and to do something on that basis" (in Sieg, 1997, p. 463). However, instead of formulating axioms for that notion in his 1934 Princeton lectures, Gödel took a second important step in further modifying Herbrand's definition. He considered as general recursive those total number theoretic functions whose values can be computed in an equational calculus, starting with general recursion equations and proceeding with very elementary replacement rules. In a 1964 letter to van Heijenoort, Gödel asserted, "… it was exactly by specifying the rules of computation that a mathematically workable and fruitful concept was obtained".[6]

Gödel had obviously defined a broad class of calculable functions, but at the time he did not think of general recursiveness as a rigorous explication of calculability.[7] Only in late 1935 did it become plausible to him, as he put it on 1 May, 1968, in a letter to Kreisel, "that my [incompleteness] results were valid for all formal systems". The plausibility of this claim rested on an observation concerning computability in the Postscriptum to his 1936-note, *On the Length of Proofs*. Here is the observation for systems S_i of i-th order arithmetic, $i > 0$.

> It can, moreover, be shown that a function computable in one of the systems S_i, or even in a system of transfinite order, is computable already in S_1. Thus, the notion "computable" is in a certain sense "absolute", while almost all metamathematical notions otherwise known (for example, provable, definable, and so on) quite essentially depend upon the system adopted. (Gödel, 1936, p. 399)

Ten years later, in his contribution to the Princeton Bicentennial Conference, Gödel formulated the absoluteness claim not just for higher-type extensions of arithmetic, but for any formal system containing arithmetic, in particular, for set theory. The philosophical significance of general recursiveness is almost exclusively attributed to its absoluteness. Connecting his remarks to a previous lecture given by Tarski, Gödel started his talk with:

> Tarski has stressed in his lecture (and I think justly) the great importance of the concept of general recursiveness (or Turing's computability). It seems to me that this importance is largely due to the fact that with this concept one has for the first time succeeded in giving an absolute definition of an interesting epistemological notion, i.e., one not depending on the formalism chosen. (Gödel, 1946, p. 150)

[6] For brief descriptions of the equational calculus see Gödel's (1934, pp. 368–369) or his (193?, pp. 166–168).

[7] Cf. his letter to Martin Davis quoted in (Davis, 1982, p. 9).

In 1965, Gödel added a footnote to this remark clarifying the precise nature of the absoluteness claim:

> To be more precise: a function of integers is computable in any formal system containing arithmetic if and only if it is computable in arithmetic, where a function *f* is called computable in *S* if there is a computable term representing *f*.

The metamathematical absoluteness claim as formulated in 1936 can readily be established for the specific theories of higher-order arithmetic. However, in order to prove the claim that functions computable in any formal system containing arithmetic are general recursive, the formal nature of the systems has to be rigorously characterized and then exploited. One can do that, for example, by imposing on such systems the recursiveness conditions of Hilbert and Bernays that were formulated in Supplement II of the second volume of their *Grundlagen der Mathematik*. When proceeding in this way one commits, however, a subtle circularity in case one simultaneously insists that the general recursive functions allow the proper mathematical characterization of formality.[8]

In Gödel's 1946 Princeton remark, "Turing's computability" is mentioned, but is listed parenthetically behind general recursiveness without any emphasis that it might play a special role. That notion becomes a focal point in Gödel's reflections only in the 1951 Gibbs Lecture where he explores the implications of the incompleteness theorems, not in their original formulation, but rather in a "much more satisfactory form" that is "due to the work of various mathematicians". He stresses, "The greatest improvement was made possible through the precise definition of the concept of finite procedure, which plays such a decisive role in these results".[9] Gödel points out that there are different ways of arriving at a precise definition of finite procedure, which all lead to exactly the same concept. However, and here is the observation on Turing,

> The most satisfactory way … [of arriving at such a definition] is that of reducing the concept of finite procedure to that of a machine with a finite number of parts, as has been done by the British mathematician Turing. (Gödel, 1951, pp. 304–305)

Gödel does not expand on this brief remark; in particular, he gives no hint of how reduction is to be understood. He also does not explain, why such a reduction is "the most satisfactory way" of getting to a precise definition or, for

---

[8] This is analyzed in section 2 of (Sieg, 1994) and with an illuminating Churchian perspective, in section 4 of (Sieg, 1997).

[9] In a footnote Gödel explains that the concept of "finite procedure" is considered to be equivalent to the concept of a "computable function of integers", i.e., a function *f* "whose definition makes it possible actually to compute *f*(*n*) for each integer *n*". The reason why that can be done is formulated as follows: "The procedures to be considered do not operate on integers but on formulas, but because of the enumeration of the formulas in question, they can always be reduced to procedures operating on integers".

that matter, why the concept of a machine with a finite number of parts is equivalent to that of a Turing machine. At this point, it seems, the ultimate justification lies in the pure and perhaps rather crude fact that finite procedures can be effected by finite machines.[10]

Gödel claims in the Gibbs Lecture (1951, p. 311) that the state of philosophy "in our days" is to be faulted for not being able to draw in a mathematically rigorous way the philosophical implications of the "mathematical aspect of the situation", i.e., the situation created by the incompleteness results. I have argued that not even the mathematical aspect had been clarified in a convincing way; after all, it crucially depended on very problematic considerations concerning a precise notion of computability.

II. Finite Machines & Computors

To bring out very clearly that the appeal to a reduction is a most significant step for Gödel, let me go back to the informative manuscript (Gödel, 193?) from the late 1930s. In it, Gödel examines general recursiveness and Turing computability, but under a methodological perspective that is completely different from the one found in the Gibbs Lecture. After having given a perspicuous presentation of his equational calculus, Gödel claims outright that it provides "the correct definition of a computable function". Thus, he seems to be fully endorsing Church's Thesis concerning general recursive functions. He adds a remark on Turing asserting, "That this really is the correct definition of m e c h a n i c a l computability was established beyond any doubt by Turing". How did Turing establish this claim? Here is Gödel's answer:

> [Turing] has shown that the computable functions defined in this way [via the equational calculus] are exactly those for which you can construct a machine with a finite number of parts which will do the following thing. If you write down any number $n_1, \ldots, n_r$ on a slip of paper and put the slip of paper into the machine and turn the crank, then after a finite number of turns the machine will stop and the value of the function for the argument $n_1, \ldots, n_r$ will be printed on the paper. (Gödel, 193?, p. 168)

The mathematical theorem stating the equivalence of Turing computability and general recursiveness plays the pivotal role at this time: Gödel does not yet focus

