Chapter 2
Wittgenstein's Diagonal Argument: A Variation on Cantor and Turing[1]

Juliet Floyd

2.1 Introduction

On 30 July 1947 Wittgenstein began writing what I call in what follows his "1947 remark"[2]:

> Turing's 'machines'. These machines are humans who calculate. And one might express what he says also in the form of games. And the interesting games would be such as brought one via certain rules to nonsensical instructions. I am thinking of games like the "racing game".[3] One has received the order "Go on in the same way" when this makes no sense,

[1]Thanks are due to Per Martin-Löf and the organizers of the Swedish Collegium for Advanced Studies (SCAS) conference in his honor in Uppsala, May 2009. The audience, especially the editors of the present volume, created a stimulating occasion without which this essay would not have been written. Helpful remarks were given to me there by Göran Sundholm, Sören Stenlund, Anders Öberg, Wilfried Sieg, Kim Solin, Simo Säätelä, and Gisela Bengtsson. My understanding of the significance of Wittgenstein's Diagonal Argument was enhanced during my stay as a fellow 2009–2010 at the Lichtenberg-Kolleg, Georg August Universität Göttingen, especially in conversations with Felix Mühlhölzer and Akihiro Kanamori. Wolfgang Kienzler offered helpful comments before and during my presentation of some of these ideas at the Collegium Philosophicum, Friedrich Schiller Universität, Jena, April 2010. The final draft was much improved in light of comments provided by Sten Lindström, Sören Stenlund and William Tait.

[2]This part of the remark is printed as §1096 of Wittgenstein et al. (1980), hereafter abbreviated RPP I. See footnote 21 below for the manuscript contexts.

[3]I have not been able to identify with certainty what this game is. I presume that Wittgenstein is thinking of a board game in which cards are drawn, or dice thrown, ~~and pieces are moved in a kind of race~~. See below for specifics.
or knobs and cranks turned so as to move pieces in a simulated horse race.

J. Floyd (✉)
Department of Philosophy, Boston University, Boston, USA
e-mail: jfloyd@bu.edu

P. Dybjer et al. (eds.), *Epistemology versus Ontology*, Logic, Epistemology, and the Unity of Science 27, DOI 10.1007/978-94-007-4435-6_2,

> say because one has got into a circle. For that order makes sense only in certain positions. (Watson.[4])

The most sustained interpretation of this remark was offered some time ago by Stewart Shanker, who argued (1987, 1998) that its primary focus is philosophy of mind, and specifically the behaviorism embedded within the cognitivist revolution that Turing spawned. Shanker maintains that Wittgenstein is committed to denying Church's thesis, viz., that all (humanly) computable functions are Turing computable. In what follows I shall leave aside Church's thesis: too many issues about it arise for me to profitably canvas the associated problems here, and Shanker is quite clear that he is reconstructing the implications of Wittgenstein's remark and not its specific, local, content. Nor shall I contest the idea – forwarded not only by Shanker, but also by Kripke and Wright (among many others) – that there are fundamental criticisms of functionalism, reductionism, and computationalism about the mind that may be drawn out of Wittgenstein's later thought.[5] Shanker is surely right to have stressed the broad context of Wittgenstein's 1947 remark, which is a lengthy exploration of psychological concepts. And Wittgenstein did investigate the sense in which any model of computation such as Turing's could be said to give us a description of how humans (or human brains or all possible computing machines) actually work, when calculating. Turing offers, not a definition of "state of mind", but what Wittgenstein thought of as a "language game", a simplified model or snapshot of a portion of human activity in language, an object of comparison forwarded for a specific analytic purpose.

Turing sent Wittgenstein an offprint of his famous (1937a) paper "On Computable Numbers, With an Application to the *Entscheidungsproblem*".[6] It contains terminology of "processes", "motions" "findings" "verdicts", and so on. This talk had the potential for conflating an analysis of Hilbert's *Entscheidungsproblem* and the purely logical notion of possibility encoded in a formal system with a description of human computation. As Shanker argues, such conflations without due attention to the idealizations involved were of concern to Wittgenstein. However, as I am confident Shanker would allow, there are other issues at stake in Wittgenstein's remark than philosophy of mind or Church's thesis. Turing could not have given a negative resolution of the *Entscheidungsproblem* in his paper if his proof had turned on a specific thesis in philosophy of mind. Thus it is of importance to stress that in his 1947 remark Wittgenstein was directing his attention, not only to psychological concepts, but to problems in the foundations of logic and mathematics, and to one problem in particular that had long occupied him, viz., the *Entscheidungsproblem*.

In the above quoted 1947 remark Wittgenstein is indeed alluding to Turing's famous (1937a) paper. He discussed its contents and then recent undecidability results with (Alister) Watson in the summer of 1937, when Turing returned to

[4] Alister Watson discussed the Cantor diagonal argument with Turing in 1935 and introduced Wittgenstein to Turing. The three had a discussion of incompleteness results in the summer of 1937 that led to Watson (1938). See Hodges (1983), pp. 109, 136 and footnote 7 below.

