Natural Computing (2006)
DOI 10.1007/s11047-006-9021-9

On mind & Turing's machines

WILFRIED SIEG
Carnegie Mellon University, Pittsburgh, PA, USA
E-mail: ws15@andrew.cmu.edu

Abstract. Turing's notion of human computability is exactly right not only for obtaining a negative solution of Hilbert's *Entscheidungsproblem* that is conclusive, but also for achieving a precise characterization of formal systems that is needed for the general formulation of the incompleteness theorems. The broad intellectual context reaches back to Leibniz and requires a focus on *mechanical* procedures; these procedures are to be carried out by human computers without invoking higher cognitive capacities. The question whether there are strictly broader notions of effectiveness has of course been asked for both cognitive and physical processes. I address this question not in any general way, but rather by focusing on aspects of mathematical reasoning that transcend mechanical procedures. Section 1 discusses Gödel's perspective on mechanical computability as articulated in his [193?], where he drew a dramatic conclusion from the undecidability of certain Diophantine propositions, namely, that mathematicians cannot be replaced by machines. That theme is taken up in the Gibbs Lecture of 1951; Gödel argues there in greater detail that the human mind infinitely surpasses the powers of any finite machine. An analysis of the argument is presented in Section 2 under the heading *Beyond calculation*. Section 3 is entitled *Beyond discipline* and gives Turing's view of intelligent machinery; it is devoted to the seemingly sharp conflict between Gödel's and Turing's views on mind. Their deeper disagreement really concerns the nature of machines, and I'll end with some brief remarks on (supra-) mechanical devices in Section 4.

Key words: absolutely unsolvable (undecidable), axiom of infinity, Church's Thesis, Diophantine problem, finite machine, general recursive function, mechanical computability, objective mathematics, subjective mathematics, Turing machine

1. Mechanical computability

[Gödel 193?] is the unpublished draft for a lecture Gödel presumably never delivered; it was written in the late 1930s. Here one finds the earliest extensive discussion of Turing and the reason why Gödel, at the time, thought Turing had established "beyond any doubt" that "this really is the correct definition of mechanical computability". Obviously, we have to clarify what "this" refers to, but first I want to give some of the surrounding context. Already in Gödel [1933] elucidated, as others had done before him, the mechanical feature of effective procedures by pointing to the possibility that machines carry them out. When insisting that the inference rules of precisely described proof methods have to be "purely formal" he explains:

> [The inference rules] 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. This has the consequence that there can never be any doubt as to what cases the rules of inference apply to, and thus the highest possible degree of exactness is obtained. (*Collected Works III*, p. 45)

During the spring term of 1939 Gödel gave an introductory logic course at Notre Dame. The logical decision problem is informally discussed and seen in the historical context of Leibniz's "Calculemus".[1] Before arguing that results of modern logic prevent the realization of Leibniz's project, Gödel asserts that the rules of logic can be applied in a "purely mechanical" way and that it is therefore possible "to construct a machine which would do the following thing":

> The supposed machine is to have a crank and whenever you turn the crank once around the machine would write down a tautology of the calculus of predicates and it would write down every existing tautology ... if you turn the crank sufficiently often. So this machine would really replace thinking completely as far as deriving of formulas of the calculus of predicates is concerned. It would be a thinking machine in the literal sense of the word. For the calculus of propositions you can do even more. You could construct a machine in form of a typewriter such that if you type down a formula of the calculus of propositions then the machine would ring a bell [if the formula is a tautology] and if it is not it would not. You could do the same thing for the calculus of monadic predicates.

Having formulated these positive results Gödel points out that "it is impossible to construct a machine which would do the same thing for the whole calculus of predicates". Drawing on the undecidability of predicate logic established by Church and Turing, he continues with a striking claim:

> 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.

I mention these matters to indicate the fascination Gödel had with the mechanical realization of logical procedures, but also his *penchant* for overly dramatic formulations concerning the human mind. He takes obviously for granted here that a mathematically satisfactory definition of mechanical procedures has been given.

Such a definition, Gödel insists in [193?], p. 166, is provided by the work of Herbrand, Church and Turing. In that manuscript he examines the relation between mechanical computability, general recursiveness and machine computability. This is of special interest, as we will see that his methodological perspective here is quite different from his later standpoint. He gives, on pp. 167–168, a perspicuous presentation of the equational calculus that is "essentially Herbrand's" and defines general recursive functions. He claims outright that it provides "the correct definition of a computable function". Then he asserts, "That this really is the correct definition of mechanical computability was established beyond any doubt by Turing." Here the referent for "this" has finally been revealed: it is the definition of general recursive functions. How did Turing establish that this is also the correct definition of *mechanical* computability? Gödel's answer 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 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. (*Collected Works III*, p. 168)

The implicit claim is clearly that a procedure is mechanical just in case it is executable by a machine with a finite number of parts. There is no indication of the structure of such machines except for the insistence that they have only finitely many parts, whereas Turing machines are of course potentially infinite due to the expanding tape.

