GÖDEL'S PHILOSOPHICAL CHALLENGE (TO TURING):
The human mind infinitely surpasses any finite machine

Wilfried Sieg

Preamble. Turing's work is connected to discussions on the foundations of mathematics, and I don't mean the polemical ones in the 1920s, but rather the substantively difficult ones that reach back to the second half of the 19th century
[1] and to the opposing views of Dedekind and Kronecker, representing an *abstract, conceptual* and a *concrete, computational* approach to mathematics. In his attempt
[2] to mediate between these two approaches, Hilbert – around the turn from the 19th to the 20th century – expanded the discussion to include "metamathematics": Dedekindian in the axiomatic theory, but Kroneckerian in its presentation. That led him, in 1917, to propose the investigation of "mathematical proof", proof
[3] taken here in a broad, informal sense.

For his finitist consistency program that began early in 1922, Hilbert insisted of course on formal presentations and articulated their significance in
[4] 1927 as follows: But does our mathematical thinking proceed according to formal rules? Can it be represented faithfully in formal theories? And, most pertinently, what are "formal theories" to begin with?. It is in this broader historical and
[5] intellectual context, I would like to present reflections of Gödel and Turing on computability and the human mathematical mind. They will be schematic, emphasizing only main themes.[1]

Both Gödel and Turing proved remarkable and important theorems. *Gödel* obtained in 1930/31 his incompleteness theorems that undermined the finitist program, but opened the way to broader versions. *Turing* proved in 1936 the undecidability of first-order logic. These results were used by Gödel to draw, most significantly in the early 1950s, conclusions concerning the human mind. Implicitly, those conclusions raised a philosophical challenge to Turing; thus the
[6] title of my talk: "To Turing" is flanked by parentheses, as the philosophical challenge issued by Gödel's mathematical results was not only a challenge to

[1] For details, if you would like to look at them and to hear also other voices, Papers: On computability; Hilbert's Proof Theory; Searching for proofs.

Turing but also to Gödel himself. Can one rigorously argue from these results, as [7] Gödel did in his 1951 Gibbs Lecture, to the claim that *the human mind infinitely surpasses any finite machine*?

To even understand this assertion one is forced to review the emergence of a rigorous notion of computability in the 1930s, as well as Gödel and Turing's [8] role in securing its conceptual foundation. The first part of my talk thus presents Overview their contributions; it is entitled *Absoluteness & computors*. The central issue is addressed in the second part, *Beyond mechanisms & discipline*, and is further explored in the third part, *Finding proofs with ingenuity*.

[9] **1. Absoluteness & computors.**

When Gödel was confronted in 1934 with Church's proposal of identifying the calculability of number-theoretic functions with their λ-definability, he viewed the proposal as "thoroughly unsatisfactory". He proposed, instead, "to state a set of axioms which would embody the generally accepted properties of this notion [calculability], and to do something on that basis". But in his contemporaneous Princeton lectures he did not formulate axioms for that notion, but rather modified Herbrand's definition of a "calculable finitist function" to arrive at the concept of a "general recursive function". At the time, he did *not* think of general recursiveness as a rigorous explication of calculability.

Only in 1935 did it become "plausible" to Gödel that his [incompleteness] results were "valid for all formal systems". The plausibility of this claim rested on an observation concerning computability (i.e., general recursiveness) he had made in the Postscriptum to his *On the length of proofs*. Here is the observation for systems S_i of i-th order arithmetic, i>0: if a function is computable in one of the systems S_i then it is calculable in S_1. Gödel concludes that this notion

> [10] … 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.

In his contribution to the Princeton Bicentennial Conference, Gödel generalized the absoluteness claim to *any formal* system containing arithmetic and attributed the philosophical significance of general recursiveness almost exclusively to its absoluteness. Gödel started his talk with a remark on Tarski's earlier lecture;

Tarski had stressed, and Gödel thinks justly, the "great importance of the concept of general recursiveness (or Turing computability)". Gödel continues:

[11] 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. (*Collected Works II*, p.150)

In the 1951 Gibbs Lecture, Turing's notion becomes a focal point in Gödel's reflections. There he explores the implications of the incompleteness theorems, not in their original formulation, but in a much more satisfactory general form. 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." This is followed by a remark concerning Turing:

[12] 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. (*Collected Works III*, pp. 304-5)

Gödel does not expand on this remark: he gives no hint of how *reduction* is to be understood and how it yields a precise definition of finite procedure; he does not explain why the concept of a machine with a finite number of parts is equivalent to that of a Turing machine.

The appreciation of Turing's work indicated in the Gibbs lecture for the first time is deepened in two brief remarks published during Gödel's lifetime after 1951: (i) the 1963 *Postscriptum* to his classical incompleteness paper and (ii) the 1964 *Postscriptum* to his Princeton Lecture Notes. The 1963 *Postscriptum* re-emphasizes the centrality of Turing's work for both incompleteness theorems, as it makes possible "a precise and unquestionably adequate definition" of the general notion of formal system. Thus, Gödel concludes:

… 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.

