Title	Gödel on Turing's analysis of computability
Sub Title	
Author	小川, 芳範(Ogawa, Yoshinori)
Publisher	Centre for Advanced Research on Logic and Sensibility The Global Centers of Excellence Program, Keio University
Publication year	2008
Jtitle	CARLS series of advanced study of logic and sensibility Vol.1, (2007.) ,p.381- 402
JaLC DOI	
Abstract	Gödel made two different and conflicting remarks on Turing's discussion of computability: in one he applauds Turing's definition of computability in terms of a Turing machine whereas in the other he criticizes Turing's argument for the definition as containing a "philosophical error" and thus "inconclusive." A main objective of my essay is to make sense of Gödel's apparently inconsistent responses to Turing's discussion and thereby elucidate what continually motivated his thinking throughout his long research career.
Notes	Part 5 : Logic and Informatics
Genre	Research Paper
URL	https://koara.lib.keio.ac.jp/xoonips/modules/xoonips/detail.php?koara_id=KO12002001-20080331-0381

23 | Gödel on Turing's Analysis of Computability

Yoshinori Ogawa[1]

[1] Centre for Advanced Research on Logic and Sensibility (CARLS), Keio University

Gödel made two different and conflicting remarks on Turing's discussion of computability: in one he applauds Turing's definition of computability in terms of a Turing machine whereas in the other he criticizes Turing's argument for the definition as containing a "philosophical error" and thus "inconclusive." A main objective of my essay is to make sense of Gödel's apparently inconsistent responses to Turing's discussion and thereby elucidate what continually motivated his thinking throughout his long research career.

1. Two Problems Concerning the Incompleteness Theorems

As is well known, in the early 1920s, Hilbert proposed a foundational program in an attempt to protect classical mathematics from the double-threat of set theoretical paradoxes and the revisionist advocacy of Brouwer and Weyl. In a nutshell, it consists in providing a consistency proof for formalized mathematics using only its finitary fragment. The idea is that the problematic, non-finitary parts of classical mathematics are nothing other than particular instances of "ideal elements," which are introduced into a system for the purpose of simplifying or generalizing our thinking about a certain subject matter, and therefore that there is no need to be concerned about the epistemological status of such formal devices, insofar as their employment never leads to incorrect finitary results, that is to say, insofar as the non-finitary

system is consistent with finitary mathematics. In this way, it is said, Hilbert's program attempts to *justify* our reference in infinitary mathematics to transcendent objects and is an epistemological (foundationalist) project.

Without contradicting such an epistemological reading of Hilbert's program, however, it may be argued that the consistency program also forms a part of Hilbert's grand plan of the "mechanization" of mathematics, which is the project of obtaining an effective procedure to solve mathematical and metamathematical problems in a uniform manner; for what is actually carried out in the Hilbert program is to replace the discourse of each branch of mathematics with a rigorously specified deductive system and to study these deductive systems syntactically by means of finitary methods. Indeed, Hilbert's foundational enterprise included among its most urgent tasks, in addition to the finitary consistency problem, the so-called *Entscheidungsproblem* of first order predicate logic and the problem of the completeness of mathematics, both of which were to be tackled by applying the same kind of systematic, "calculatory treatment (*rechnerische Behandlung*)" to syntactic items. As is all too well known, however, in 1931, precisely when Hilbert's project began to look promising, Gödel's paper containing two incompleteness theorems appeared. Very crudely, its main results are as follows:

1. The formal system of mathematics with a certain expressive power cannot be both consistent and complete. (The first incompleteness theorem)
2. The formal system of mathematics with a certain expressive power cannot show its consistency if it is, in fact, consistent. (The second incompleteness theorem)

In short, formal mathematics, if consistent, is incomplete, and its consistency cannot be established in it. Thus, Gödel's theorems are often said to have demonstrated, once and for all, the infeasibility of the Hilbert program insofar as it is characterized by the search for a finitary consistency proof.

The story, however, is not that simple, and there are two things to be recognized here: one is concerned with the formalizability of *meta*mathematics, and the other with the *generalization* of the first incompleteness theorem. An important element of Gödel's proof consists in showing how a formal system can "talk about" its own language and deductive system. More specifically,

he managed to construct, *within* a formal system of arithmetic P,[1] a metamathematical sentence G that "says" that G is not provable in P. Once it is shown that there is such a self-referential sentence in P, it follows that G is not provable in P if P is, in fact, consistent. Now, what is to be seen is that this, i.e., the assertion "If P is consistent, then G is not provable in P," has an analogue in P via the Gödel numbering, which assigns a number to each character of the formal language. The antecedent "P is consistent" can be expressed in P by a formula (call it "ConP" for short) asserting that "For every x, x is not the Gödel number of a proof of a contradiction," and the consequent "G is not provable in P," as we saw above, is equivalent to G. Thus, the metamathematical statement "If P is consistent, then G is not provable in P" corresponds in P to "ConP → G." And this latter is provable in P, as the entire proof for the first incompleteness theorem has an analogue in P via the Gödel numbering. Let us now assume that P's consistency has been proved in P. Since this means that "ConP" is provable in P, it follows from "ConP → G" and "ConP" with a use of modus ponens that G is provable in P. But this would result in a contradiction since, as has already been shown, G is not provable if P is consistent. Hence, if P is, in fact, consistent, then the assumption that P's consistency is provable in P is incorrect, and consequently the consistency of this deductive system cannot be established in *that* system, much less in its finitary fragment. Gödel's second incompleteness theorem, then, implies that P's consistency cannot be established insofar as (what is put forward as) a consistency proof for P is expressible in P. Accordingly, the feasibility of Hilbert's project of finding a finitary consistency proof depends upon the formalizability in P of the Hilbertian finitary methods.