[5] Kripke (1982), Wright (2001), Chapter 7. See also Gefwert (1998).

[6] See Hodges (1983), p. 136. Cf. Turing (1937c).

Cambridge between years at Princeton.[7] Since Wittgenstein had given an early formulation of the problem of a decision procedure for all of logic,[8] it is likely that Turing's (negative) resolution of the *Entscheidungsproblem* was of special interest to him. These discussions preceded and, I believe, significantly stimulated and shaped Wittgenstein's focused work on the foundations of mathematics in the period 1940–1944, especially his preoccupation with the idea that mathematics might be conceived to be wholly *experimental* in nature: an idea he associated with Turing. Moreover, so far as we know Wittgenstein never read Turing's "Computing Machinery and Intelligence" Turing (1950), the paper that injected the AI program, and Church's thesis, into philosophy of mind.[9] Instead, in 1947 Wittgenstein was recalling discussions he had had with Watson and Turing in 1937–1939 concerning problems in the foundations of mathematics.

In general, therefore, I agree with Sieg's interpretation of Turing's model in relation to Wittgenstein's 1947 remark. Sieg cites it while arguing, both that Turing was not the naive mechanist he is often taken to be, and also that Wittgenstein picked up on a feature of Turing's analysis that was indeed crucial for resolving the *Entscheidungsproblem*.[10] What was wanted to resolve Hilbert's famous problem was an analysis of the notion of a "definite method" in the relevant sense: a "mechanical procedure" that can be carried out by human beings, i.e., computers, with only limited cognitive steps (recognizing a symbolic configuration, seeing that one of finitely many rules applies, shifting attention stepwise to a new symbolic configuration, and so on).[11] An analysis like Turing's that could connect the notion with (certain limited aspects of possible) *human* cognitive activity was, then, precisely what was wanted. The human aspect enters at one pivotal point, when Turing claims that a human computer can recognize only a bounded number of different discrete configurations "at a glance", or "immediately".[12] Sieg's conceptual analysis explains what makes Turing's analysis of computability more vivid, more pertinent and (to use Gödel's word) more epistemologically satisfying than Church's or

[7] Hodges (1983), p. 135; cf. Floyd (2001).

[8] In a letter to Russell of later November or early December 1913; see R. 23 in McGuinness (2008) or in Wittgenstein (2004). For a discussion of the history and the philosophical issues see Dreben and Floyd (1991).

[9] Malcolm queried by letter (3 November 1950, now lost) whether Wittgenstein had read "Computing Machinery and Intelligence", asking whether the whole thing was a "leg pull". Wittgenstein answered (1 December 1950) that "I haven't read it but I imagine it's <u>no</u> leg-pull". (Wittgenstein (2004), McGuinness (2008), p. 469).

[10] Sieg (1994), p. 91; Sieg (2008), p. 529.

[11] The *Entscheidungsproblem* asks, e.g., for an algorithm that will take as input a description of a formal language and a mathematical statement in the language and determine whether or not the statement is provable in the system (or: whether or not a first-order formula of the predicate calculus is or is not valid) in a finite number of steps. Turing 1937a offered a proof that there is no such algorithm, as had, albeit with a different proof, the earlier Church (1936).

[12] As Turing writes (1937a, p. 231), "the justification lies in the fact that the human memory is necessarily limited"; cf. §9 of the paper.

Gödel's extensionally equivalent demarcations of the class of recursive functions, though without subscribing to Gödel's and Church's own accounts of that epistemic advantage.[13]

It is often held (e.g., by Gödel[14]) that Turing's analogy with a human computer, drawing on the assumption that a (human) computer scans and works with only a finite number of symbols and/or states, involves strong metaphysical, epistemological and/or psychological assumptions that he intended to use to *justify* his analysis. From the perspective adopted here, this is not so. Turing's model only makes explicit certain characteristic features earmarking the concept that is being analyzed in the specific, Hilbertian context (that of a recognizeable *step within* a computation or a formal system, a "definite procedure" in the relevant sense). It is not a thesis in philosophy of mind or mathematics, but instead an assumption taken up in a spirit analogous to Wittgenstein's idea that a proof must be perspicuous (*Übersichtlich, Übersehbar*), i.e., something that a human being can take in, reproduce, write down, communicate, verify, and/or articulate *in some systematic way or other*.[15]

If we look carefully at the context of Wittgenstein's 1947 remark, we see that it is Turing's *argumentation* as such that he is considering, Turing's *use* of an abstract model of human activity to make a diagonal argument, and not any issue concerning the explanation or psychological description of human mental activity as such. This may be seen, not only by emphasizing, as Sieg does, that Turing's analysis requires no such general description, but also by noticing that immediately after this 1947 remark Wittgenstein frames a novel "variant" of Cantor's diagonal argument.

The purpose of this essay is to set forth what I shall hereafter call *Wittgenstein's Diagonal Argument*. Showing that it *is* a distinctive argument, that it is a *variant* of Cantor's and Turing's arguments, and that it *can* be used to make a proof are my primary aims here. Full analysis of the 1947 remarks' significance within the context of Wittgenstein's philosophy awaits another occasion, though in the final section I shall broach several interpretive issues.