The literal reading of the argument for the claim "this really is the correct definition of mechanical computability was established beyond any doubt by Turing" amounts to this. The equational calculus characterizes the computations of number-theoretic functions and provides thus "the correct definition of computable function". That the class of computable functions is co-extensional with that of *mechanically* computable ones is then guaranteed by "Turing's proof" of the equivalence between general recursiveness and machine computability.[2] Consequently, the definition of general recursive functions via the equational calculus characterizes correctly the mechanically computable functions. Without any explicit reason for the first step in this argument, it can only be viewed as a direct appeal to Church's Thesis.

If we go beyond the literal reading and think through the argument in parallel to Turing's analysis in his [1936], then we can interpret matters as follows. Turing considers arithmetic calculations done by a computor. He argues that they involve only very elementary processes; these processes can be carried out by a Turing machine operating on strings of symbols. Gödel, this interpretation maintains, also considers arithmetic calculations done by a computor; these calculations can be reduced to computations in the equational calculus. This first step is taken in parallel by Gödel and Turing and is based on a conceptual analysis; cf. the next paragraph. The second step connects calculations of a computor to computations of a Turing machine. This connection is established by mathematical arguments: Turing simply states that machines operating on finite strings can be proved to be equivalent to machines operating on individual symbols, i.e., to ordinary Turing machines; Gödel appeals to "Turing's proof" of the fact that general recursiveness and machine computability are equivalent.

Notice that in Gödel's way of thinking about matters at this juncture, the *mathematical theorem* stating the equivalence of general recursiveness and machine computability plays the pivotal role: It is not Turing's analysis that is appealed to by Gödel but rather "Turing's proof". The central analytic claim my interpretation attributes to Gödel is hardly argued for. On p. 13 Gödel just asserts, "... by analyzing in which manner this calculation [of the values of a general recursive function] proceeds you will find that it makes use only of the two following rules." The two rules as formulated here allow substituting numerals for variables and equals for equals. So, in some sense, Gödel seems to think that the rules of the equational calculus provide a way of "canonically" representing steps in calculations and, in addition, that his characterization of recursion is the most general one.[3] The latter is imposed by the requirement that function values have to be calculated, as pointed out in [1934], p. 369 top; the former is emphasized much later in a letter to van Heijenoort of April 23, 1963, where Gödel distinguishes his definition from Herbrand's. His definition, Gödel asserts, brought out clearly what Herbrand had failed to see, namely "that the computation (for all computable functions) proceeds by exactly the same rules". (*Collected Works V*, p. 308) By contrast, Turing shifts from arithmetically meaningful steps to symbolic processes that underlie them and can be taken to satisfy restrictive boundedness as well as locality conditions. These conditions cannot be imposed directly on arithmetic steps and are certainly not satisfied by computations in the equational calculus.

2. Beyond calculation

In [193?] Gödel begins the discussion by reference to Hilbert's "famous words" that "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 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; ..." Then he continues, in a way that is similar to the striking remark in the Notre Dame Lectures, "i.e., it will never be possible to replace the mathematician by a machine, even if you confine yourself to number-theoretic problems." (pp. 164–165)

The succinct argument for this conclusion is refined in the Gibbs Lecture of 1951. In the second and longer part of the lecture, Gödel gave the most sustained defence of his Platonist standpoint drawing the "philosophical implications" of the situation presented by the incompleteness theorems.[4] "Of course," he says polemically, "in consequence of the undeveloped state of philosophy in our days, you must not expect these inferences to be drawn with mathematical rigor." The mathematical aspect of the situation, he claims, can be described rigorously; it is formulated as a disjunction, "Either mathematics is incompletable in this sense, that its evident axioms can never be comprised in a finite rule, that is to say, the human mind (even within the realm of pure mathematics) infinitely surpasses the powers of any finite machine, or else there exist absolutely unsolvable Diophantine problems of the type specified ..." Gödel insists that this fact is both "mathematically established" and of "great philosophical interest". He presents on pages 11–13 an argument for the disjunction and considers its conclusion as "inevitable".