Gödel was well aware of this circumstance and thus wrote towards the end of his 1931 paper that the second incompleteness theorem does not contradict Hilbert's "formalist viewpoint:"

> For this viewpoint presupposes only the existence of a consistency proof
> in which nothing but finitary means of proof is used, and it is conceivable

---

[1]   P is the version of the system of Russell and Whitehead's *Principia Mathematica* in Gödel's 1931 paper.

that there exist finitary proofs that cannot be expressed in the formalism
of P (or of M and A). (Gödel 1931, p. 195)

Very interestingly, von Neumann and Herbrand, who were supposed to be
in Hilbert's camp, were the first to criticize Gödel's reservation about the extent
of his second theorem's implications for Hilbert's consistency program, and
Gödel responded to their objection in letters.[2] It is not clear how these
interactions affected Gödel's thinking but, in any case, by about late 1933,
Gödel changed his views significantly. In his lecture to the Mathematical
Association of America in that year, he presented a formal system A, which
"is based exclusively on the method of complete induction in its definitions
as well as in its proofs," and stated that "all the attempts for a proof for freedom
from contradiction undertaken by Hilbert and his disciples" can be carried
out in A.[3] Gödel went on to say that all the finitary proofs complying with
the requirements of A can easily be expressed in a formal system of arithmetic
and thereby, in effect, declared Hilbert's program bankrupt.

The second and related problem is concerned with the generalization of
the first incompleteness theorem, and Gödel's efforts to find a satisfactory
solution to this problem led him to the problem of *computability*. As its title
indicates, the results of Gödel's 1931 paper are restricted in the sense that the

---

[2] Gödel's thinking around this time is expressed clearly in his letters of Nov. 1930
and Jan. 1931 to Herbrand:

> Clearly, I do not claim either that it is certain that some finitist proofs are not
> formalizable in *Principia Mathematica*, even though intuitively I tend toward
> this assumption. In any case, a finitist proof not formalizable in *Principia
> Mathematica* would have to be quite extraordinarily complicated, and on this
> purely practical ground there is very little prospect of finding one; but that, in
> my opinion, does not alter anything about the possibility in principle. (quoted in
> Sieg 2005, p.179)

[3] As Sieg observes in his 2005, "the restrictive characteristics of the system A . . .
include the requirement that notions have to be decidable and functions must be
calculable." And when Gödel claims that such notions and functions can always be
defined by complete induction, "definition by complete induction" is to be understood
as definition by recursion, which is by no means restricted to primitive recursion.
(Sieg 2005, fn. 7, pp. 179-180)

existence of a formally undecidable formula asserted in it seems applicable only to Russell and Whitehead's and related systems.[4] That is, strictly speaking, Gödel 1931 maintains only that *an* axiomatic deductive system of arithmetic is (syntactically) incomplete. An obvious question arises, then, whether or not the theorem applies to formal systems other than P. Toward the end of section 2, right after the proof for the first incompleteness theorem is given, Gödel notes that his proof appeals to no properties peculiar to the Russell-Whitehead system. More precisely, the properties used in it are said to be the following two, and those only:

1. The class of axioms and the rules of inference (i.e. the relation "immediate consequence") are recursively definable via the Gödel numbering.
2. Every (primitive) recursive function is numeralwise expressible within P.

Gödel's position, then, is that the first incompleteness theorem is applicable not only to P but also to any axiomatic system that has these two properties. For our purpose there are two things to be recognized here. On the one hand, the second condition, i.e., the numeralwise expressibility of (primitive) recursive function within the relevant system, implies that the theorem applies to a system which has expressive power equal to (or greater than) that of elementary arithmetic. On the other hand, concerning the first condition, i.e., the recursive definability of the syntax of the relevant system, Gödel writes that it "is fulfilled in general by every system whose rules of inference are the usual ones and whose axioms (as in P) result from substitution in finitely many schemata," and, for these reasons, concludes that the theorem "holds for a very wide class of formal systems," including all those which arise from the PM system and the Zermelo-Fraenkel set theory by addition of finitely many axioms. The question is whether the applicability of the first incompleteness theorem is confined to P (and its related systems) or can be extended to a formal system

---

[4]   The title of the paper in the German original reads "Über formal unentscheidbare Sätze der Principia Mathematica und verwandter Systeme I [On Formally Undecidable Propositions of Principia Mathematica and Related Systems I]"

in general.[5] Gödel's task, then, would be to show, by analyzing the concept of formal system, that every formal system satisfies the first condition, that is, that the syntax of a formal system is definable, through the Gödel numbering, by primitive recursive functions (or some other relations expressible in elementary arithmetic).[6]