[10] In his (1933, p. 45) Gödel describes the constructivity requirements on theories and explicates the purely formal character of inference rules. The latter "refer only to the outward structure of the formulas, not to their meaning, so that they could be applied by someone who knew nothing about mathematics, or by a machine". He also asserts there, "thus the highest possible degree of exactness is obtained".

on Turing's analysis as being the basis for a reduction of mechanical calculability to (Turing) machine computability.[11]

The appreciation of Turing's work indicated in the Gibbs Lecture for the first time is deepened in other writings of Gödel. Perhaps, it would be better to say that Turing's work appears as a topic of perceptive, but also quite aphoristic remarks. Indeed, there are only three such remarks that were published during Gödel's lifetime after 1951: (i) the *Postscriptum* to the 1931 incompleteness paper, (ii) the *Postscriptum* to the 1934 Princeton Lecture Notes, and (iii) the 1972 note *A Philosophical Error in Turing's Work*. The latter note appeared in a slightly different version in Wang's book from 1974. In the sequel, I will refer to the "1972-note" and the "1974-note", though I am convinced that the first note is the later one.

The brief *Postscriptum* added to (Gödel, 1931) in 1963 emphasizes the centrality of Turing's work for both incompleteness theorems; here is the text:

> In consequence of later advances, in particular of the fact that due to A. M. Turing's work a precise and unquestionably adequate definition of the general notion of formal system can now be given, a completely general version of Theorems VI and XI is now possible. That is, it can be proved rigorously that in every consistent formal system that contains a certain amount of finitary number theory there exist undecidable arithmetic propositions and that, moreover, the consistency of any such system cannot be proved in the system. (Gödel, 1931, p. 195)

In the more extended *Postscriptum* written a year later for his Princeton Lecture Notes, Gödel repeats this remark almost verbatim, but then states a reason why Turing's work provides the basis for a "precise and unquestionably adequate definition of the general concept of formal system": "Turing's work gives an analysis of the concept of 'mechanical procedure' (alias 'algorithm' or 'computation procedure' or 'finite combinatorial procedure'). This concept is shown to be equivalent with that of a 'Turing machine'" (Gödel, 1934, pp. 369–370).

In a footnote attached to the last sentence Gödel refers to (Turing, 1936) and points to its ninth section, where Turing argues for the adequacy of his machine concept. Gödel emphasizes that previous equivalent definitions of computability, including general recursiveness and λ-definability, "are much less suitable for our purposes". However, he does not elucidate the special character of Turing computability in this context or any other context I am aware of, and he does not indicate either, how he thought an analysis proceeded or how the equivalence

[11] In the spring of 1939, Gödel gave a logic course at the University of Notre Dame and argued for the superiority of the human mind over machines via the undecidability of the decision problem for predicate logic; the latter is put into the historical context of Leibniz's *Calculemus*! He claims: "So here already one can prove that Leibnitzens [sic!] program of the *calculemus* cannot be carried through, i.e. one knows that the human mind will never be able to be replaced by a machine already for this comparatively simple question to decide whether a formula is a tautology or not". The conception of machine is as in (193?)—an office calculator with a crank.

between the (analyzed) concept and Turing computability could be shown. In the next paragraph, I will give a very condensed version of Turing's important argument, though I note right away that Turing did not view it as proving an equivalence result of the sort Gödel described.[12]

Call a human computing agent who proceeds mechanically a computor; such a computor operates deterministically on finite, possibly two-dimensional configurations of symbols when performing a calculation.[13] Turing aims to isolate the most basic steps taken in calculations, i.e., steps that need not be further subdivided. This goal requires that the configurations on which the computor operates be immediately recognizable. Joining this demand with the evident limitation of the computor's sensory apparatus leads to the "boundedness" of configurations and the "locality" of operations:

(B) There is a fixed finite bound on the number of configurations a computor can immediately recognize; and

(L) A computor can change only immediately recognizable (sub-) configurations.

As Turing considers the two-dimensional character of configurations as inessential for mechanical procedures, the calculations of the computor, satisfying the boundedness and locality restrictions, are directly captured by Turing machines operating on strings; the latter can provably be mimicked by ordinary two-letter Turing machines.[14]

So, it seems we are naturally and convincingly led from calculations of a computor on two-dimensional paper to computations of a Turing machine on a linear tape. Are these machines in the end, as Turing's student Gandy put it, nothing but codifications of computors? Is Gandy right when claiming in (1980, p. 124) that Turing's considerations provide (the outline of) a proof for the claim, "What can be calculated by an abstract human being working in a routine way is computable?" Does Turing's argument thus secure the conclusiveness and generality of the limitative mathematical results, respect their broad intellectual

[12] I have analyzed Turing's argument in other papers (e.g., 1994; 2002). My subsequent discussion takes Turing machines in the way in which Post defined them in (1947), namely, as production systems. That has the consequence that states of mind are physically represented, quite in Turing's spirit; cf. part III of section 9 in his paper (1936) and the marvelous discussion in (Turing, 1954).

[13] That captures exactly the intellectual problematic and context: the *Entscheidungsproblem* was to be solved mechanically by us; formal systems were to guarantee intersubjectivity on a minimal, mechanically verifiable level between us.

[14] It should be noted that step-by-step calculations in the equational calculus cannot be carried out by a computor satisfying these restrictive conditions: arbitrarily large numerals have to be recognized and arbitrarily complex terms have to be replaced by their numerical values—in one step.

context and appeal only to mechanical procedures that are carried out by humans without the use of higher cognitive capacities?

Turing himself found his considerations mathematically unsatisfactory. Indeed, he took two problematic steps by (i) starting the analysis with calculations on two-dimensional paper (this is problematic as possibly more general configurations and procedures should be considered) and (ii) dismissing, without argument, the two-dimensional character of paper as "no essential of computation". However, a restricted result is rigorously established by Turing's considerations: Turing machines can carry out the calculations of computors—as long as computors not only satisfy (*B*) and (*L*), but also operate on linear configurations; this result can be extended to extremely general configurations, K-graphs.[15] But even then, there is no proof of Turing's Thesis.