The disjunction is called in footnote 15 a theorem that holds for finitists and intuitionists as an implication. Here is the appropriate implication: If the evident axioms of mathematics can be comprised in a finite rule, then there exist absolutely unsolvable Diophantine problems. Let us establish this implication by adapting Gödel's considerations for the disjunctive conclusion; the argument is brief. Assume the axioms that are evident for the human mind can be comprised in a finite rule "that is to say", for Gödel, a Turing machine can list them. Thus there exists a mechanical rule producing all the evident axioms for "subjective" mathematics, which is by definition the system of all humanly demonstrable mathematical propositions.[5] On pain of contradiction with the second incompleteness theorem, the human mind cannot prove the consistency of subjective mathematics. (This step is of course justified only if the inferential apparatus for subjective mathematics is given by a mechanical rule, and if subjective mathematics satisfies all the other conditions for the applicability of the second theorem.) Consequently, the Diophantine problem corresponding to the consistency statement cannot be proved either in subjective mathematics. That justifies Gödel's broader claim that it is undecidable "not just within some particular axiomatic system, but by *any* mathematical proof the human mind can conceive". (p. 13) In this sense the problem is *absolutely undecidable* for the human mind. So it seems that we have established the implication. However, the very first step in this argument, indicated by "that is to say", appeals to the precise concept of "finite procedure" as analyzed by Turing. Why is "that is to say" justified for Gödel? – To answer this question, I examine Gödel's earlier remarks about finite procedures and finite machines.[6]

Gödel stresses in the first paragraph of the Gibbs Lecture that the incompleteness theorems have taken on "a much more satisfactory form than they had had originally". The greatest improvement was made possible, he underlines, "through the precise definition of the concept of finite procedure, which plays a decisive role in these results". Though there are a number of different ways of arriving at such a definition which all lead to "exactly the same concept", the most satisfactory way is that taken by Turing when "reducing the concept of a finite procedure to that of a machine with a finite number of parts". Gödel does not indicate the character of, or an argument for, the reduction of finite procedures to procedures effected by a machine with a finite number of parts, but he states explicitly that he takes finite machine "in the precise sense" of a Turing machine. (p. 9) This reduction is pivotal for establishing the central implication rigorously, and it is thus crucial to understand and grasp its *mathematical* character. How else can we assent to the claim that the implication has been established mathematically as a theorem? In his [1964] Gödel expressed matters quite differently: there he asserts that Turing in [1936] gave an analysis of mechanical procedures and showed that the analyzed concept is equivalent to that of a Turing machine. The claimed equivalence is viewed as central for obtaining "a precise and unquestionably adequate definition of the general concept of formal system" and for supporting, I would like to add in the current context, the mathematical cogency of the argument for the implication.

Gödel neither proved the mathematical conclusiveness of the reduction nor the correctness of the equivalence. So let us assume, for the sake of the argument, that the implication has been mathematically established and see what conclusions of great philosophical interest can be drawn. There is, as a first background assumption, Gödel's deeply rationalist and optimistic perspective that denies the consequent of the implication. That perspective, shared with Hilbert as we saw in Section 1, was articulated in [193?], and it was still taken in the early 1970s. Wang reports in [1974], pp. 324–325, that Gödel agreed with Hilbert in rejecting the possibility that there are number-theoretic problems undecidable for the human mind. Our task is then to follow the path of Gödel's reflections on the first alternative of his disjunction or the negated antecedent of our implication. That assertion states: There is no finite machine (i.e. no Turing machine) that lists all the axioms of mathematics which are evident to the human mind. Gödel argues for two related conclusions: (i) the working of the human mind is not reducible to operations of the brain, and (ii) the human mind infinitely surpasses the powers of any finite machine.[7]

A second background assumption is introduced to obtain the first conclusion: The brain, "to all appearances", is "a finite machine with a finite number of parts, namely, the neurons and their connections". (p. 15) As finite machines are taken to be Turing machines, brains are consequently also considered as Turing machines. That is reiterated in Wang, p. 326, where Gödel views it as very likely that "The brain functions basically like a digital computer." Together with the above assertion this allows Gödel to conclude in the Gibbs Lecture, "the working of the human mind cannot be reduced to the working of the brain".[8] In Wang it is taken to be in conflict with the commonly accepted view, "There is no mind separate from matter." That view is for Gödel a "prejudice of our time, which will be disproved scientifically (perhaps by the fact that there aren't enough nerve cells to perform the observable operations of the mind)". Gödel uses the notion of a finite machine in an extremely general way when considering the brain as a finite machine with a finite number of parts. It is here that the identification of finite machines with Turing machines becomes evidently problematic: Is it at all plausible to think that the brain has a similarly fixed structure and fixed program as a particular Turing machine? The argumentation is problematic also on different grounds; namely, Gödel takes "human mind" in a more general way than just the mind of any one individual human being. Why should it be then that mind is realized through any particular brain?