2. Formal Systems and Computability

Now the very title of Gödel's 1934 Princeton lectures, "On Undecidable Propositions of Formal Mathematical Systems" indicates that his efforts were turned precisely to this direction. The printed version of Gödel's lectures begins with the sentence "A formal mathematical system is a system of symbols together with rules for employing them," and goes on to state that the syntax of a formal system must be "constructive." That is, according to Gödel, it is required of a formal system that there be a "finite procedure" for deciding whether a given sequence of symbols is a well-formed formula and one for deciding whether a given well-formed formula is an immediate consequence (by a rule of inference) of a given sequence of wffs and so on. In a lecture he gave in 1933, Gödel emphasizes as "the outstanding feature of the rules of inference" that "they are purely formal, i.e., 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*."[7] Gödel's point is that it is essential for a formal system to be finite and mechanical,[8] and thus it appears the concept of formal system, whose exact content was left

---

[5]   More precisely, any ω-consistent system containing elementary arithmetic.
[6]   Sieg reports in his 1997 that Gödel and Church had a brief exchange on the issue of the theorem's generality already in 1932. Bernays too, according to Sieg, expressed the same concern in his letter to Church in 1934 that Gödel's theorems might not be applicable to Church's systems because some very special features of the *Principia Mathematica* seemed to be needed in Gödel's proof. See Sieg 1997, p. 161 and especially fn. 13 on that page.
[7]   Gödel 1933, p. 45. My italics.
[8]   See also Gödel's remarks in the postscriptum added to his (1931) in 1963 that "a formal system can simply be defined to be any mechanical procedure for producing formulas, called provable formulas" and that it is essential for the concept of formal system that "reasoning is completely replaced by mechanical operations on formulas." (Gödel 1986-95, vol.1, p. 370)

unspecified in the 1931 paper, was now given a clearer characterization by Gödel in terms of "computability by a finite, mechanical procedure."

Yet, all these notions employed in the characterization are, in fact, no clearer than that of formal system in their being *informal* notions, and accordingly cannot provide an appropriate basis for Gödel's goal of generalizing the first incompleteness theorem. Rather, what is required here is a rigorous characterization of the concept of computability itself. Needless to say, this is given by the so-called Church-Turing thesis, which identifies the informal concept of computability or effective calculability with a rigorous mathematical notion like recursiveness or Turing-machine computability. Indeed, as is often mentioned, the printed version of Gödel's Princeton lectures contains an assertion that looks very much like the Church-Turing thesis. In section 2 he offers a formulation equivalent to the "easy" half of the Church-Turing thesis by saying that (primitive) recursive functions "can be computed by a finite procedure" and adds in a footnote that "the converse [i.e., computability $\rightarrow$ recursiveness] seems to be true if, besides [primitive] recursions . . . recursions of other forms (e.g., with respect to two variables simultaneously) are admitted."[9] This, however, cannot be taken to mean that Gödel actually anticipated the Church-Turing thesis. For one thing, Gödel himself later explicitly denied such an interpretation,[10] and, for another, it has been reported that at right about this time Gödel dismissed Church's proposal of defining

---

[9] The footnote continues with the following words:

> This cannot be proved, since the notion of finite computation is not defined, but it serves as a heuristic principle (Gödel 1934, p. 348).

[10] In a letter of February 15, 1965 to Martin Davis, Gödel writes:

> It is *not true* that footnote 3 is a statement of Church's Thesis. The conjecture stated there only refers to the equivalence of 'finite (computation) procedure' and 'recursive procedure.' However, I was, at the time of these lectures, not at all convinced that my concept of recursion comprises all possible recursions. (Davis 1982, p. 8)

Incidentally, in the last section of Gödel 1934 titled "General recursive functions" Gödel attempts, by introducing what is today known as Herbrand-Gödel computability, to capture the notion of recursiveness and thus point to a direction toward the rigorization of the concept of computability.

"effective calculability" by "lambda-definability" as "thoroughly unsatisfactory."[11] It would seem, then, that Gödel was still in a search for a rigorous definition of "formal system" and hence had yet to achieve the desired goal of generalizing the first incompleteness theorem as late as in 1934.

But it did not take long before Gödel changed his view. In a draft of a lecture, written probably in early 1938, he says this:

> When I first published my paper about undecidable propositions the result could not be pronounced in this generality, because for the notions of mechanical procedure and formal system no mathematically satisfactory definition had been given at that time. This gap has since been filled by Herbrand, Church and Turing (Gödel 193?, p. 166).