The methodological difficulties can be avoided by taking an alternative approach, namely, to characterize a Turing Computor axiomatically as a discrete dynamical system and to show that any system satisfying the axioms is computationally reducible to a Turing machine (Sieg, 2002; 2009a). No appeal to a thesis is needed; rather, that appeal has been replaced by the task of recognizing the correctness of axioms for an intended notion. This way of extracting from Turing's analysis clear axiomatic conditions and then establishing a representation theorem seems to follow Gödel's suggestion to Church in 1934; it also seems to fall, in a way, under the description Gödel gave of Turing's work, when arguing that it analyzes the concept "mechanical procedure" and that "this concept is shown to be equivalent with that of a Turing machine".[16]

With the conceptual foundations in place, we can examine how Gödel and Turing thought about the fact that humans transcend the limitations of any particular Turing machine (with respect to the first incompleteness theorem). They chose quite different paths: Gödel was led to argue for the existence of humanly effective, non-mechanical procedures and continued to identify finite machines with Turing machines; thus, he "established" our topical claim that the human mind infinitely surpasses any finite machine. Turing, by contrast, was led to the more modest demand of releasing computors and machines from the strict discipline of carrying out procedures mechanically and providing them with room for initiative. Let us see what that amounts to.

III. Beyond Mechanisms & Discipline

Gödel's paper (193?) begins by referring to Hilbert's famous words, "for any precisely formulated mathematical question a unique answer can be found".

---

[15] The underlying methodological matters are discussed in (Sieg & Byrnes, 1996), where K-graphs were introduced as a generalization of the graphical structures considered in (Kolmogorov & Uspenski, 1963).

[16] In (Martin, 2005), a particular (and insightful) interpretation of Gödel's view on mathematical concepts is given. It is developed with special attention to the concept of set, but it seems to be adaptable to the concept of computability. Cf. the summary on pp. 223–224.

Those words are taken to assert that for any mathematical proposition A there is a proof of either A or not-A, "where by 'proof' is meant something which starts from evident axioms and proceeds by evident inferences". He argues that the incompleteness theorems show that something is lost when one takes the step from this notion of proof to a formalized one:

> [I]t is not possible to formalise (sic!) mathematical evidence even in the domain of number theory, but the conviction about which Hilbert speaks remains entirely untouched. Another way of putting the result is this: it is not possible to mechanise (sic!) mathematical reasoning […]. (Gödel, 193?)

And that means for Gödel that "it will never be possible to replace the mathematician by a machine, even if you confine yourself to number-theoretic problems" (pp. 164–165). Gödel took this deeply rationalist and optimistic perspective still in the early 1970s: Wang reports that Gödel rejected the possibility that there are number theoretic problems undecidable for the human mind (Wang, 1974, pp. 324–325).[17]

Gödel's claim that it is impossible to mechanize mathematical reasoning is supported in the Gibbs Lecture by an argument that relies primarily on the second incompleteness theorem; see the detailed analyses in (Feferman, 2006a) and (Sieg, 2007, Section 2). This claim raises immediately the question, "What aspects of mathematical reasoning or experience defy formalization?" In his 1974-note, Gödel points to two "vaguely defined" processes that may be sharpened to systematic and effective, but non-mechanical procedures; namely, the process of defining recursive well-orderings of integers for larger and larger ordinals of the second number class and that of formulating stronger and stronger axioms of infinity. The point is reiterated in the modified formulation of the 1972-note, where Gödel, on p. 305, considers another version of his first theorem that may be taken "as an indication for the existence of mathematical yes or no questions undecidable for the human mind". However, he points to a *fact* that in his view weighs against such an interpretation: "There *do* exist unexplored series of axioms which are analytic in the sense that they only explicate the concepts occurring in them". As an example, he again presents axioms of infinity, "which only explicate the content of the general concept of set". These reflections on axioms of infinity and their impact on provability are foreshadowed in (Gödel, 1947, p. 182), where Gödel asserts that the current axioms of set theory "can be supplemented without arbitrariness by new axioms which are only the natural continuation of the series of those [axioms of infinity] set up so far". So, there may be a completeness theorem stating, "every proposition expressible in set theory is decidable from the present axioms plus some true assertion about the largeness of the universe of all sets".

[17] For a broad discussion of Gödel's reflections on "absolutely unsolvable problems", cf. (Feferman, 2006a; Kennedy, van Atten, 2004; 2009).

Though Gödel calls the existence of an unexplored series of axioms a fact, he asserts also that the process of forming such a series does not yet form a "well-defined procedure which could actually be carried out (and would yield a non-recursive number-theoretic function)", because it would require "a substantial advance in our understanding of the basic concepts of mathematics" (Gödel, 1972, p. 306). A *prima facie* startlingly different reason for not yet having a precise definition of such a procedure is given in the 1974-note, p. 325: it would require "a substantial deepening of our understanding of the basic operations of the mind". That is only prima facie different, as Gödel's 1972-note connects such a procedure with the dynamic development of the human mind.

> [M]ind, in its use, is not static, but constantly developing, i.e., 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, 1972, p. 306)[18]

Gödel continues:

> There may exist systematic methods of actualizing this development, which could form part of the procedure. Therefore, although at each stage the number and precision of the abstract terms at our disposal may be finite, both [...] may converge toward infinity in the course of the application of the procedure.

The procedure mentioned as a plausible candidate for satisfying this description is again the process of forming ever stronger axioms of infinity.

The notes (1972) and (1974) are very closely connected, but there is a subtle and yet, it seems to me, substantive difference. In the 1974-note the claim that the number of possible states of mind may converge to infinity is a consequence of the dynamic development of mind. That claim is followed by a remark that begins in a superficially similar way as the first sentence of the above quotation, but ends with a quite different observation: "Now there may exist systematic methods of accelerating, specializing, and uniquely determining this development, e.g. by asking the right questions on the basis of a mechanical procedure" (Gödel 1974, p. 325).

---

[18] Gödel's brief exploration of the issue of defining a non-mechanical, but effective procedure is preceded in this note by a severe critique of Turing. He *assumes* that Turing's argument in the 1936 paper was to show that "mental procedures cannot go beyond mechanical procedures" and considers it as inconclusive, because Turing neglects the dynamic nature of mind. However, simply carrying out a mechanical procedure does not, and indeed should not, involve an expansion of our understanding. Turing viewed the restricted use of mind in computations undoubtedly as static. I leave that misunderstanding out of the systematic considerations in the main text. The appeal to finiteness of states of mind when comparing Gödel's and Turing's perspectives is also pushed into the background as it is not crucial at all for the central issues under discussion: there does not seem to be any disagreement.