As a contribution to the occasion of this volume, I dedicate my observations to Per Martin-Löf. He is a unique mathematician and philosopher in having used proof-theoretic semantics to frame a rigorous analysis of the notions of judgment and proposition at work in logic, and in his influential constructive type theory.[16] I like to think he would especially appreciate the kind of "variant" of the Cantor proof that Wittgenstein sketches.

[13] See Sieg (2006a, b). Compare Gandy (1988). On Gödel's attitude, see footnote 28 below.

[14] See the note Gödel added to his "Some remarks on the undecidability results" (1972a), in Gödel (1990), p. 304, and Webb (1990). Gödel (somewhat unfairly) accuses Turing of a "philosophical error" in failing to admit that "*mind, in its use, is not static, but constantly developing*", as if the appropriateness of Turing's analysis turns on denying that mental states might form a continuous series.

[15] Wittgenstein's notion of *perspicuousness* has received much attention. Two works which argue, as I would, that it does not involve a restrictive epistemological thesis or reductive anthropologism are Marion (2011) and Mühlhölzer (2010).

[16] See, e.g., Martin-Löf (1984, 1996).

In presenting Wittgenstein's Diagonal Argument I proceed as follows. First (2.1), I briefly rehearse the Halting Problem, informed by a well-known application of diagonal argumentation. While that argument itself does not, strictly speaking, appear in Turing's (1937a) paper, a closely related one does, at the beginning of its §8 (Sect. 2.2.2). However, Turing frames another, rather different argument immediately afterward, an argument that appeals to the notion of computation by machine in a more concrete way, through the construction of what I shall call a *Pointerless Machine* (Sect. 2.2.3). Next (3) I present Wittgenstein's Diagonal Argument, arguing that it derives from his reading of Turing's §8. And then (4) I present a "positive" version of Russell's paradox that is analogous to Wittgenstein's and Turing's arguments and which raises interesting questions of its own. Finally (5), I shall canvas a few of the philosophical and historical issues raised by these proofs.

2.2 Three Diagonal Arguments

2.2.1 The Halting Problem

Though it does not, strictly speaking, occur in Turing (1937a), the so-called "Halting Problem" is an accessible and well-known example of diagonal argumentation with which we shall begin.[17]

The totality of Turing machines in one variable can be enumerated. In his (1937a) Turing presented his machine model in terms of "skeleton tables" and associated with each particular machine a unique "description number" (**D.N.**), thus Gödelizing; nowadays it is usual to construe a Turing machine as a set of quadruples. In the modern construal, a Turing machine t has as its input-output behavior a partial function $f : N \to N$ as ollows: t is presented with an initial configuration that codes a natural number j according to a specified protocol, and t then proceeds through its instructions. In the event that t goes into a specified halt state with a configuration that codes a natural number k according to protocol, then $f(j) = k$ and f is said to *converge at* j, written "$f(j)\downarrow$". Otherwise, f is said to *diverge at* j, written "$f(j)\uparrow$". In general, f is partial because of the latter possibility.

Enumerating Turing machines as t_i, we have corresponding partial functions $f_i : N \to N$, and a partial function $g : N \to N$ is said to be *computable* if it is an f_i. The set of Turing machines is thus definable and enumerable, but represents the set of *partial* computable functions. Because of this, it is not possible to diagonalize out

[17] Turing's argument in 1937a in §8 is not formulated as a halting problem; this was done later, probably by Martin Davis in a lecture of 1952. For further details on historical priority, see *http://en.wikipedia.org/wiki/Halting_problem#History_of_the_halting_problem* and Copeland (2004), p. 40 n 61.

of the list of computable functions, as it is from a list of, e.g., real numbers in binary representation (as in Cantor's 1891 argument). In other words, the altered diagonal sequence, though it may be defined as a function, is not a computable function in the Turing sense.

The last idea is what is to be proved. (Once the equivalence to formal systems is made explicit, this result yields Turing's negative resolution of the *Entscheidungsproblem*.)

To fix ideas, consider a binary array, conceived as indicating *via* "↑" that Turing machine t_i diverges on input j, and *via* "↓" that it converges on input j. Each t_i computes a partial function $f_i : N \to N$ on the natural numbers, construed as a binary sequence.

$$
\begin{array}{llllll}
t_1 & \uparrow & \uparrow & \downarrow & \downarrow & \uparrow & \ldots \\
t_2 & \downarrow & \downarrow & \uparrow & \uparrow & \downarrow & \ldots \\
t_3 & \downarrow & \downarrow & \downarrow & \uparrow & \uparrow & \ldots \\
t_4 & \uparrow & \uparrow & \uparrow & \uparrow & \downarrow & \ldots \\
t_5 & \downarrow & \uparrow & \downarrow & \uparrow & \downarrow & \ldots \\
\ldots & & & & & &
\end{array}
$$

Cantor's method of diagonal argument applies as follows. As Turing showed in §6 of his (1937a), there is a universal Turing machine UT_1. It corresponds to a partial function $f(i, j)$ of two variables, yielding the output for t_i on input j, thereby simulating the input-output behavior of every t_i on the list. Now we construct D, the Diagonal Machine, with corresponding one-variable function which on input i computes UT_1 (i, i). D is well-defined, and corresponds to a well-defined (computable, partial) function.