Here Gödel mentions three people, but for him it was Turing's 1936 paper "On Computable Numbers, with an Application to the *Entscheidungsproblem*" that filled the "gap." Immediately after the above quoted passage, Gödel states that "the essential point is to define what a procedure is" since formal system is a mechanical procedure for producing provable formulas. Now, the procedures in question, which operate on syntactic items are, via the Gödel coding, reducible to procedures operating on integers, continues Gödel, so the concept to be specified is the concept of "computable function." He then provides mathematical definitions of computable functions as those functions whose values can be computed in an equational calculus and writes, "that this really is the correct definition of mechanical computability was established beyond any doubt by Turing" (ibid, p.168). The question, of course, is why Gödel thinks this is so. His answer to this question is as follows:

> He [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 slip of paper and put the slip of paper into the machine and turn the crank, then after a finite number

---

[11] See Sieg 1997, pp. 159-160, Davis 1982, pp. 12-13.

of turns the machine will stop and the values of the function for the argument $n_1, \ldots n_r$ will be printed on the paper. (Gödel 193?)

Gödel's implicit claim is Turing has shown that all mechanical procedures can be carried out by "a machine with a finite number of parts," which is really simply to reiterate Turing's thesis.

Although put in a different methodological perspective, very much the same story is told by Gödel at the outset of his Gibbs Lecture in 1951. There he considers the implications of some basic theorems on the foundations of mathematics, including the incompleteness theorems, whose "greatest improvement was made possible through the precise definition of the concept of finite procedure, which plays a decisive role in these results." He acknowledges the existence of several different ways of arriving at such a definition and points out that nevertheless they all lead to exactly the same concept. Then comes a brief remark on Turing:

> The most satisfactory way, in my opinion, is that of reducing the concept of finite procedure to that of machine with a finite number of parts, as has been done by the British mathematician Turing. (Gödel 1951, pp. 304-305)

But once again, this is only to reiterate the result of Turing's analysis, and Gödel's remark on Turing ends there. What, then, is Gödel' reason for thinking Turing's definition to be "the most satisfactory" one? What does Gödel think makes Turing's reduction so convincing? Gödel remains completely silent on this, and commentators seem to be at a loss at this point. Even such a knowledgeable and thorough commentator as Wilfried Sieg can do nothing but wonder. Sieg writes thus:

> There is no explanation of why such a reduction is the most satisfactory way of getting to a precise definition or, for that matter, of 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 reduced to computations of finite machines. (Sieg 2006, p. 196)[12]

3. Turing's Analysis and the "Stumbling Block"

As it turns out, Sieg's interpretation (or lack thereof) on this question forms a part of his global thesis that in a deep sense Gödel failed to recognize the genuinely distinctive character of Turing's analysis. Indeed, in Sieg's view, the failure is not limited to Gödel but ascribable to many of the leading researchers in the relevant fields of the time. In this connection, Sieg speaks of a "stumbling block" which they faced in the analysis of the notion of computation. To make the story short, many of them explicated the informal notion of effective calculability for number-theoretic functions by the formal computability of their values in a deductive formalism. But they were unable, in the end, to provide a systematic reason for insisting that the elementary steps of an effective procedure (governing derivations in a logic) must be recursive without presupposing the relevant property of the formalism.[13] On Sieg's account, Turing overcame this stumbling block by making a shift from formal-logically meaningful steps to symbolic steps underlining them and thereby appropriately brought in human computers in such a manner that the relevant property, i.e., recursiveness, of computation can be derived from (or guaranteed by) the (finite) character of human processing capacities, when proceeding mechanically.[14]

Now, to get back to Sieg's interpretation of Gödel's "failure," our question

---

[12] On Gödel's claim in his (193?) that the correctness of characterizing the computations of number-theoretic functions via the equational calculus was established by Turing, Sieg similarly writes, "I cannot see at all how Turing's reductive steps can be adapted to argue for Gödel's analytic claim." (Sieg 2006, p. 199)

[13] In this connection, it is to be noted that in his (1936) Gödel maintains that the notion of computability is "absolute":

> . . . a function computable in one of the system Si, or even in a system of transfinite order, is computable already in S1. 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. (*Collected Works I*, p. 399)

However, as Sieg points out, "the absoluteness was achieved, ironically, only relative to the description of the 'formal' system Si; the stumbling block shows up exactly here." (Sieg 1994, p. 88)

then is whether Gödel too was unable to overcome the stumbling block. In thinking about this question, there is another remark by Gödel on Turing which we must consider. In what was presented as a footnote to his 1964 postscript to the 1934 lectures (on the occasion of their reproduction in Davis 1965), Gödel made a remark, which he titled "*A philosophical error in Turing's work.*" It reads like this:

> Turing in his [1936, section 9] gives an argument which is supposed to show that mental procedures cannot go beyond mechanical procedures. However, this argument is inconclusive. What Turing disregards completely is the fact that *mind, 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. 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 (and, therefore, also Turing's number of *distinguishable states of mind*) may *converge toward infinity* in the course of the application of the procedure. (Gödel 1972a)

Given the fact that this remark was appended by Gödel as a footnote to his 1964 postscript to the Princeton lectures, we naturally wonder if (and how) we can make sense of Gödel's intent at all. For the postscript in question begins with the following words:

> 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 concept of formal system can now be given, the existence of undecidable arithmetical propositions and the non-demonstrability of the

---

[14] According to Sieg's articulation, the central restrictive conditions include:
   (Boundedness) There is a fixed bound on the number of configurations a computer can immediately recognize.
   (Locality) A computer can change only immediately recognizable (sub-) configurations.