I do not fully understand these enigmatic observations, but three points can be made. First, mathematical experience has to be invoked when asking the right questions; second, aspects of that experience may be codified in a mechanical procedure and serve as the basis for asking the right questions; third, the answers may involve abstract terms that are introduced by the non-mechanical mental procedure. We should not dismiss or disregard Gödel's methodological remark that "asking the right questions on the basis of a mechanical procedure" may be part of a systematic method to push forward the development of mind.[19] Even this very limited understanding allows us to see that Gödel's reflections overlap with Turing's proposal for investigating matters in a more empirical and directly computational manner.

Much of Turing's work of the late 1940s and early 1950s explicitly deals with mental processes. But nowhere is it claimed that the latter cannot go beyond mechanical ones. Mechanical processes are still made precise as Turing machine computations; in contrast, machines that might exhibit intelligence have a more complex structure than Turing machines and, most importantly, interact with their environment. Conceptual idealization and empirical adequacy are now being sought for quite different purposes, and one might even say that Turing is actually trying to capture what Gödel described when searching for a broader concept of humanly effective calculability, namely, "… that mind, in its use, is not static, but constantly developing". In his paper *Intelligent Machinery*, Turing states:

> If the untrained infant's mind is to become an intelligent one, it must acquire both discipline and initiative. So far we have been considering only discipline [via the universal machine]. […] But discipline is certainly not enough in itself to produce intelligence. 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 to try and copy it in machines. (Turing, 1948, p. 21)[20]

How, in particular, can we transcend discipline when doing mathematics? Turing provided a hint already in his 1939-paper, where ordinal logics are introduced to expand formal theories in a systematic way; (cf. Feferman, 1988; 2006b) for informative discussions. In that paper, his Ph.D. thesis written under the di-

[19] There seems to be also a connection to remarks in his (1947, pp. 182–183), where Gödel points out that there may be "another way" (apart from judging its intrinsic necessity) to decide the truth of a new axiom. This other way consists in inductively studying its success, "that is, its fruitfulness in consequences and in particular in 'verifiable' consequences, i.e., consequences demonstrable without the new axiom, whose proofs by means of the new axiom, however, are considerably simpler and easier to discover, and make it possible to condense into one proof many different proofs".

[20] In his (1950, p. 459), Turing points out, in a similar spirit: "Intelligent behaviour presumably consists in a departure from the completely disciplined behaviour involved in computation, but a rather slight one, which does not give rise to random behaviour, or to pointless repetitive loops".

rection of Church, Turing distinguishes between ingenuity and intuition. He observes that in formal logics their respective roles take on a greater definiteness. Intuition is used for "setting down formal rules for inferences which are always intuitively valid", whereas ingenuity is to "determine which steps are the more profitable for the purpose of proving a particular proposition". He notes:

> In pre-Gödel times it was thought by some that it would be possible to carry this programme to such a point that all the intuitive judgements of mathematics could be replaced by a finite number of these rules. The necessity for intuition would then be entirely eliminated. (Turing, 1939, p. 209)

That intuition cannot be eliminated, on account of the first incompleteness theorem, is emphasized in Turing's letters to Max Newman from around 1940 that have been reprinted in (Copeland, 2004, pp. 211–216). After all, one can determine the truth of the Gödel sentence, say, for ZF set theory, despite the fact that it is independent of ZF. Providing a general reason for such a determination, Turing writes, "… there is a fairly definite idea of a true formula which is quite different from the idea of a provable one" (p. 215). Eight years later, in his (1948, p. 107), Turing formulated at the very outset reasons given by some for asserting, "it is not possible for machinery to show intelligent behaviour [sic!]". One of the reasons is directly related to the limitative theorems. They are assumed to show that when machines are used for "determining the truth or falsity of mathematical theorems […] then any given machine will in some cases be unable to give an answer at all". This inability of any particular machine is contrasted with human intelligence that "seems to be able to find methods of ever-increasing power for dealing with such problems 'transcending' the methods available to machines" (Turing, 1948, p. 108).

It is thus not surprising that Turing takes in his paper (1950, pp. 444–445) the mathematical objection to his view quite seriously. He considers the objection as based on the limitative results, in particular Gödel's theorems, which are understood by some as proving "a disability of machines to which the human intellect is not subject". Turing gives two responses. The short one states that the objection takes for granted, without any sort of proof, that the human intellect is not subject to the limitations to which machines provably are. However, Turing thinks that the objection cannot be dismissed quite so lightly and proceeds to a second response. It acknowledges the superiority of the human intellect with respect to a single machine (we can recognize the truth of "its" Gödel sentence), but Turing views that as a petty triumph. The reason for this is formulated succinctly as follows: "There would be no question of triumphing simultaneously over all machines. In short, then, there might be men cleverer than any given machine, but then there might be other machines cleverer again, and so on" (Turing, 1950, p. 445).

Turing does not offer a proof of the claim that there is "no question of triumphing simultaneously over all machines". It is precisely here that Gödel's "fact" concerning a humanly effective, but non-mechanical procedure seems to be in conflict with Turing's assertion.[21] If the "fact" were a fact, then it would sustain the objection successfully. Can one go beyond claim and counterclaim? Or, even better, can one use the tension as an inspiration for concrete work that elucidates the situation?

IV. Finding Proofs (With Ingenuity)

Let us return, as a first positive step towards bridging the gap between claim and counterclaim, to Turing's distinction between ingenuity and intuition. Intuition is explicitly linked to the incompleteness of formal theories and provides an entry point to exploiting, through computational work, a certain parallelism between Turing's and Gödel's considerations, when the latter are based on mechanical procedures. Copying the *residue* in machines is the common task at hand. It is a difficult one in the case of mathematical thinking, and Gödel would argue an impossible one, if machines are particular Turing machines. Turing would agree, of course. Before we can start copying, we have to discover partially the nature of the residue; one might hope to begin doing that through proposals for finding proofs in mathematics.

In his lecture to the London Mathematical Society and in *Intelligent Machinery*, Turing calls for heuristically guided intellectual searches and for initiative that includes, in the context of mathematics, proposing new intuitive steps. Such searches and the discovery of novel intuitive steps would be at the center of "research into intelligence of machinery". Let me draw a diagram: the formal theory FT_i has been expanded to the proof theoretically stronger theory FT_{i+1}; the theories are presented via Turing machines M_i and M_{i+1}, respectively.

$$\begin{array}{c}\boxed{FT_{i+1} \text{ is given by Turing machine } M_{i+1}} \\ \uparrow \\ \boxed{FT_{i} \text{ is given by Turing machine } M_{i}}\end{array}$$

[21] "Seems", as Turing pits individual men against particular machines, whereas Gödel pits the "human mind" against machines. This aspect is also briefly discussed in the first letter to Newman in (Copeland, 2004, p. 215): if one moves away from considering a particular machine and allows machines with different sets of proofs, then "by choosing a suitable machine one can approximate 'truth' by 'provability' better than with a less suitable machine, and can in a sense approximate it as well as you please".