The proposition that the working of the human mind cannot be reduced to the working of the brain is thus not obtained as a "direct" consequence of the incompleteness theorems, but requires additional substantive assumptions: (i) there are no Diophantine problems the human mind cannot solve, (ii) brains are finite machines with finitely many parts, and (iii) finite machines with finitely many parts are Turing machines. None of these assumptions is uncontroversial; what seems not to be controversial, however, is Gödel's more open formulation in [193?] that it is not possible to mechanize mathematical reasoning. That raises immediately the question, what aspects of mathematical reasoning or experience defy formalization? In his note [1974] that was published in Wang, pp. 325–326, Gödel points to two "vaguely defined" processes that may lead 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 was reiterated in a modified formulation [Gödel, 1972d] that was published only later in *Collected Works II*, p. 306. The [1972d] formulation of this note is preceded by [1972c], where Gödel gives *Another version of the first undecidability theorem* that involves number theoretic problems of Goldbach type. This version of the theorem may be taken, Gödel states, "as an indication for the existence of mathematical yes or no questions undecidable for the human mind" (p. 305). However, he points to a *fact* that "weighs against this interpretation'', namely, that ''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 points also here to axioms of infinity, ''which only explicate the content of the general concept of set'' (p. 306). If the existence of such effective, non-mechanical procedures is taken as a fact or, more cautiously, as a third background assumption, then Gödel's second conclusion is established: The human mind, indeed, infinitely surpasses the power of any finite machine.

Though Gödel calls the existence of an ''unexplored series'' of axioms of infinity a *fact*, he also views it as a ''vaguely defined'' procedure and emphasizes that it requires further mathematical experience; after all, its formulation can be given only once set theory has been developed ''to a considerable extent''. In the note [1972d] Gödel suggests that the process of forming stronger and stronger axioms of infinity 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''. In the note [1974], Gödel offers a *prima facie* startlingly different reason for not yet having a precise definition of such a procedure: it ''would require a substantial deepening of our understanding of the basic operations of the mind'' (p. 325).

Gödel's *Remarks before the Princeton bicentennial conference* in [1946] throw some light on this seeming tension. Gödel discusses there not only the role axioms of infinity might play in possibly obtaining an absolute concept of demonstrability, but he also explores the possibility of an absolute mathematical ''definition of definability''. What is most interesting for our considerations here is the fact that he considers a restricted concept of *human definability* that would reflect a human capacity, namely, ''comprehensibility by our mind''. That concept should satisfy, he thinks, the ''postulate of denumerability'' and in particular allow us to define (in this particular sense) only countably many sets. ''For it has some plausibility that all things conceivable by us are denumerable, even if you disregard the question of expressibility in some language'' (p. 3). That requirement, together with the related difficulty of the definability of the least indefinable ordinal, does not make such a concept of definability ''impossible, but only [means] that it would involve some extramathematical element concerning the psychology of the being who deals with mathematics.'' Obviously, Turing brought to bear on his definition of computability, most fruitfully, an extramathematical feature of the psychology of a human computor.[9] Gödel viewed that definition in [1946], the reader may recall, as the first "absolute definition of an interesting epistemological notion" (p. 1). His reflections on the possibility of absolute definitions of demonstrability and definability were encouraged by the success in the case of computability. Can we obtain by a detailed study of *actual* mathematical experience a deeper "understanding of the basic operations of the mind" and thus make also a "substantial advance in our understanding of the basic concepts of mathematics"?

3. Beyond discipline

Gödel's brief exploration in [1972d] of the issue of defining a non-mechanical, but effective procedure is preceded by a severe critique of Turing. The critical attitude is indicated already by the descriptive and harshly judging title of the note, *A philosophical error in Turing's work*. The discussion of Church's thesis and Turing's analysis is in general fraught with controversy and misunderstanding, and the controversy begins often with a dispute over what the intended informal concept is. When Gödel spotted a philosophical error in Turing's work, he *assumed* that Turing's argument in the 1936 paper was to show that "mental procedures cannot go beyond mechanical procedures". He considered the argument as 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. (*Collected Works II*, p. 306)

Turing did not give a conclusive argument for Gödel's claim, but then it has to be added that he did not intend to argue for it. Simply carrying out a mechanical procedure does not, indeed, should not involve an expansion of our understanding. Turing viewed the restricted use of mind in computations undoubtedly as static; after all, it seems that this feature contributed to the good reasons for replacing states of mind of the human computor by "more definite physical counterparts" in section 9, part III, of his classical paper.