> consistency of a system in the same system can now be proved rigorously for every consistent formal system containing a certain amount of finitary number theory.
>
> Turing's work gives an analysis of the concept of 'mechanical procedure' (alias 'algorithm' or 'computation procedure'). The concept is shown to be equivalent with that of a 'Turing machine.' (Davis 2004, pp. 71-72)

How could Gödel praise Turing's definition of computability, on the basis of which he maintains the generality of the incompleteness theorem, and, at the same time, criticizes Turing's argument for this definition as containing a "philosophical error" and "inconclusive"? Is Sieg right on the mark after all? Or indeed, was Gödel not only unable to recognize the genuinely distinctive character of Turing's analysis but even confused in some way?

4. The Finite and Mechanical Character of Formal System

Against Sieg's interpretation, Oron Shagrir recently put forward an alternative reading, which is "more charitable" to Gödel. Here I quote his account at some length:

> My suggestion is that we take him as holding the view that the finite and mechanical character of computation is not matter of the human condition, but of the epistemic role of computation in the foundations of mathematics, and, in particular, in the finitistic program of Hilbert. Let me explain. There is a major difference between the historical contexts in which Turing and Gödel worked. Turing tackled the Entscheidungsproblem as an interesting mathematical problem worth solving; he was hardly aware of the fierce foundational debates. Gödel, on the other hand, was passionately interested in the foundations of mathematics. Though not a student of Hilbert, his work was nonetheless deeply entrenched in the framework of Hilbert's finitisic program, whose main goal was to provide a meta-theoretic finitary proof of the consistency of a formal system "containing a certain amount of finitary number theory." In this foundational context, a formal mathematical system is just another name for this finite and mechanical procedure. Thus the

procedure's finite and mechanical nature is a given and not open to question. Its finite and mechanical nature is underwritten by its role in the foundational project, which is defining a formal mathematical system. (Shagrir 2006, pp. 411-412)

As we saw above, Gödel was led to the problem of computability in his attempt to generalize the incompleteness theorems, and, for this purpose, it was necessary to obtain a rigorous characterization of a formal system and thus of a mechanical procedure. That is why he welcomed Turing's reduction of the concept of mechanical computational procedure to that of a machine with a finite number of parts. But why does Gödel think that a formal system is a finite and mechanical procedure *in the first place*? Isn't that because he understands the concept of formal system within the framework of Hilbert's foundational project? As Shagrir points out, in the context of Hilbert's foundational project, formal systems of mathematics constitute domains which are governed by finite and mechanical procedures. Hilbert thought it possible to establish the consistency of mathematics by an evidential and thus consistent means *precisely because* the whole mathematical practice can be represented in such formal domains. Therefore, if Gödel's thinking is actually conducted in this particular problematic, then the finite and mechanical character of a formal system is a given and not open to question. Shagrir's account fits quite well with all this and makes Gödel's "silence" intelligible.[15]

How does Shagrir explain Gödel's other, negative remark on Turing, then? As we just saw, on his account, the finite and mechanical character of a formal system is *not* something one must argue for. It would seem to follow that, in putting forward an argument for the definition, Turing is doing something uncalled for. Indeed, this is precisely how Shagrir interprets Gödel's remark:

Turing's error, on this account, is anchoring the procedure's finite and mechanical character in the human condition, specifically, in the number of states of mind. Turing's analysis, according to Gödel, does not establish that a computation procedure is a finite and mechanical procedure, for

---

[15] But would this not imply that Gödel did not need Turing's (or anybody's) argument at all to accept the Church-Turing thesis? I shall come to this question shortly.

this is not questionable at all. . . . Gödel used the notion of a *mechanical* and *finite* procedure/computation before he encountered Turing's analysis, at about the time he rejects Church's proposal as "thoroughly unsatisfactory." By Gödel's lights, even if it turns out that a human has infinite memory or can carry out infinitely many steps in finite time, this would not change either the definition of computability or that of a formal system. (Ibid, pp. 412-413)

Well enough. This, however, cannot be the end of the story. Why? Because, after all, Gödel's objection is *not* that Turing is doing something unnecessary, and there must be some other, perhaps, deeper reason that motivates Gödel to ascribe the "philosophical" error to Turing. Now, Shagrir is well aware of the point and provides an explanation. According to him, Gödel's concern is that "any system constrained by the finiteness conditions set down by Turing cannot transcend the computable." (Shagrir 2006, p.406). If these conditions do apply to the human mind, it would follow that it cannot surpass the power of a Turing machine. And to avoid such a conclusion, Gödel rejects one of the conditions Turing refers to in his analysis, i.e., the constraint on the number of distinguishable states of mind. So, on Shagrir's account, what ultimately explains Gödel's criticism against Turing is his view on the power of the human mind.