The transition from one theory to the next and, correspondingly, from one Turing machine to the next is non-mechanical for Gödel as well as for Turing. In Gödel's case, unfolding the explication of the concept of set by a non-mechanical method is the basis for a humanly effective procedure. Even if Gödel's method would take into account a mechanical procedure of the character described above, in the end, it would present a new and stronger axiom of infinity; it is in this sense that the method could be viewed as uniform. For Turing, it seems, the addition of intuitive steps (outside of his ordinal logics) is principally based on the analysis of machine learning and computer experimentation.[22] It would be closely tied to the particulars of a situation without the connecting thread of Gödel's method and, thus, it would not be uniform. In addition, Turing emphasizes at a number of places that a random element be introduced into the development of machines, thus providing an additional feature that releases them from strict discipline and facilitates a step from M_i to M_{i+1}.

What is striking is that both Gödel and Turing make "completeness claims": at the end of the second paragraph of section III, I quoted Gödel's remark from his 1947-paper that every set theoretic statement is decidable from the current axioms together with "a true assertion about the largeness of the universe of all sets"; in note 20, Turing's remark is quoted that by choosing a suitable machine one can approximate "truth" by "provability" and "in a sense approximate it [truth] as well as you please". That is highly speculative in both cases; slightly less speculatively, Turing conjectured:

> As regards mathematical philosophy, since the machines will be doing more and more mathematics themselves, the centre of gravity of the human interest will be driven further and further into philosophical questions of what can in principle be done etc. (1947, p. 103)

This expectation has not been borne out yet, and Gödel would not be surprised. However, he could have cooperated with Turing on the "philosophical questions of what can in principle be done" and, to begin with, they could have agreed terminologically that there is a human mind whose working is not reducible to the working of any particular brain. They could have explored and, possibly argued about, Turing's contention in his (1951, p. 472) "that machines can be constructed which will simulate the behaviour (sic!) of the human mind very closely". Indeed, Turing had taken a step toward a concept of human mind, when he emphasizes at the end of *Intelligent Machinery*, "the isolated man does not develop any intellectual power", and then argues:

---

[22] Copeland, in his (2006), gives much the same interpretation. He remarks on p. 168: "In his post-war writing on mind and intelligence […] the term "intuition" drops from view and what comes to the fore is the closely related idea of *learning*—in the sense of devising and discovering—new methods of proof".

> It is necessary for him to be immersed in an environment of other men, whose techniques he absorbs during the first twenty years of his life. He may then perhaps do a little research of his own and make a very few discoveries which are passed on to other men. From this point of view the search for new techniques must be regarded as carried out by the human community as a whole, rather than by individuals. (p. 127)

Turing calls this, appropriately enough, a *cultural search* in contrast to the more limited *intellectual searches* possible for individual men or machines. To build machines that think serves also another purpose as Turing explained in a 1951 radio broadcast: "The whole thinking process is still rather mysterious to us, but I believe that the attempt to make a thinking machine will help us greatly in finding out how we think ourselves" (Turing, 1951b, p. 486).

For the study of human thinking mathematics is a marvelous place to start. Where else do we find an equally rich body of rigorously organized knowledge that is structured for both intelligibility and discovery? Turing, as we saw above, had high expectations for machines' progress in doing mathematics; but it is still extremely difficult for them to "mathematize" on their own. Newman, in a radio debate with Braithwaite, Jefferson, and Turing, put the general problem very well:

> Even if we stick to the reasoning side of thinking, it is a long way from solving chess problems to the invention of new mathematical concepts or making a generalisation (sic!) that takes in ideas that were current before, but had never been brought together as instances of a single general notion. (Turing, 1952, p. 498)

The important question is whether we can gain, by closely studying *mathematical practice*, a deeper understanding of fundamental concepts, techniques and methods of mathematics and, in that way, advance our understanding of the capacities of the mathematical mind as well as of basic operations of the mind. This question motivates a more modest goal, namely, formulating strategies for an automated search: not for proofs of new results, but for proofs that reflect logical and mathematical understanding; proofs that reveal their intelligibility and that force us to make explicit the *ingenuity* required for a successful search.[23] The logical framework for such studies must include a *structural*

[23] This involves undoubtedly reactions to Turing's remarks and impatient questions in a letter to Newman: "In proofs there is actually an enormous amount of sheer slogging, a certain amount of ingenuity, while in most cases the actual 'methods of proof' are quite well known. Cannot we make it clearer where the slogging comes in, where there is ingenuity involved, and what are the methods of proof"? (Copeland, 2004, p. 213). Abramson, in his (2008), emphasizes insightfully the significance of Lady Lovelace's objection. In the context here, his emphasis pointed out to me that Turing (1950, p. 451), views "the mere working out of consequences from data and general principles" as a "virtue" and as a "source for surprises". Turing articulates that important perspective after having called "false" the assumption that "as soon as a fact is presented to a mind all consequences of the fact spring into the mind simultaneously with it".

theory of proofs that extends proof theory through (i) articulating structural features of derivations and (ii) exploiting the meaning of abstract concepts; both aspects are crucial for finding humanly intelligible proofs.[24] We will hopefully find out what kind of broad strategies and heuristic ideas will emerge, what is the necessary ingenuity. In this way, we will begin to uncover part of Turing's residue and part of what Gödel considered as humanly effective, but not mechanical, in each case "by asking the right questions on the basis of a mechanical procedure" (Gödel, 1974, p. 325).

The very last remark in (Turing, 1954) comes back, in a certain sense, to the mathematical objection. Turing views the limitative results as being "mainly of a negative character, setting bounds to what we can hope to achieve purely by reasoning". Characterizing in a new way the residue that has to be discovered and implemented to construct intelligent machinery, Turing continues, "These, and some other results of mathematical logic may be regarded as going some way towards a demonstration, within mathematics itself, of the inadequacy of 'reason' unsupported by common sense". This is as close as Turing could come to agree with Gödel's dictum "The human mind infinitely surpasses any finite machine", if "finite machine" is identified with "Turing machine".