We suppose now that we can define a "Contrary" Turing machine C that reverses the input-output behavior of D as follows: C, with the initial configuration coding j, first proceeds through the computation of $\mathrm{D}(j)$ and then follows this rule:

$$
\begin{aligned}
(*) \quad & \text{If } \mathrm{D}(j)\downarrow, \text{then } \mathrm{C}(j) = \uparrow; \\
& \text{If } \mathrm{D}(j)\uparrow, \text{then } \mathrm{C}(j) = 1
\end{aligned}
$$

In other words, if $\mathrm{D}(j)$ converges then proceed to instructions that never halt, and if $\mathrm{D}(j)$ diverges, then output the code for 1 and enter the halting state.

But there is a contradiction with assuming that this rule can be followed, or implemented by a machine that is somewhere on the list of Turing machines. Why? If C were a Turing machine, it would be t_k for some k. Then consider t_k on input k. By rule (*), *if* t_k converges on k, then it diverges on k; but if it diverges on k, then it converges on k. So t_k converges on k if and only if it diverges on k. This contradiction indicates that our supposition was false.

Rule (*) assumes Halting Knowledge, i.e., that machine C can reach a conclusion about the behavior of D on any input j, and follow rule (*). But to have such knowledge requires going through all the (possibly) infinitely many steps of the D machine. And that is not itself a procedure that we can express by a rule for a one-variable Turing machine. In other words Halting Knowledge is not Turing computable.

Classical philosophical issues about negation in infinite contexts – the worry about what it means to treat a completed totality of steps as just another step – emerge. Turing himself acknowledged as much. In (1937b) he published some corrections to his (1937a) paper. The first fixed a flaw in a definition pointed out by Bernays, thereby narrowing a reduction class he had framed for the Decision Problem. The second, also stimulated by Bernays, made his analysis more general, showing that his definition of "computable number" serves independently of a choice of logic. Turing wrote to Bernays (22 May 1937) that when he wrote the original paper of (1937a), "I was treating 'computable' too much as one might treat 'algebraic', with wholesale use of the principle of excluded middle. Even if this sounds harmless, it would be as well to have it otherwise" (1937d). In his (1937b) correction he modified the means by which computable numbers are associated with computable sequences, citing Brouwer's notion of an overlapping choice sequence, as Bernays suggested he do.[18] This avoids what Turing calls a "disagreeable situation" arising in his initial arguments: although the law of the excluded middle may be invoked to show that a Turing machine *exists* that will compute a function (e.g., the Euler constant), we may not have the means to *describe* any such machine (Turing 1937b, p. 546). The price of Turing's generalization is that real numbers no longer receive unique representations by means of sequences of figures. The payoff is that his definition's applicability no longer depends upon invoking the law of the excluded middle in infinite contexts. The loss, he explains, "is of little theoretical importance, since the [description numbers of Turing machines] are not unique in any case" and the "totality of computable numbers [remains] unaltered" (Turing 1937b, p. 546). In other words, his characterization of the computable numbers is robust with respect to its representation by this or that formal system, this or that choice of logic, or any specific analysis of what a real number really *is*. Today we would say that the class of computable numbers is *absolute* with respect to its representation in this or that formal system.[19] And this too is connected with

[18] Cf. Bernays to Turing 24 September 1937 (Turing 1937d). The corrections using Brouwer's notion of an overlapping sequence are explained in Petzold (2008), pp. 310ff. Petzold conjectures that conversations with Church at Princeton (or with Weyl) may have stimulated Turing's interest in recasting his proof, though he suspects that "Turing's work and his conclusions are so unusual that … he wasn't working within *anyone's* prescribed philosophical view of mathematics" (2008, p. 308). I agree. But in terms of possible influences on Turing, Bernays should be mentioned, and Wittgenstein should be added to the mix. The idea of expressing a rule as a table-cum-calculating device read off by a human being was prevalent in Wittgenstein's philosophy from the beginning, forming part of the distinctive flavor in the air of Cambridge in the early 1930s, and discussed explicitly in his Wittgenstein (1980).

[19] Gödel, concerned with his own notion of general recursiveness when formulating the absoluteness property (in 1936) later noted the importance of this notion in connection with the independence of Turing's analysis from any particular choice of formalism. He remarked that with

the anthropomorphic quality of his model. For it is not part of the ordinary activity of a human computer, or the general concept of a person working *within* a formal system of the kind involved, to take a stance on the law of the excluded middle.