There are two problems with Shagrir's account, I think. One is concerned with the timing of Gödel's criticism. Given that Gödel's criticism was motivated by his belief in the nature of the human mind, we might wonder since when he had held such a belief. Shagrir explains this point by referring to Gödel's conversation with Hao Wang. In Gödel's opinion, reports Wang, one of the most interesting, rigorously proved results about minds and machines can be expressed by the following disjunction:

> Either the human mind surpasses all machines (to be more precise: it can decide more number theoretical questions than any machine or else there exist number theoretical questions undecidable for the human mind. (Wang 1974, p. 324)

As Shagrir points out, already in his 1951 Gibbs Lecture, Gödel talked of the disjunction, contending it follows from the incompleteness results.[16] One

question we might want to ask Shagrir then is this. Why didn't Gödel say anything about Turing's "philosophical error" then and there? Or, to put it differently, why did Gödel have to wait so long (for at least 13 years) to express his view, and in the form of a footnote to a postscript? Considering the importance Gödel attaches to the question as to the power of the human mind, it seems very strange that he had kept it to himself for so long.

Second, as we saw above, if Shagrir's account is on the right track, for Gödel the finite and mechanical nature is underwritten by its role in the Hilbertian foundational project and hence a given and not open to question. But does this mean that Gödel did not need Turing's (or anybody's) argument at all to accept the Church-Turing thesis? On this point Shagrir says he is not suggesting that Gödel sees no connection between mechanical computability and human computability:

> Quite the contrary: the epistemic context requires that a human be able, at least in principle, to follow the computation procedure, i.e., to check whether a configuration of symbols constitutes a formal proof or not. Gödel praises Turing precisely for this, for analyzing the concept of a human who follows a finite and mechanical procedure. (Shagrir 2006, p. 413)

On Shagrir's account, Gödel understood, whenever he praised Turing's definition, the latter's analysis in Turing (1936) to provide an argument or some sort of justification insofar as it is an argument for defining mental procedures *in calculating humans* in terms of a machine with a finite number of parts. In other words, Gödel took Turing's claim there to be a conditional or restricted one. The question, then, is this. Recall that Gödel claims in his remark "A philosophical error in Turing's work" that Turing gives on the second page of section 9 "an argument which is supposed to show that mental

---

[16] There the disjunction is formulated this way:

> Either ... the human mind (even within the realm of pure mathematics) infinitely surpass the power of any finite machine, or else there exist absolutely unsolvable Diophantine problems of the type. (Gödel 1951, p. 310)

procedures cannot go beyond mechanical procedures. However, this argument is inconclusive." And recall also that, in the postscript to which this remark was added, Gödel praised Turing's definition of computability. How can Gödel think that one and the same argument has two different conclusions, one general, one conditional?

5. Finiteness and Mechanicalness

In my opinion, for the solution of this interpretative puzzle, another piece is needed. First of all, it must be admitted that, in his negative remark on Turing, Gödel assumes that Turing, in his (1936), makes a claim concerning *general* mental processes. Gödel strongly disagrees with this claim and rejects it. But why? It is because Gödel thinks that there exist finite and *non-mechanical* mental procedures, by means of which the mind can surpass infinitely the powers of any finite machine. Gödel argues for this thesis using the dilemma which he says follows from the incompleteness results, but I won't discuss his argument here. Instead, what I want to do is draw your attention to the following, rather unexciting but very important fact: in order to speak of the possibility of finite and non-mechanical procedures, for Gödel the two terms "finite" and "mechanical" cannot be interchangeable. In this connection, it is to be recognized that, in the postscript to the 1934 Princeton lectures, Gödel points precisely to this distinction. There he says that a formal system can simply be defined to be any mechanical procedure for producing provable formulas and also that the equivalence of finite computation and general recursiveness he alluded to in the footnote can be considered correct "if 'finite procedure' is understood to mean 'mechanical procedure'." He also makes it sure to add a parenthetical remark that "the question of whether there exist finite *non-mechanical* procedures not equivalent with any algorithm, has nothing whatsoever to do with the adequacy of the definition of 'formal system' and of 'mechanical procedure'."[17] But exactly when did Gödel start making

---

[17] He then further states that "the result mentioned in this postscript do not establish any bounds for the powers of human reason, but rather for the potentialities of pure formalism in mathematics." Incidentally, Gödel's negative remark about Turing's analysis was meant to be a footnote to the word "mathematics" in this passage.

a clear distinction between the two terms? Confining ourselves just to those papers we have mentioned here, it can be seen that in 1934 Gödel says "finite procedure" and "finite computation," in 1938 "mechanical procedure," in 1951 he reverts to "finite procedure."

It is my contention that Gödel came to the clear distinction some time after the publication of his 1958 so-called *Dialectica* article, at the very earliest. In this paper, which he titled "On an extension of finitary mathematics which has not yet been used,"[18] Gödel introduces a distinction in the concept of *finitary mathematics* between two components: the *constructive* element demands that objects and facts referred to in mathematics must be exhibitable or obtainable by construction or proof, whereas the *specifically finitistic* element demands, in addition to the constructiveness requirement, that mathematical objects and facts must be given in concrete mathematical intuition understood in the sense of Hilbert's foundational program, which would seem to imply that mathematics is to be concerned only with combinatorial properties and relations of concrete objects such as symbols. Gödel then proposes that the second, specifically finitistic, element be dropped and suggests to explore the possibility of the finitary mathematics which is based upon insights into abstract concepts, which have as their content "thought structure" or "thought content." According to him, in the proofs of propositions about these mental objects, insights are needed which are derived from a reflection upon the meanings involved.[19] Is Gödel here proposing a finitary mathematics which is distinct from that of Hilbert's? In the 1958 paper itself, Gödel writes nothing to that effect. But, later in the process of revising the English translation of the paper, Gödel made a remarkable comment. In the 1958 paper a footnote which refers to Hilbert's 1926 paper "On the infinite" was attached to the passage where the first occurrence of the term "finitary mathematics" is found,