Acknowledgments

Versions of this essay were read at the *Computability in Europe* conference in 2006, as the Gödel Lecture at the *Colloquium Logicum 2006*, and at the Logic Seminar of the Mathematics Department, University of Lisbon. It is closely based on previous papers of mine (Sieg, 2006; 2007), but I argue here for a parallelism between Gödel's and Turing's considerations for transcending purely mechanical processes in mathematics. I revised and completed the manuscript in March 2009, while I was a Fellow at the Swedish Collegium for Advanced Study in Uppsala, but made some stylistic changes two years later. Finally, remarks by Jack Copeland and Oron Shagrir stimulated a sharpening and expansion of points in the last section; that version served as the basis for my talk at the Studia Logica Conference "Church's Thesis: Logic, Mind and Nature" in Krakow (June 3–5, 2011).

[24] I have been pursuing a form of such a structural proof theory for quite a number of years. Central considerations and results are presented in (Sieg, 2010); there I also pointed out connections with Greek mathematics and the radical transformation of mathematics in the nineteenth century, as described in (Stein, 1988). A fully automated proof search method for (classical) first-order logic has been implemented in the AProS system. The overall project, addressing strategic search and dynamic tutoring, is being extended now also to elementary set theory; it is described at http://www.phil.cmu.edu/projects/apros/, and AProS is downloadable from that site.

REFERENCES

Abramson, D. (2008). Turing's Responses to Two Objections. *Minds & Machines*, *18*(2), 147–167.

Ackermann, W. (1924). Begründung des 'tertium non datur' mittels der Hilbertschen Theorie der Widerspruchsfreiheit. *Mathematische Annalen*, *93*, 1–36.

Church, A. (1935). An Unsolvable Problem of Elementary Number Theory. *Bulletin of the AMS*, *41*, 332–333.

Copeland, J. (2004). *The Essential Turing: The Ideas That Gave Birth to the Computer Age*. Oxford: Clarendon Press.

Copeland, J. (2006). Turing's Thesis. In: A. Olszewski, J. Wolenski, R. Janusz (Eds.), *Church's Thesis After 70 Years* (pp. 147–174). Frankfurt: Ontos Verlag.

Davis, M. (1982). Why Gödel Did Not Have Church's Thesis. *Information and Control*, *54*, 3–24.

Dawson, John W. (2006). *Gödel and the Origin of Computer Science*. Lecture at CiE 2006, Swansea.

Dedekind, R. (1888). *Was sind und was sollen die Zahlen?* Braunschweig: Vieweg.

Feferman, S. (1988). Turing in the Land of O(z). In: R. Herken (Ed.), *The Universal Turing Machine—A Half-Century Survey* (pp. 113–147). Oxford: Oxford University Press.

Feferman, S. (2006a). Are There Absolutely Unsolvable Problems? Gödel's Dichotomy. *Philosophia Mathematica*, *14*(2), 134–152.

Feferman, S. (2006b). Turing's Thesis. *Notices of the AMS*, *53*(10), 1200–1205.

Gandy, R. (1980). Church's Thesis and Principles for Mechanisms. In: J. Barwise, H. J. Keisler, K. Kunen (Eds.), *The Kleene Symposium* (pp. 123–148), Amsterdam: North-Holland Publishing Company.

Gödel, K. (1931). Über formal unentscheidbare Sätze der *Principia Mathematica* und verwandter Systeme I [On Formally Undecidable Propositions of Principia Mathematica and Related Systems I]. *Monatshefte für Mathematik und Physik*, *38*, 173–198.

Gödel, K. (1933). The Present Situation in the Foundations of Mathematics. In: K. Gödel, *Collected Works III* (pp. 45–53). Oxford: Oxford University Press.

Gödel, K. (1934). On Undecidable Propositions of Formal Mathematical Systems. In: K. Gödel, *Collected Works I* (pp. 346–371). Oxford: Oxford University Press.

Gödel, K. (1936). Über die Länge von Beweisen. *Ergebnisse eines mathematischen Kolloquiums*, Heft 7, 23–24.

Gödel, K. (1939). Finding Aid (Notre Dame Lecture on Logic, Spring 1939). In: K. Gödel, *Collected Works IV–V* (pp. 527–528). Oxford: Oxford University Press.

Gödel, K. (193?). Undecidable Diophantine Propositions. In: K. Gödel, *Collected Works III* (pp. 164–175). Oxford: Oxford University Press.

Gödel, K. (1946). Remarks Before the Princeton Bicentennial Conference on Problems in Mathematics. In: K. Gödel, *Collected Works II* (pp. 150–153). Oxford: Oxford University Press.

Gödel, K. (1947). What is Cantor's Continuum Problem? In: K. Gödel, *Collected Works II* (pp. 167–187). Oxford: Oxford University Press.

Gödel, K. (1951). Some Basic Theorems on the Foundations of Mathematics and Their Implications. In: K. Gödel, *Collected Works III* (pp. 304–323). Oxford: Oxford University Press.

Gödel, K. (1964). Postscriptum for [1934]. In: K. Gödel, *Collected Works I* (pp. 369–371). Oxford: Oxford University Press.

Gödel, K. (1972). Some Remarks on the Undecidability Results. In: K. Gödel, *Collected Works II* (pp. 305–306). Oxford: Oxford University Press.

Gödel, K. (1974). Note. In: H. Wang, *From Mathematics to Philosophy* (pp. 325–326). London: Routledge & Kegan Paul.

Gödel, K. (1986). *Collected Works I*. Oxford: Oxford University Press.

Gödel, K. (1990). *Collected Works II*. Oxford: Oxford University Press.

Gödel, K. (1995). *Collected Works III*. Oxford: Oxford University Press.

Gödel, K. (2003). *Collected Works IV–V*. Oxford: Oxford University Press.

Herbrand, J. (1931). On the Consistency of Arithmetic. In: W. Goldfarb (Ed.), *Jacques Herbrand, Logical Writings* (pp. 282–298). Cambridge, Mass.: Harvard University Press.

Herken, R. (Ed.). (1988). *The Universal Turing Machine—A Half-Century Survey*. Oxford: Oxford University Press.

Hilbert, D. (1900). Über den Zahlbegriff. *Jahresbericht der Deutschen Mathematiker Vereinigung*, *8*, 180–194.

Hilbert, D. (1905). Über die Grundlagen der Logik und der Arithmetik. In: A. Krazer (Ed.), *Verhandlungen des 3. Internationalen Mathematiker-Kongresses: in Heidelberg vom 8. bis 13* (pp. 174–185). Leipzig: Teubner.