2.2.2 Turing's First Argument

Turing's (1937a) definitions are as follows. A ***circle-free machine*** is one that, placed in a particular initial configuration, prints an infinite sequence of 0's and 1's (blank spaces and other symbols are regarded by Turing as aids to memory, analogous to scratch paper; only these scratch symbols are ever erased). A ***circular machine*** fails to do this, never writing down more than a finite number of 0s and 1s. (Unlike a contemporary Turing Machine, then, for Turing the *satisfactory* machines print out infinite sequences of 0's and 1's, whereas the *unsatisfactory* ones "get stuck" (see footnote 26).) A ***computable number*** is a real number differing by an integer from a number computed by a circle-free machine (i.e., its decimal (binary) expansion will, in the non-integer part, coincide with an infinite series of 0's and 1's printed by some circle-free machine); this is a real number whose decimal (binary) expression is said to be ***calculable by finite means***. A ***computable sequence*** is one that can be represented (computed) by a circle-free machine.

The First Argument begins §8. Turing draws a distinction between the application of Cantor's original diagonal argument and the version of it he will apply in his paper:

> It may be thought that arguments which prove that the real numbers are not enumerable would also prove that the computable numbers and sequences cannot be enumerable. [n. Cf. Hobson, *Theory of functions of a real variable* (2^{nd} ed., 1921), 87, 88]. It might, for instance, be thought that the limit of a sequence of computable numbers must be computable. This is clearly only true if the sequence of computable numbers is defined by some rule.
> Or we might apply the diagonal process. "If the computable sequences are enumerable, let α_n be the n-th computable sequence, and let $\phi_n(m)$ be the m-th figure in α_n. Let β be the sequence with $1-\phi_n(n)$ as its n-th figure. Since β is computable, there exists a number K such that $1-\phi_n(n) = \phi_K(n)$ all n. Putting $n = K$, we have $1 = 2\phi_K(K)$, $i.e.$ 1 is even. This is impossible. The computable sequences are therefore not enumerable".

The argument Turing offers in quotation marks purports to show that the computable numbers are not enumerable in just the same way as the real numbers are not, according to Cantor's original diagonal argument. (We should notice that

---

Turing's analysis of computability "one has for the first time succeeded in giving an absolute definition of an interesting epistemological notion, i.e., one not depending on the formalism chosen" (Gödel here means a formal system of the relevant (recursively axiomatizeable, finitary language) kind). See Gödel's 1946 "Remarks before the Princeton bicentennial conference on problems in mathematics", in Gödel (1990), pp. 150–153; Compare his Postscriptum to his 1936a essay "On the Length of Proofs", Ibid., p. 399. See footnote 28, and Sieg (2006a, b), especially pp. 472ff.

its structure is reminiscent of the Contrary Machine, framed in the Halting Problem above, which switches one kind of binary digit to another, "negating" all the steps along the diagonal.) However, Turing responds:

> The fallacy in this argument lies in the assumption that β is computable. It would be true if we could enumerate the computable sequences by finite means [JF: i.e., by means of a circle-free machine], but the problem of enumerating computable sequences is equivalent to the problem of finding out whether a given number is the D.N of a circle-free machine, and we have no general process for doing this in a finite number of steps. In fact, by applying the diagonal process argument correctly, we can show that there cannot be any such general process.

This "correct" application of the diagonal argument is, globally, a *semantic* one in the computer scientist's sense: it deals with sequences (e.g. β) and the nature of their possible characterizations. The "fallacy" in thinking that Cantor's diagonal argument *can* apply to show that the computable numbers are not enumerable (i.e., in the original, Cantorian sense of enumerable as "countable") is that we will, as it turns out, be able to reject the claim that the sequence β is computable. So there is no diagonalizing out. The assumption that α_n, the enumeration of computable sequences, is enumerable *by finite means* is false. Turing's First Argument rejects that claim (much as in the Halting Argument above) by producing the contradiction he describes: it follows from treating the problem of enumerating all the computable sequences by finite means (i.e., by a circle-free machine) as "equivalent" to the problem of finding a general process for determining whether a given arbitrary number is or is not the description number of a circle-free machine. This, Turing writes – initially without argument – we cannot carry out in every case in a finite number of steps.

However, Turing immediately writes that this First Argument, "though perfectly sound", has a "disadvantage", namely, it may nevertheless "leave the reader with a feeling that 'there must be something wrong"'. Turing has remained so far little more than intuitive about our inability to construct a circle-free machine that will determine whether or not a number is the description number of a circle-free machine, and he has not actually shown how to reduce the original problem to that one. At best he has leaned on the idea that an infinite tape cannot be gone through in a finite number of steps. While this is fine so far as it goes, Turing asks for something else, something more rigorous.

2.2.3 The Argument from the Pointerless Machine

Turing immediately offers a second argument, one which, as he says, "gives a certain insight into the significance of the idea "circle-free"". I shall call it the *Argument from the Pointerless Machine* to indicate a connection with Wittgenstein's idea of logic as comprised, at least in part, of tautologies, i.e., apparently sensical sentences which are, upon further reflection, *sinnlos*, directionless, like two vectors which when added yield nothing but a directionless point with "zero" directional information.[20] Since Turing's is the first in print ever to *construct* a machine model to argue over computability in principle, it is of great historic importance, and so worth rehearsing in its own right. More importantly for my purposes here, *it* is the argument that Wittgenstein's 1947 diagonal argument phrased in terms of games.