---

[18] To be precise, the original German title is "Über eine bisher noch nicht benützte Erweiterung des finiten Standpunkts," which is different from the one mentioned here. The latter is Gödel's own translation when he revised the English translation of the 1958 paper, which was done by someone else. Hereafter for all the quotations from Gödel 1958, I use the text of Gödel 1972.

[19] I do not discuss here what sort of epistemology Gödel has in mind or what exactly it involves, nor how it is related to the consistency proof he presents in the *Dialectica* paper.

but in about 1968 Gödel added a new footnote to the previous one and wrote as follows:

> Note that it is Hilbert's insistence on concrete knowledge that makes finitary mathematics so surprisingly weak and excludes many things that are just as incontrovertibly evident to everybody as finitary number theory. [. . .] There is nothing in the term "finitary" which would suggest a restriction to concrete knowledge. Only Hilbert's special interpretation of it introduces this restriction. (Gödel 1972, p. 272)

What Gödel has in mind here must correspond to the (finitary) mathematics whose evidence is based upon our insight into higher-type, abstract objects. This can be confirmed by the fact that Gödel, in the postscript to the 1934 lectures, attaches a footnote to the expression "finite *non-mechanical* procedures" which describes the procedures in question as "such as involve the use of abstract terms on the basis of their meaning."[20] All these facts seem to me to point to one thing. That is, in distinguishing the two component parts in the concept of finitary mathematics, Gödel was indeed trying to arrive at a finitary mathematics, which is different from Hilbert's finitistic attitude. When this distinction is made, "finite" becomes distinguishable from "mechanical," and thus it becomes possible to talk of the possibility of finite and non-mechanical procedures.

6. Conclusion

So what does all this mean as regards Gödel's remark about Turing? What we should recognize is that once "finite" is distinguished from "mechanical," we have in hand two different types of finite procedures, mechanical and non-mechanical. This in turn means that for Gödel there are two different kinds of computability. That Gödel indeed has such a conception of computability

---

[20]   The (mental) procedures which Gödel mentions in his negative remark on Turing must also be considered to correspond to those he talks of in the *Dialectica* article. It is by no means a mere coincidence that these remarks all come from the same period, ca 1964.

can be confirmed by the following remark, which he made in his letter of 14 August, 1964 to Jean van Heijenoort:

> In my opinion Church's Thesis is unquestionably correct for mechanical, but perhaps incorrect for *intuitionistic*, computability (as I clearly stated in the postscript to my 1934 lectures). (*Collected Works 5*, p. 316, my italics)

The answer to our interpretative puzzle, I believe, is contained here. That is, corresponding to the double-meaning of the term "computability," Gödel did have two different ways of reading Turing's analysis. When the term is understood in the sense of "finite, mechanical procedure," Gödel completely agrees with Turing's definition and can endorse his argument for the definition. On the other hand, Gödel rejects Turing thesis and his analytic argument for the thesis if "computability" is understood to mean "finite, non-mechanical procedure."

Finally, what does this double-conception of computability say about Gödel's thinking? It is most important that we understand what ultimately motivates Gödel to introduce the distinction in question. I think it is no other than his interest in what he alluded to in his 1931 paper as a finitary consistency proof which is unformalizable in P, the formal system of arithmetic. Gödel might have thought that even if the project of a finitary consistency proof in the sense of Hilbert is bankrupt, this does not necessarily mean the bankruptcy of a finitary consistency proof *as such*. Indeed, in the revised version of the *Dialectica* article, Gödel writes, "due to the lack of a precise definition of either concrete or abstract evidence there exists, today, no rigorous proof for the insufficiency (even for the consistency proof of number theory) of finitary mathematics." (Gödel 1972, p. 273)[21] Behind this reservation lies Gödel's strong belief in the solvability of mathematical problems, which Hilbert too held, and it also means that Gödel' thinking was motivated, up till the very end of his research career, by philosophical and especially epistemological concerns.