In the *Postscriptum* to his Princeton Notes, Gödel repeats this remark almost verbatim, but then expresses why Turing's work plays such an important role:

[13] 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".

Though Gödel emphasizes that equivalent definitions like recursiveness and λ-definability are much less suitable for his purpose, he does not elucidate the special character of Turing computability. In particular, he does not indicate, how he thought an analysis proceeded or how the equivalence between the (analyzed) concept and Turing computability could be proved.

Let us turn then directly to Turing's work in section 9 of his *On computable numbers*. Call a human computing agent who proceeds mechanically a *computor;* such a computor is taken to operate deterministically on finite, possibly two-dimensional configurations of symbols when performing a calculation. 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 a "finite bound" on the number of configurations and to the "locality" of operations. As Turing considers the two-dimensional character of configurations as inessential for
[14] calculations, he can be taken to have shown: computors that satisfy boundedness and locality conditions and operate deterministically on strings ARE *string machines;* two-letter Turing machines can mimic the latter.

Are Turing machines then, as Gandy put it, *codifications* of computors? Is Gandy right when claiming that Turing provided (the outline of) a proof for the
[15] claim, "What can be calculated by an abstract human being working in a routine way is computable"? Turing himself found his considerations mathematically unsatisfactory: indeed, they do not constitute a *proof* of a general claim for arbitrary *configurations*! This difficulty can be avoided by characterizing the
[16] structural (axiomatic) notion of a *Turing Computor* as a discrete dynamical system satisfying abstract boundedness and locality conditions.[2] Then one can show that a Turing machine can carry out the computations of any model of the axioms. No appeal to a thesis is needed: that appeal is replaced by recognizing the correctness of axioms for the intended notion: *human computing agent*!

This way of extracting from Turing's analysis axiomatic conditions and then establishing a representation theorem seems to follow Gödel's suggestion to Church; it also seems to fall under the description Gödel gave of Turing's work, when arguing that it gives an analysis of the concept mechanical procedure, and that "this concept is shown to be equivalent with that of a Turing machine". With

[2] … like notion of group, field, topological space …

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.

[17] **2. Beyond mechanisms & discipline.**

Gödel's [193?] begins by referring to Hilbert's famous words: "For any precisely formulated mathematical question a unique answer can be found." He takes these words to mean 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
[18] from this notion of proof to a formalized one: "... it is not possible to formalize 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 mechanize mathematical reasoning; ..." That means for Gödel, "it will never be possible to replace the mathematician by a machine, even if you confine yourself to number-theoretic problems."

In his 1972-note *A philosophical error in Turing's work* Gödel considers a 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". Then he points to a *fact* that weighs against this 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 presents axioms of infinity, "which only explicate the content of the general concept of set". Gödel calls the existence of an unexplored series of axioms a *fact*, but asserts 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)": it would require "a substantial advance in our understanding of the basic concepts of mathematics". Such an advance is then connected with the dynamic development of the human mind.

> [19] ... *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.

In a closely related note Gödel remarks:

[20] 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.

I have to admit that I don't fully understand these enigmatic observations; nevertheless, three points can be made: (1) mathematical experience is invoked when asking the right questions; (2) aspects of that experience may be codified in a mechanical procedure and serve as the basis for asking those questions; (3) answers may involve abstract terms that are introduced by the non-mechanical mental procedure. 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 deals explicitly with mental processes without claiming that they are all mechanical. Turing machines are used to model mechanical processes; in contrast, machines that are to exhibit intelligence have a more complex structure, in particular, they interact with their environment. Indeed, one might say that Turing is trying to capture what Gödel described as a central feature of humanly effective calculability: mind, in its use, is not static, but constantly developing. In his *Intelligent machinery* Turing states that an untrained infant's mind must acquire discipline (as being a sort of universal machine):

[21] 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. (p. 21)

How can we transcend discipline, when doing mathematics? Turing provides a hint in his 1939 paper, where ordinal logics are introduced to expand formal theories systematically. Turing distinguishes there between *ingenuity* and *intuition*. In formal logics, he observes, they have definite roles: 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:

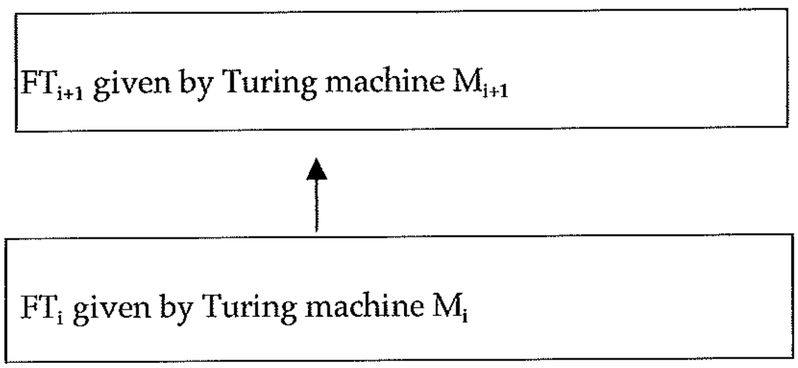
[22] In pre-Gödel times it was thought by some that it would be possible to carry this programme to such a point that all the intuitive judgements of mathematics could be replaced by a finite number of these rules. The necessity for intuition would then be entirely eliminated. (p. 209)