Hilbert, D. (1923). Die logischen Grundlagen der Mathematik. *Mathematische Annalen*, *88*, 151–165.

Hilbert, D., Bernays, P. (1921/2). Grundlagen der Mathematik. in: W. Ewald, W. Sieg (Eds.), *David Hilbert's Lectures on the Foundations of Arithmetic and Logic 1917–1933* (pp. 431–521). Springer.

Hilbert, D., Bernays, P. (1934). *Grundlagen der Mathematik I*. Berlin: Springer Verlag.

Hilbert, D., Bernays, P. (1939). *Grundlagen der Mathematik II*. Berlin: Springer Verlag.

Kennedy, J., van Atten, M. (2004). Gödel's Modernism: On Set-Theoretic Incompleteness. *Graduate Faculty Philosophy Journal*, *25*(2), 289–349.

Kennedy, J., van Atten, M. (2009). "Gödel's Modernism: On Set-Theoretic Incompleteness", Revisited. In: S. Lindström et al. (Eds.), *Logicism, Intuitionism, and Formalism—What Has Become of Them?* (Synthese Library 341, pp. 303–355). Springer.

Kolmogorov, A., Uspenski, V. (1963). On the Definition of an Algorithm. *AMS Translations*, *21*(2), 217–245.

Martin, D. A. (2005). Gödel's Conceptual Realism. *Bulletin of Symbolic Logic*, *11*(2), 207–224.

Post, E. (1947). Recursive Unsolvability of a Problem of Thue. *J. Symbolic Logic*, *12*, 1–11.

Sieg, W. (1994). Mechanical Procedures and Mathematical Experience. In: A. George (Ed.), *Mathematics and Mind* (pp. 71–117). Oxford: Oxford University Press.

Sieg, W. (1997). Step by Recursive Step: Church's Analysis of Effective Calculability. *Bulletin of Symbolic Logic*, *3*(2), 154–180.

Sieg, W. (1999). Hilbert's Programs: 1917–1922. *Bulletin of Symbolic Logic*, *11*(2), 1–44.

Sieg, W. (2002). Calculations by Man and Machine: Conceptual Analysis. In: W. Sieg, R. Sommer, C. Talcott (Eds.), *Reflections on the Foundations of Mathematics* (Lecture Notes in Logic 15, pp. 390–409). Cambridge University Press.

Sieg, W. (2006). Gödel on Computability. *Philosophia Mathematica*, *14*(2), 189–207.

Sieg, W. (2007). On Mind & Turing's Machines. *Natural Computing*, *6*, 187–205.

Sieg, W. (2009a). On Computability. In: A. D. Irvine (Ed.), *Philosophy of Mathematics* (pp. 535–630). Amsterdam: North Holland.

Sieg, W. (2009b). Hilbert's Proof Theory. In: D. M. Gabbay, J. Woods (Eds.), *Handbook of the History of Logic. Volume 5: Logic from Russell to Church* (pp. 321–384). Amsterdam: North Holland.

Sieg, W. (2010). Searching for Proofs (And Uncovering Capacities of the Mathematical Mind). In: S. Feferman, W. Sieg (Eds.), *Proofs, Categories and Computations* (pp. 189–215). London: College Publications.

Sieg, W., Byrnes, J. (1996). K-Graph Machines: Generalizing Turing's Machines and Arguments. In: P. Hájek (Ed.), *Godel '96: Logical Foundations of Mathematics, Computer Science and Physics* (Lecture Notes in Logic 6, pp. 98–119). Berlin: Springer Verlag.

Sieg, W., and D. Schlimm (2005). Dedekind's Analysis of Number: Systems and Axioms. *Synthese*, *147*, 121–170.

Stein, H. (1988). Logos, Logic, and Logistiké: Some Philosophical Remarks on Nineteenth-Century Transformation of Mathematics. In: W. Aspray, P. Kitcher (Eds.), *History and Philosophy of Modern Mathematics* (vol. XI of Minnesota Studies in the Philosophy of Science, pp. 238–259). Minneapolis: University of Minnesota Press.

Turing, A. M. (1936). On Computable Numbers, With an Application to the Entscheidungsproblem. *Proc. London Math. Soc.* (series 2), *42*, 230–265.

Turing, A. M. (1939). Systems of Logic Based on Ordinals. *Proc. London Math. Soc.* (series 2), *45*, 161–228.

Turing, A. M. (1947). Lecture to the London Mathematical Society on 20 February 1947. In: D. C. Ince (Ed.), *Collected Works of A.M. Turing—Mechanical Intelligence* (87–105). Amsterdam: North Holland.
Turing, A. M. (1948). Intelligent Machinery. In: D. C. Ince (Ed.), *Collected Works of A.M. Turing—Mechanical Intelligence* (107–127). Amsterdam: North Holland.
Turing, A. M. (1950). Computing Machinery and Intelligence. *Mind*, *59*, 433–460.
Turing, A. M. (1951a). Intelligent Machinery, A Heretical Theory [Radio broadcast]. Printed in J. Copeland, *The Essential Turing: The Ideas That Gave Birth to the Computer Age* (pp. 472–475). Oxford: Clarendon Press.
Turing, A. M. (1951b). Can Digital Computers Think? [Radio broadcast]. Printed in J. Copeland, *The Essential Turing: The Ideas That Gave Birth to the Computer Age* (pp. 482–486). Oxford: Clarendon Press.
Turing, A. M. (1952). Can Automatic Calculating Machines Be Said to Think? [Radio discussion of A. Turing, R. Braithwaite, G. Jefferson, and M. Newman. Printed in J. Copeland, *The Essential Turing: The Ideas That Gave Birth to the Computer Age* (pp. 494–506). Oxford: Clarendon Press.
Turing, A. M. (1954). Solvable and Unsolvable Problems. *Science News*, *31*, 7–23.
van Heijenoort, J. (1968). *From Frege to Gödel—A Source Book in Mathematical Logic, 1879–1931*. Cambridge, Mass.: Harvard University Press.
Wang, H. (1974). *From Mathematics to Philosophy*. London: Routledge & Kegan Paul.