References

Church, Alonzo.   1936 "An Unsolvable Problem of Elementary Number Theory,"
    American Journal of Mathematics, 58, 345-363; reprinted in Davis 2004, 89-107.
                1956 *Introduction to Mathematical Logic*, Princeton, Princeton
    University Press.
Davis, Martin.   1965 *The Undecidable*, New York, Raven Press.
                1982 "Why Gödel didn't have Church's Thesis," *Information and
    Control*, 54, 3-24.
                2004 *The Undecidable*, New York, Dover Publications Inc.; a
    corrected republication of Davis 1965.
Gödel, Kurt.   1931   "Über formal unentscheidbare Sätze der Principia
    Mathematica und verwant der Systeme, I" *Monatshefte für Mathematik und Physik*,
    38, 173-198; reprinted in *Collected Works*, I, 144-195.
                1933 "The Present Situation in the Foundations of Mathematics,"
    Lecture delivered a the meeting of the Mathematical Association of America in
    Cambridge Massachusetts, December 29-30, 1933; reprinted in *Collected Works*,
    III.
                1934 "On Undecidable Propositions of Formal Mathematical
    Systems," lecture notes, Princeton; reprinted in Davis 2004, 39-71, and *Collected*

---

[21] In this context, Gödel's letter of 2 August, 1962 to Leon Rappaport is also of great importance. There he writes:

> Nothing has been changed lately in my results or their philosophical consequences, but perhaps some misconceptions of them have been dispelled or weakened. My theorem only shows that the *mechanization* of mathematics, i.e., the elimination of the *mind* and of *abstract* entities, is impossible, if one wants to have a satisfactory foundation and system of mathematics.
>
> I have not proved that there are mathematical questions undecidable for the human mind, but only that there is no *machine* (or *blind formalism*) that can decide all number theoretical questions (even of a certain special kind).
>
> Likewise it does not follow from my theorems that there are no *convincing* consistency proofs for the usual mathematical formalisms, notwithstanding that such proofs must use modes of reasoning not contained in those formalisms. What is practically certain is that there are, for the classical formalisms, no conclusive *combinatorial* consistency proofs (such as Hilbert expected to give), i.e., no consistency proofs that use only concepts referring to finite combinations of symbols and not referring to any infinite totality of such combinations.
>
> I have published lately (see Dial., vol. 12 (1958) p. 280) a consistency proof for number theory which probably for many mathematicians is just as convincing as would be a combinatorial consistency proof, which however uses certain abstract concepts (in the sense explained in this paper).

Works, I, 346-371.

          193? "Undecidable Diophantine Propositions," reprinted in *Collected Works*, I, 164-175.

          1951 "Some Basic Theorems on the Foundations of Mathematics and their Implications," in *Collected Works*, III.

          1958 "On a Hitherto Unutilized Extension of the Finitary Standpoint," reprinted in *Collected Works*, II, 241-251.

          1972 "On an extension of finitary mathematics which has not yet been used," reprinted in *Collected Works*, II, 271-279.

          1972a "Some remarks on the undecidability results," reprinted in *Collected Works*, II, 305-306.

          1986-1995 *Collected Works*, Vols. I-III. Oxford, Oxford University Press.

Hilbert, David, and Wilhelm Ackermann. 1928 *Grundzüge der theoretischen Logik*, Berlin, Springer.

Kleene, Stephen. 1936 "General Recursive Functions of Natural Numbers," *Mathematische Annalen*, 112, 5, 727-742; reprinted in Davis 2004, 236-252.

Ogawa, Yoshinori. 2004 "The Pursuit of Rigor: Hilbert's axiomatic method and the objectivity of mathematics," *Annals of the Japan Association for Philosophy of Science*, Vol.12, No.2, 29-48.

Shagrir, Oron. 2006 "Gödel on Turing on Computability," in Adam Olszewski, Jan Wolenski and Robert Janusz eds., *Church's Thesis after 70 Years*, Ontos-Verlag, 393-419.

Shapiro, Stewart. 1980 "On the Notion of Effectiveness," *History and Philosophy of Logic*, 1, 209-230.

          1981 "Understanding Church's Thesis," *Journal of Philosophical Logic*, 10, 353-365.

          1983 "Remarks on the Development of Computability," *History and Philosophy of Logic*, 4, 203-220.

          1994a "Reasoning, Logic and Computation," *Philosophia Mathematica, 3*, vol.2, 31-51.

          1994b "Metamathematics and Computability," in Ivor Grattan-Guinness, ed., *Companion Encyclopedia of the History and Philosophy of the Mathematical Sciences*. London, Routledge.

Sieg, Wilfried. 1994 "Mathematical Procedures and Mathematical Experience," in Alexander George ed., *Mathematics and Mind*, Oxford University Press.

          1997 "Step by Recursive Step: Church's Analysis of Effective Calculability," *Bulletin of Symbolic Logic* 3, 154-180.

          2005 "Only Two Letters: the Correspondence between Herbrand and Gödel," *Bulletin of Symbolic Logic* 11, 172-184.

          2006 "Gödel on Computability," *Philosophia Mathematica* (III) 14, 189-207.

Sieg, Wilfried and Daniele Mundici. 1994 "Paper Machines," *Philosophia Mathematica*, 3, vol.2, 5-30.

Turing, Alan. 1936 "On Computable Numbers, with an Application to the Entscheidungsproblem," Proceedings of the London Mathematical Society, 42, 230-265; reprinted in Davis 2004, 116-151.

van Heijenoort, J. 1967 *From Frege to Gödel: a source book in mathematical logic*, 1879-1931. Cambridge, MA, Harvard University Press.

Wang, Hao. 1974 *From Mathematics to Philosophy*, London, Routledge & Kegan Paul.