Gödel's first theorem says of course that intuition in Turing's sense can't be eliminated. Turing takes this *mathematical objection* to his view concerning intelligent machinery seriously. He gives two responses in his most famous paper [1950, pp. 444-5]. The short one states that the objection takes for granted, without any proof, that the human intellect is not limited in the way machines provably are. However, Turing thinks that the objection cannot be dismissed so lightly and gives a second response. He acknowledges the superiority of the human intellect w.r.t. a single machine (we can recognize the truth of "its" Gödel sentence), but views that as a petty triumph:

[23] 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 does not offer a proof of the claim that there is "no question of triumphing simultaneously over *all* machines". [Obvious conflict with Gödel's assertion.] But let us analyze matters a bit more closely and consider this diagram (where you have to add an upward arrow):

[24]



A given formal theory FT_i has been expanded to theory FT_{i+1}; the theories are presented via Turing machines M_i and M_{i+1}. The transition from one theory to the next and, correspondingly, from one Turing machine to the next is given in Gödel's case by a non-mechanical method that takes into account aspects of human mathematical experience including the understanding of abstract notions. In Turing's case the addition of new intuitive steps is similarly based on a computer's experience and learning; Turing emphasizes at a number of places that a random element be introduced, thus providing an additional feature in the development of machines that releases them from strict discipline. – As a matter of fact, Turing conjectured in [1947, p. 103]:

[25] 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.

This expectation has not been borne out yet, and Gödel would not be surprised. However, he could cooperate with Turing on the "philosophical questions of what can in principle be done". Indeed, they could agree, terminologically, that there is a human mind whose working is not reducible to the working of any particular brain: Turing had taken a step towards such a concept of human mind, when he emphasized at the end of *Intelligent Machinery*, "the isolated man does not develop any intellectual power":

> 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.

Turing calls this, appropriately enough, a *cultural search* in contrast to the more limited *intellectual searches* possible for individual men or machines. Finally, Gödel and Turing could explore and, possibly argue about, Turing's contention in [1951, p. 472] "that machines can be constructed, which will simulate the behavior of the human mind very closely". I take this as an inspiration towards concrete work in a more concrete context.

[26] **3. Finding proofs with ingenuity.**

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? The important question
[27] is whether we can gain, *by closely studying actual mathematical practice*, a deeper understanding of fundamental techniques and methods of mathematics. With that purpose in mind, let us return to Turing's distinction between intuition and ingenuity. Intuition is linked to the incompleteness of formal theories and provides an entry point to exploiting the parallelism between Turing and Gödel's considerations on expansions of theories! Ingenuity is involved in the more limited "intellectual searches" and is a capacity to "determine which steps are the more profitable for the purpose of proving a particular proposition".

Formulate strategies for automated proof search: not yet for proofs of new results, but for proofs that reflect logical and mathematical understanding; proofs that force us to make explicit the *ingenuity* required for a successful automated search. This involves reactions to Turing's remarks and impatient questions in a letter to Newman [Copeland 2004, p. 213]:

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?

[28] The logical framework for such studies must include a *structure theory of proofs*, i.e., an extension of proof theory that (i) articulates structural features of proofs, (ii) exploits the meaning of concepts, and (iii) makes explicit leading ideas or methods of proofs.

Underlying logic:

Intercalation calculus is properly between sequent and nd calculi; it was introduced to make possible the direct search for normal natural deduction proofs; indeed, ic-proofs are isomorphic to such proofs.

Case studies in Local Axiomatics:

[29] (i) Gödel's incompleteness theorems: automated proof at an abstract level assuming representability and derivability conditions axiomatically; we are now in the process of verifying these conditions (for ZF).[3]

[30] (ii) Cantor Bernstein theorem: we have taken a first important step by formally verifying the theorem starting with Zermelo's axioms - using intelligibly organized proofs; many of the intermediate lemmata are already found automatically.[4]

I am quite confident that we will begin in this way to make clearer where proofs involve ingenuity. I also hope that it is a beginning 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". Let me make one final remark: Gödel and Turing's views on "minds & machines" are usually seen in opposition to each other; Gödel's "fact" concerning a humanly effective, but non-mechanical procedure is of course rather striking. If one focuses on the real challenge presented by the incompleteness theorems, then one finds that Gödel and Turing pursue parallel

---

[3] The proof given is very close to Gödel's "first" proof of the result he had at the Königsberg Conference in September of 1930 – without arithmetization.

[4] We discovered that there is EXACTLY one proof modulo a single ingenious step; that step is taken in different ways to make explicit an inductively specified set – through approximation from below, from above or as a fixed-point of a monotone operator.

approaches with complementary goals, but different global perspectives. Turing views, in the very last sentence of his (1954), the limitative results as "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
[31] 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".

[32] THANK YOU!

[33]-[36] LOCAL AXIOMATICS III
(surrounding the Pythagorean
Theorem)