Turing's second argument is intended to isolate more perspicuously the difficulty indicated in his First Argument. It works by considering how to define a machine $\mathcal{H}$, using an enumeration of all Turing machines, to directly compute a certain sequence, β', whose digits are drawn from the $\phi_n(n)$ along the diagonal sequence issuing from the enumeration of all computable sequences α_n. Recall from 1.2 above that α_n is the nth computable sequence in the enumeration of computable sequences (i.e., those sequences computable by a circle-free machine); $\phi_n(m)$ is the mth figure in α_n. β, used in the First Argument, is the "contrary" sequence consisting of a series of 0's and 1's issuing from a switch of 0 to 1 and vice versa along the diagonal sequence, $\phi_n(n)$. By contrast β' is the sequence whose nth figure is the output of the *nth* circle-free machine on input n: it corresponds to $\phi_n(n)$, which we may think of as the *positive* diagonal sequence. Its construction will make clear how it is the way in which one conceives of the enumeration of α_n (by finite means or not by finite means) that matters.

The Turing machines may be enumerated, for each has a "standard" description number k. Now suppose that there is a definite process for deciding whether an arbitrary number is that of a circle-free machine, i.e., that there is a machine $\mathcal{D}$ which, given the standard description number k of an arbitrary Turing machine $\mathcal{M}$, will test to see whether k is the number of a circular machine or not. If $\mathcal{M}$ is circular, $\mathcal{D}$ outputs on input k "u" (for "unsatisfactory"), and if $\mathcal{M}$ is circle-free, $\mathcal{D}$ outputs on k "s" (for "satisfactory"). $\mathcal{D}$ enumerates α_n by finite means. Combining $\mathcal{D}$ with the universal machine $\mathcal{U}$, we may construct a machine $\mathcal{H}$. $\mathcal{H}$ is designed to compute the sequence β'. But it turns out to be (what I call) a *Pointerless Machine*, as we may see from its characterization.

$\mathcal{H}$ proceeds as follows to compute β'. Its motion is divided into sections. In the first $N-1$ sections the integers $1, 2, \ldots N-1$ have been tested by $\mathcal{D}$. A certain number of these, say $R(N-1)$, have been marked "s", i.e., are description numbers of circle-free machines. In the Nth section the machine $\mathcal{D}$ tests the number N. If N is satisfactory, then $R(N) = 1+R(N-1)$ and the first $R(N)$ figures of the sequence whose description number is N are calculated. $\mathcal{H}$ writes down the $R(N)$th figure of this sequence. This figure will be a figure of β', for it is the output on n of the nth circle-free Turing machine in the enumeration of α_n by finite means that $\mathcal{D}$ is assumed to provide. Otherwise, if N is not satisfactory, then $R(N) = R(N-1)$ and the machine goes on to the $(N+1)$th section of its motion.

$\mathcal{H}$ is circle-free, by the assumption that $\mathcal{D}$ exists. Now let K be the D.N. of $\mathcal{H}$. What does $\mathcal{H}$ do on input K? Since K is the description number of $\mathcal{H}$, and $\mathcal{H}$ is circle-free, the verdict delivered by $\mathcal{D}$ cannot be "u". But the verdict also cannot be

[20] Compare the discussion in Dreben and Floyd (1991).

"s". For if it were, $\mathcal{H}$ would write down as the Kth digit of β' the Kth digit of the sequence computed by the Kth circle-free machine in α_n, namely by $\mathcal{H}$ itself. But the instruction for $\mathcal{H}$ on input K would be "calculate the first R(K) $= R(K-1)+1$ figures computed by the machine with description number K (that is, $\mathcal{H}$) and write down the $R(K)$th". The computation of the first $R(K)-1$ figures would be carried out without trouble. But the instructions for calculating the $R(K)$th figure would amount to "calculate the first $R(K)$ figures computed by $\mathcal{H}$ and write down the $R(K)$th". This digit "would never be found", as Turing says. For at the Kth step, it would be "circular", contrary to the verdict "s" and the original assumption that $\mathcal{D}$ exists ((1937a), p. 247). For its instructions at the Kth step amount to the "circular" order "do what you do".

The First Argument and Turing's Argument from the Pointerless Machine are constructive arguments in the classical sense: neither invokes the law of the excluded middle to reason about infinite objects. Moreover, as Turing's (1937b) correction showed, each may be set forth without presuming that standard machine descriptions are associated uniquely with real numbers, i.e., without presupposing the application of the law of excluded middle here either. Finally, both are, like the Halting argument, computability arguments: applications of the diagonal process in the context of Turing Machines.