Even in his work of the late 1940s and early 1950s that deals explicitly with mental processes, Turing does not argue that mental procedures cannot go beyond mechanical procedures. Mechanical processes are, as a matter of fact, still made precise as Turing machine computations; machines that might exhibit intelligence have, in contrast, a more complex structure than Turing machines. Conceptual idealization and empirical adequacy are now being sought for quite different purposes, and Turing is trying to capture clearly what Gödel found missing in his analysis for a broader concept of humanly effective calculability, namely, "... that mind, in its use, is not static, but constantly developing".[10] Gödel continued the above remark in this way:

> 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.

The particular procedure mentioned as a plausible candidate for satisfying this description is the process of forming stronger and stronger axioms of infinity. We saw that the two notes, [1972d] and [1974], are very closely connected. However, there is one subtle and yet substantive difference. In [1974] the claim that the number of possible states of mind may converge to infinity is obtained as a consequence of the dynamic development of mind. That claim is then followed by a remark that begins, in a superficially similar way, as the first sentence in the above quotation:

> 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.

Clearly, I don't have a full understanding of these enigmatic observations, but there are three aspects that are clear enough. 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 the right questions; third, the answers may involve abstract terms that are incorporated into 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. It allows us, even on the basis of a very limited understanding, to relate Gödel's reflections tenuously with Turing's proposal for investigating matters. Prima facie their perspectives are radically different, as Gödel proceeds by philosophical argument and broad, speculative appeal to mathematical experience, whereas Turing suggests attacking the problem largely by computational experimentation. That standard view of the situation is quite incomplete. In the paper *Intelligent machinery* written about 10 years after [1939], Turing states what is really the central problem of cognitive psychology:

> 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, W.S.]. ... 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? A hint is provided in Turing's 1939 paper, where he 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. (p. 209)

The distinction between ingenuity and intuition, but also the explicit link of intuition to incompleteness, provides an entry to exploit through concrete computational work the "parallelism" of Turing's and Gödel's considerations. Copying the residue in machines is the task at hand. It is extremely difficult in the case of mathematical thinking, and Gödel would argue it is an impossible one, if machines are Turing machines. Turing would agree. Before we can start copying, we have to discover at least partially the nature of the residue, with an emphasis on "partially", through some restricted proposals for finding proofs in mathematics. Let us look briefly at the broad setting.

Proofs in a formal logic can be obtained uniformly by a patient search through an enumeration of all theorems, but additional intuitive steps remain necessary because of the incompleteness theorems. Turing suggested particular intuitive steps in his ordinal logics; his arguments are theoretical, but connect directly to the discussion of actual or projected computing devices that appears in his *Lecture to London Mathematical Society* and in *Intelligent Machinery*. In these papers he calls for intellectual searches (i.e., heuristically guided searches) and initiative (that includes, in the context of mathematics, proposing new intuitive steps). However, he emphasizes (1947, p. 122):

> 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.

Gödel and Turing, it seems, could have cooperated on the philosophical questions of what can in principle be done. They also could have agreed, so to speak terminologically, that there is a human mind whose working is not reducible to the working of any particular brain. Towards the end of *Intelligent Machinery* Turing emphasizes, "the isolated man does not develop any intellectual power", and argues:

> It is necessary for him to be immersed in an environment of other men, whose techniques he absorbs during the first 20 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* and contrasts it with more limited, *intellectual searches*. Such searches, Turing says definitionally, can be carried out by individual brains. In the case of mathematics they would include searches through all proofs and would be at the centre of "research into intelligence of machinery". Turing had high expectations for machines' progress in mathematics; indeed, he was unreasonably optimistic about their emerging capacities. Even now it is a real difficulty to have machines do mathematics on their own: work on Gödel's "theoretical" questions has to be complemented by sustained efforts to meet Turing's "practical" challenge. I take this to be one of the ultimate motivations for having machines find proofs in mathematics, i.e., proofs that reflect logical as well as mathematical understanding.