NEW REFERENCES (FOR THE POSTSCIPTUM)

Bernays, P. (1930). Die Philosophie der Mathematik und die Hilbertsche Beweistheorie. In: P. Bernays, *Abhandlungen zur Philosophie der Mathematik* (pp. 17–61). Darmstadt: Wissenschaftliche Buchgesellschaft.
Davis, M., Sieg, W. (2015). Conceptual Confluence in 1936: Post and Turing. In: G. Sommaruga, T. Strahm (Eds.), *Turing's Revolution* (pp. 3–27). Berlin: Springer.
Gentzen, G. (1936). Die Widerspruchsfreiheit der reinen Zahlentheorie. *Mathematische Annalen*, *112*(1), 493–565.
Hilbert, D. (1918). Axiomatisches Denken. *Mathematische Annalen*, *78*, 405–415.
Hilbert, D., Ackermann, W. (1928). *Grundzüge der theoretischen Logik*. Berlin: Springer.
Hilbert, D., Bernays, P. (1917–18). Prinzipien der Mathematik. In: W. Ewald, W. Sieg (Eds.), *David Hilbert's Lectures on the Foundations of Arithmetic and Logic 1917–1933* (pp. 59–221). Berlin: Springer.
Sieg, W. (2013). Hilbert's Programs and Beyond. Oxford: Oxford University Press.
Sieg, W. (2014). The Ways of Hilbert's Axiomatics: Structural and Formal. *Perspectives on Science*, *22*(1), 133–157.

Sieg, W. (2018). What is the *Concept* of Computation? In: F. Manea, R.G. Miller, D. Nowotka (Eds.), *Sailing Routes in the World of Computation*, *Proceedings of CiE 2018* (pp. 386–396). Springer LNCS 10936.

Sieg, W., Morris, R. (2018). Dedekind's Structuralism: Creating Concepts and Deriving Theorems. In: E. H. Reck (Ed.), *Logic, Philosophy of Mathematics, and Their History: Essays in Honor of W.W. Tait* (pp. 251–301). London: College Publications.

Sieg, W., Walsh, P. (2019). Natural Formalization: Deriving the Cantor-Bernstein Theorem in ZF. *The Review of Symbolic Logic*. doi:10.1017/S175502031900056X

Sieg, W., Derakhshan, F. (2020). Human-Centered Automated Proof Search. Manuscript submitted for publication.

Sieg, W., Szabo, M., McLaughlin, D. (2016). Why Post Did [Not] Have Turing's Thesis. In: E. Omodeo, A. Policriti (Eds.), *Martin Davis on Computability, Computational Logic, and Mathematical Foundations* (pp. 175–208). Berlin: Springer.

Postscriptum

This essay was originally published in the volume *Computability—Turing, Gödel, Church, and Beyond*, MIT Press 2013. It is reprinted here with the permission of MIT Press. The current version is not literally the same essay, as I made a few minor stylistic changes. Three developments in my own thinking, since the completion of the essay in 2011, are worthwhile to point out and to describe briefly in this Postscriptum. The first provides a stronger connection to the past, the second is a further deepening of the analysis of the concept of computability, and the third yields a systematic connection to the future from the perspective of 2011.

There is then, first of all, a deeper historical understanding of the methodological basis for the investigations of Gödel and Turing. The crucial building blocks for that basis were provided by the radically new structuralist conception of mathematics in the work of Dedekind and Hilbert and the dramatically expanded reach of logic primarily through Frege's efforts; (Sieg & Morris, 2018). The mathematical work and the logical work were hardly connected when they were created during the last thirty years of the nineteenth century. After Whitehead and Russell had reshaped logic through *Principia Mathematica*, the two building blocks were joined and received a rigorous mathematical description in (Hilbert & Bernays, 1917–1918). These lectures are the beginning of modern mathematical logic and opened the door for metamathematical investigations in the 1920s; they are also, via (Hilbert & Ackermann, 1928), the backdrop for Gödel and Turing. The emergence of metamathematics took place during the first thirty years of the twentieth century; it is incisively described in (Bernays, 1930). Many people have contributed to a deeper historical understanding that is reflected in the first half of my book *Hilbert's Programs and Beyond*. The shift from structural to formal axiomatics, absolutely central for Gödel and Turing, is elucidated in (Sieg, 2014). Book and paper contain, of course, references to the rich literature.

The second development is a sharpening of my structural axiomatic approach in order to characterize computability as an abstract mathematical concept. That is alluded to in this essay at the end of Section II. It has an historical component that brings out the significance of Post's work (Sieg, Szabo & McLaughlin, 2016); it also uncovers the deep conceptual confluence of Post's and Turing's work in 1936, presented in (Davis & Sieg, 2015). Finally, in a paper that was dedicated to Davis' ninetieth birthday (Sieg, 2018), I raised and sought to answer the key methodological question, "What is the concept of computation?" Drawing on my earlier work, the concise answer is given in terms of computable dynamical systems. This is done against the background of two classes of mathematical results generalizing the considerations of Section I (Gödel's Absoluteness) and of Section II (Turing's Reducibility). The set theoretic formulation of the abstract concept "computable dynamical system" is waiting for an illuminating category theoretic characterization.

We finally come to the third development since 2011. It concerns neither the historical background for Sections I and II nor the axiomatic sharpening of the concept of computation. It is rather connected to the comparative analysis of Gödel's and Turing's suggestions for transcending mechanical procedures in Sections III and IV. The goals of that development are described in broad strokes in the penultimate paragraph of the essay and have been pursued within my AProS Project that is mentioned in Note 23. The latter seeks to find strategies for the automated search for humanly intelligible proofs in constructive and classical logic, but also in meta-mathematics (Gödel's incompleteness theorems) and set theory (the Cantor-Bernstein Theorem). My views on "natural formalization within a foundational frame" and "human-centered automated proof search" are at the center of and operative in (Sieg & Walsh, 2019), respectively (Sieg & Derakhshan, 2020).

The relevant theoretical perspective is this: formalizing mathematical practice is central for the significance of proof theoretic investigations, be they concerned with the consistency problem of formal theories or with the "mining" of particular proofs. We use refined, conceptually organized formal frameworks to reflect deep structures of mathematical proofs. Thus, we aim for a theory of proofs in which "ordinary" proofs are treated as objects of investigation. That is in the spirit of the pioneers. Hilbert remarked in (1918), "[w]e must—that is my conviction—take the concept of the specifically mathematical proof as an object of investigation". In just this spirit, Gentzen thought in his (1936, p. 499) that one can obtain only through formalization a "rigorous treatment of proofs" and emphasized then most strongly, "[t]he objects of proof theory shall be the proofs carried out in mathematics proper".