But the Argument from the Pointerless Machine is more concrete than either the First Argument or the Halting Argument. And it is distinctive in not asking us to build the application of negation *into* the machine. The Pointerless Machine is one we construct, and then watch and trace out. The difficulty it points to is not that $\mathcal{H}$ gives rise to the possibility of constructing another contrary sequence which generates a contradiction. Instead, the argument is semantic in another way. The Pointerless Machine $\mathcal{H}$ gives rise to a command structure which is empty, tautologous, senseless. It produces, not a contradiction, but an empty circle, something like the order "Do what you are told to do". In the context at hand, this means that $\mathcal{H}$ cannot *do* anything. As Wittgenstein wrote in 1947, a command line "makes sense only in a certain positions".

2.3 Wittgenstein's Diagonal Argument

Immediately after his 1947 about Turing's "Machines" being "humans who calculate", Wittgenstein frames a diagonal argument of his own. This "expresses" Turing's argument "in the form of games", and should be counted as a part of that first remark.

> A variant of Cantor's diagonal proof:
> Let N=F (k, n) be the form of the law for the development of decimal fractions. N is the nth decimal place of the kth development. The diagonal law then is: N=F (n,n) = Def F′ (n).
> To prove that F′(n) cannot be one of the rules F (k, n).
> Assume it is the 100th. Then the formation rule of F′ (1) runs F (1, 1), of F′(2) F (2, 2) etc.

> But the rule for the formation of the 100th place of F′(n) will run F (100, 100); that is, it tells us only that the hundredth place is supposed to be equal to itself, and so for n = 100 it is *not* a rule.
> [I have namely always had the feeling that the Cantor proof did two things, while appearing to do only one.]
> The rule of the game runs "Do the same as. . . " – and in the special case it becomes "Do the same as you are doing".[21]

As we see, it is the Argument from the Pointerless Machine which Wittgenstein is translating into the vocabulary of language games in 1947. The reference to Turing and Watson is not extraneous. Moreover, the argument had a legacy. Wittgenstein was later credited by Kreisel with "a very neat way of putting the point" of Gödel's use of the diagonal argument to prove the incompleteness of arithmetic, in terms of the empty command, "Write what you write" (1950, p. 281n).[22]

Let us rehearse Wittgenstein's argument, to show that it constitutes a genuine proof. Wittgenstein begins by imagining a "form" of law for enumerating the "decimal fractions" (*Dezimalbrüchen*). We may presume that Wittgenstein has the rational numbers in mind, and in the case of the rational numbers, we know that such a law or rule (e.g., a listing) can exhaustively enumerate the totality. As Cantor showed, this is not true for the totality of real numbers. But the argumentation Wittgenstein sets forth applies whether the presentation of the list exhausts a set or not: all it assumes is that the presentation utilizes the expression of rules for the development of decimal fractions, a way of "developing" or writing them out that utilizes a countable mode of expression. Moreover, Wittgenstein's German speaks of decimal expansion development (*Entwicklung von Dezimalbrüchen*), and ordinarily in German this terminology (*Dezimalbruchentwicklung*) is taken to cover expansions of real numbers as well.[23] So Wittgenstein may well have had (a subset of) the real numbers, e.g., the computable real numbers, in mind as well. "Form" here assumes a space of *possible* representations: it means that we may imagine an enumeration in any way we like, and Wittgenstein does not restrict its presentation. He is articulating, in other words, a generalized *form* of diagonal argumentation. The argument is thus generally applicable, not only to decimal expansions, but to any purported listing or rule-governed expression of them; it does not rely on any particular notational device or preferred spatial arrangements of signs. In that sense, Wittgenstein's argument appeals to no picture, and it is not essentially

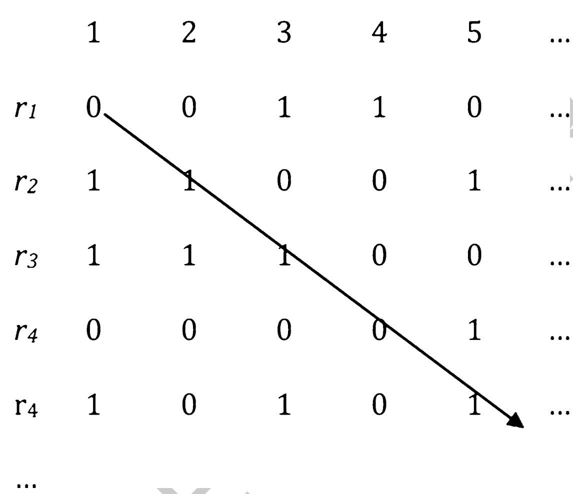
[21] Wittgenstein (1999), MS 135 p. 118; the square brackets indicate a passage later deleted when the remark [illegible] its way into Wittgenstein (1999), TS 229 §1764, published at RPP I §1097. (At *Zettel* §695 only this second remark [illegible] thereby separating it from the mention of Turing and Watson, hereafter Z). The argument as written here occurs here with "F" replacing the original "ϕ", following the typescript.

[22] See also Stenius (1970) for another general approach to the antinomies distinguishing between contradictory rules (that cannot be followed) and contradictory concepts (e.g., "the round square") that is explicitly based on a reading of Wittgenstein (in this case, the *Tractatus*).