When focusing on proof search in mathematics it may be possible to use and expand logical work, but also draw on experience of actual mathematical practice. I distinguish two important features of the latter: (i) the refined conceptual organization internal to a given part of mathematics, and (ii) the introduction of new abstract concepts that cut across different areas of mathematics.[11] Logical formality per se does not facilitate the finding of arguments from given assumptions to a particular conclusion. However, strategic considerations can be formulated (for natural deduction calculi) and help to bridge the gap between assumptions and conclusion, suggesting at least a very rough structure of arguments. These logical structures depend solely on the syntactic form of assumptions and conclusion; they provide a seemingly modest, but in fact very important starting-point for strategies that promote automated proof search in mathematics.

Here is a pregnant general statement that appeals primarily to the first feature of mathematical practice mentioned above: *Proofs provide explanations of what they prove by putting their conclusion in a context that shows them to be correct*.[12] The deductive organization of parts of mathematics is the classical methodology for specifying such contexts. "Leading mathematical ideas" have to be found, proofs have to be planned: I take this to be the axiomatic method turned dynamic and local.[13] This requires undoubtedly the introduction of heuristics that reflect a deep understanding of the underlying mathematical subject matter. The broad and operationally significant claim is, that we have succeeded in isolating the leading ideas for a part of mathematics, if that part can be developed by machine – automatically, efficiently, and in a way that is, furthermore, easily accessible to human mathematicians.[14] This feature can undoubtedly serve as a springboard for the second feature I mentioned earlier, one that is so characteristic of the developments in modern mathematics, beginning in the second half of the 19th century: the introduction of abstract notions that do not have an intended interpretation, but rather are applicable in may different contexts. (Notions like group, field, topological space.) The above general statement concerning mathematical explanation can now be directly extended to incorporate also the second feature of actual mathematical experience. Turing might ask, can machines be educated to make such reflective moves on their own?

It remains a deep challenge to understand better the very nature of reasoning. A marvelous place to start is mathematics; where else do we find such a rich body of systematically and rigorously organized knowledge that is structured for intelligibility and discovery? The appropriate logical framework should undoubtedly include a *structure theory of (mathematical) proofs*. Such an extension of mathematical logic and in particular of proof theory interacts directly with a sophisticated automated search for humanly intelligible proofs. How far can this be pushed? What kind of broader leading ideas will emerge? What deeper understanding of basic operations of the mind will be gained? – We'll hopefully find out and, thus, uncover with strategic ingenuity part of Turing's residue and capture also part of what Gödel considered as "humanly effective", but not mechanical – "by asking the right questions on the basis of a mechanical procedure".

4. (Supra-) Mechanical devices

Turing machines codify directly the most basic operations of a human computor and can be realized as physical devices, up to a point. Gödel took for granted that finite machines just are (computationally equivalent to) Turing machines. Similarly, Church in [1937a, b] claimed that Turing machines are obtained by natural restrictions from machines occupying a finite space and with working parts of finite size; he viewed the restrictions "of such a nature as obviously to cause no loss of generality". In contrast to Gödel and Church, Gandy did not take this equivalence for granted and certainly not as being supported by Turing's analysis. He characterized machines informally as discrete mechanical devices that can carry out massively parallel operations. Mathematically Gandy machines are discrete dynamical systems satisfying boundedness and locality conditions that are physically motivated; they are provably not more powerful than Turing machines. Clearly one may ask: Are there plausible broader concepts of computations for physical systems? If there are systems that carry out supra-Turing processes they cannot satisfy the physical restrictions motivating the boundedness and locality conditions for Gandy machines. That is, such systems must violate either the upper bound on signal propagation or the lower bound on the size of distinguishable atomic components.[15] In *Paper machines*, Mundici and I diagnosed matters concerning physical processes in the following way. Every mathematical model of physical processes comes with at least two problems, "How accurately does the model capture physical reality, and how efficiently can the model be used to make predictions?" What is distinctive about modern developments is the fact that, on the one hand, computer simulations have led to an emphasis on algorithmic aspects of scientific laws and, on the other hand, physical systems are being considered as computational devices that process information much as computers do. It seems, ironically, that the mathematical inquiry into paper machines has led to the point where (effective) mathematical descriptions of nature and (natural) computations for mathematical problems coincide.

How could we have physical processes that allow *supra-Turing computations*? If harnessed in a machine, we would have a genuinely supra-mechanical device. However, we want to be able to *effectively determine* mathematical states from other such states – that "parallel" physical states, i.e., we want to make predictions and do that in a sharply intersubjective way. If that would not be the case, why would we want to call such a physical process a computation and not just an oracle? Wouldn't that undermine the radical intersubjectivity computations were to insure? There are many fascinating open issues concerning mental and physical processes that may or may not have adequate computational models. They are empirical, broadly conceptual, mathematical and, indeed, richly interdisciplinary.

Notes

1 This is [Gödel 1939]. As to the character of these lectures, see Dawson (1997), p. 135.

2 In Turing's [1936] general recursive functions are not mentioned. Turing established in an Appendix to his paper the equivalence of his notion with λ-definability. As Church and Kleene had already proved the equivalence of λ-definability and general recursiveness, "Turing's Theorem" is thus established for Turing computability.

3 This is obviously in contrast to the view he had in 1934 when defining general recursive functions; cf. [Sieg 2006].

4 That standpoint is formulated at the very end of the lecture as follows: p. 38 (CW III, 322/3): "Thereby [i.e., the Platonistic view] I mean the view that mathematics describes a non-sensual reality, which exists independently both of the acts and

[[of]] the dispositions of the human mind and is only perceived, and probably perceived very incompletely, by the human mind."

5. This is in contrast to the case of "objective" mathematics, the system of all true mathematical propositions, for which one cannot have a "well-defined system of correct axioms" (given by a finite rule) that comprises all of it. – In [Wang 1974], pp. 324-326, Gödel's position on these issues is (uncritically) discussed. The disjunction is presented as one of "two most interesting rigorously proved results about minds and machines" and is formulated as follows: "Either the human mind surpasses all machines (to be more precise: it can decide more number theoretic questions than any machine) or else there exist number theoretical questions undecidable for the human mind."

6 Boolos'*Introductory Note* to the Gibbs Lecture, in particular Section 3, gives a different perspective on difficulties in the argument.

7 This does not follow just from the fact that for every Turing machine that lists evident axioms, there is another axiom evident to the human mind not included in the list. Turing had tried already in his 1939 paper, *Ordinal Logics*, to overcome the incompleteness results by strengthening theories systematically. He added consistency statements (or reflection principles) and iterated this step along constructive ordinals; Feferman perfected that line of investigation, cf. his [1988]. Such a procedure was also envisioned in [Gödel, 1946], pp. 1–2.

8 Cf. also note 13 of the Gibbs Lecture and the remark on p. 17.

9 Cf. Parsons' informative remarks in the *Introductory Note* to [Gödel, 1946] on p. 148.

10 [Gödel 1972d] may be viewed, Gödel mentions, as a note to the word "mathematics" in the sentence, "Note that the results mentioned in this postscript do not establish any bounds of the powers of human reason, but rather for the potentialities of pure formalism in mathematics." This sentence appears in the 1964 Postscriptum to the Princeton Lectures Gödel gave in 1934; *Collected Works I*, pp. 369–371. He states in that Postscriptum also that there may be "finite non-mechanical procedures" and emphasizes, as he does in many other contexts, that such procedures would "involve the use of abstract terms on the basis of their meaning". (Note 36 on p. 370 of *Collected Works I*) Other contexts are found in volume III of the *Collected Works*, for example, the Gibbs Lecture (p. 318 and note 27 on that very page) and a related passage in "Is mathematics syntax of language?" (p. 344 and note 24) These are systematically connected to Gödel's reflections surrounding (the translation of) his Dialectica paper [1958] and [1972]. A thorough discussion of these issues cannot be given here; but as to my perspective on the basic difficulties, see the discussion in Section 4 of my paper "Beyond Hilbert's Reach?".

11 That is, it seems to me, still far removed from the introduction of "abstract terms" in Gödel's discussions. They are also, if not mainly, concerned with the introduction of new mathematical objects. Cf. note 10.

12 That is a classical observation; just recall the dual experiences of Hobbes and Newton with the Pythagorean Theorem, when reading Book 1 of Euclid's *Elements*.

13 Saunders MacLane articulated such a perspective and pursued matters to a certain extent in his Göttingen dissertation. See his papers [1935] and [1979].

[14] To mention one example: in an abstract setting, where representability and derivability conditions, but also instances of the diagonal lemma are taken for granted as axioms, Gödel's proofs can be found fully automatically; see Sieg and Field, (2005). The leading ideas used to extend the basic logical strategies are very natural; they allow moving between object and meta-theoretic considerations via provability elimination and introduction rules.

[15] My [2002a, b] discuss the methodological issues and Gandy machines in detail. For a general and informative discussion concerning "hypercomputation", see Davis's paper [2004]. A specific case of "computations" beyond the Turing limit is presented through Siegelmann's (1997) ANNs (artificial neural nets): they perform hypercomputations only if arbitrary reals are admitted as weights. Finally, there is the complex case of quantum computations: they allow a significant speed-up for example in Shore's algorithm, but if I understand matters correctly the current versions don't go beyond the Turing limit.

References

Church A (1937a) Review of [Turing 1936] Journal of Symbolic Logic 2(1): 40–41

Church A (1937b) Review of [Post 1936] Journal of Symbolic Logic 2(1): 42–43

Davis M (2004) The myth of hypercomputation. In: Teuscher C (ed) Alan Turing: Life and Legacy of a Great Thinker, pp. 195–211. Springer

Dawson J (1997) Logical Dilemmas. A.K. Peters, Wellesley, Massachusetts

Feferman S (1988) Turing in the land of O(z). In: Herken (ed) 1988, pp. 113–147

Gandy R (1980) Church's Thesis and principles for mechanisms. In: Barwise J, Keisler HJ and Kunen K (eds) The Kleene Symposium, pp. 123–148. North-Holland Publishing Company, Amsterdam

Gödel, K (1933) The present situation in the foundations of mathematics. Collected Works III, 45–53

Gödel K (1934) On undecidable propositions of formal mathematical systems. Collected Works I: 346–371

Gödel, K (193?) Undecidable Diophantine propositions. Collected Works III: 164–175

Gödel, K (1946) Remarks before the Princeton bicentennial conference on problems in mathematics. Collected Works II: 150–153

Gödel, K (1951) Some basic theorems on the foundations of mathematics and their implications. Collected Works III: 304–323

Gödel, K (1964) Postscriptum for [1934]. Collected Works I: 369–371

Gödel, K (1972a) Some remarks on the undecidability results. Collected Works II: 305–306

Gödel, K (1972b) The best and most general version of the unprovability of consistency in the same system; First of the notes [1972a]

Gödel, K (1972c) Another version of the first undecidability result; Second of the notes [1972a]

Gödel, K (1972d) A philosophical error in Turing's work; Third of the notes [1972a]

Gödel, K (1974) Note in [Wang 1974], 325–326

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

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

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

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

MacLane, S (1934) Abgekürzte Beweise im Logikkalkul. Ianaugural-Dissertation, Göttingen

MacLane, S (1935a) Logical analysis of mathematical structure; The Monist 45: 118–130

MacLane, S (1979a) late return to a thesis in logic. In: Kaplansky I (ed) Saunders MacLane – Selected Papers, pp. 63-66. Springer-Verlag

Mundici D and Sieg W (1995) Paper machines. Philosophia Mathematica 3: 5–30

Post E (1936) Finite combinatory processes Formulation I. Journal of Symbolic Logic 1: 103–105

Sieg, W (1994) Mechanical procedures and mathematical experience. In: George A (ed) Mathematics and Mind, pp. 71–117. Oxford University Press

Sieg W (2002a) Calculations by man and machine: conceptual analysis Lecture Notes in Logic 15: 390–409

Sieg W (2002b) Beyond Hilbert's reach?. In: Malalment DB (eds) Reading Natural Philosophy, pp. 363–405. Open Court, Chicago

Sieg, W (2006) Gödel on computability. Philosophia Mathematica 14: 189–207

Sieg W and Field C (2005) Automated search for Gödel's proofs. Annals of Pure and Applied Logic 133: 319–338

Siegelmann, HT (1997) *Neural Networks and Analog Computation – Beyond the Turing Limit*. Birkhäuser

Turing, A (1936) On computable numbers, with an application to the Entscheidungsproblem. Proceedings of the London Mathematical Society, Series 2 (42), 230–265; reprinted in Davis 1965

Turing, A (1939) Systems of logic based on ordinals. Proceedings of the London Mathematical Society, Series 2 45, 161–228; reprinted in Davis 1965

Turing, A (1947) Lecture to the London Mathematical Society on 20 February 1947. In: Ince DC (ed) Collected Works of A.M. Turing – Mechanical Intelligence, pp. 87–105. North Holland, 1992

Turing, A (1948) Intelligent machinery. In: Ince DC (ed) Collected Works of A.M. Turing – Mechanical Intelligence, pp. 107–127. North Holland, 1992

Turing A (1950) Computing machinery and intelligence Mind 59: 433–460

Turing A (1954) Solvable and unsolvable problems Science News 31: 7–23

Wang H (1974) From Mathematics to Philosophy. Routledge & Kegan Paul, London