[23] On the German see http://de.wikipedia.org/wiki/Dezimalbruch and http://de.wikipedia.org/wiki/Dezimalsystem#Dezimalbruchentwicklung.

diagrammatical or representational, though it may be diagrammed (and of course, insofar as it is a *logical* argument, its logic may be represented formally).[24] Like Turing's arguments, it is free of a direct tie to any particular formalism. Unlike Turing's arguments, it explicitly invokes the notion of a language-game and applies to (and presupposes) an everyday conception of the notions of *rules* and the *humans who follow them*.[25] Every line in the diagonal presentation above is conceived as an instruction or command, analogous to an order given to a human being.

To fix ideas, let us imagine an enumeration of decimal fractions in the unit interval in binary decimal form. Now let $\mathrm{N} = \mathrm{F}(n, n) = \mathrm{Def}\ \mathrm{F}'(n)$, whose graph is given by the diagonal line in the picture below.



The rule for computing $\mathrm{F}'(n)$ is clear: go down the diagonal of this list, picking off the value of r_n on input n. This rule appears to be perfectly comprehensible and is in *that* sense well defined. But it is not determined, in the sense that at each and every step we know what to do with it. Why? Wittgenstein's "variant" of Cantor's Diagonal argument – that is, of Turing's Argument from the Pointerless Machine – is this.

Assume that the function F′ is a development of one decimal fraction on the list, say, the 100th. The "rule for the formation" here, as Wittgenstein writes, "will run F(100, 100)." But this

[24]Recall that in his earlier 1938 remarks on the Cantor diagonal argument Wittgenstein was preoccupied with the idea that the proof might be thought to depend upon interpreting a particular kind of picture or diagram in a certain way. Wittgenstein (1978) Part II. There are many problematic parts of these remarks, and I hope to discuss them in another essay. For now I remark only that they are much earlier than the 1947 remarks I am discussing here, written down in the immediate wake of his summer 1937 discussions with Watson and Turing.

[25]Though Turing himself would write that "these [limitative] results, 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". Turing (1954), p. 23.

> . . . tells us only that the hundredth place is supposed to be equal to itself, and so for $n = 100$ it is not a rule. The rule of the game runs "Do the same as. . . " – and in the special case it becomes "Do the same as you are doing". (RPP I §1097, quoted above).

We have here an order that, like Turing's $\mathcal{H}$ machine, "has got into a circle" (cf. RPP I §1096, quoted above).[26] If one imagines drawing a card in a board game that says "Do what this card tells you to do", or "Do what you are doing", I think we have a fair everyday representation of the kind of phenomenon upon which Wittgenstein draws.

Wittgenstein's form of circle is, unlike Turing's, explicitly expressed in terms of a tautology. And Turing's argument is distinctive, upon reflection, precisely in producing a tautology of a certain sort. In a sense, Wittgenstein is *literalizing* Turing's model, bringing it back down to the everyday, and drawing out the anthropomorphic, command-aspect of Turing's metaphors.

I have said that Wittgenstein presents a genuine proof in his 1947 remark, and I have been willing to regard it as a "variant" of Cantor's diagonal argumentation. But a qualification is in order. The argument cannot survive construal in terms of a purely extensional way of thinking, and that way of thinking is required for the context in which Cantor's argument is forwarded, a context in which infinite objects are reasoned about and with. What is shown in Wittgenstein's argument is that on the assumption, F′(100) cannot be computed. But not because of the task being infinite. Instead, we are given a rule, that, as Wittgenstein writes, "is *not* a rule" in the same sense. There is, extensionally speaking, something which *is* the value of F(100,100) in itself, and it is either 0 or 1. But if we ask *which* digit it is, we end up with the answer, "F(100,100)", which doesn't say one way or the other what it is, because that will depend upon the assumption that this sequence is the value of F′(100) at 100. The diagonal rule, in other words, cannot be applied at this step. And we have no other means of referring to the *it* that is either 0 or 1 by means of any other rule or articulation on the list that we can *follow*.

One outcome of both Turing's and Wittgenstein's proofs is that the extensional point of view is not or exclusive as a perspective in the foundations of mathematics. Wittgenstein's version of the Argument from the Pointerless Machine shows that the particular rule, F′(n), cannot be identified with any of the rules on the list, because it cannot be applied if we try to think of it as a particular member of the list. The

[26] Watson uses the metaphor that the machine "gets stuck" (Watson 1937, p. 445), but I have not found that metaphor either in Wittgenstein or Turing: it is rather ambiguous, and does not distinguish Turing's First Argument from that of the Pointerless Machine. Both Watson and Turing attended Wittgenstein's 1939 lectures at Cambridge; see (Wittgenstein 1989) where the metaphor of a contradiction "jamming" or "getting stuck" is criticized. I assume this is in response to a worry about the way of expressing things found in Watson 1937. He worries that the machine metaphor may bring out a perspective on logic that is either too psychologistic, or too experimental. He emphasizes, characteristically, that instead what matters if we face a contradiction is that we do not recognize any action to be the fulfillment of a particular order, we say, e.g., that it "makes no sense". As he writes in the 1947 remarks considered here, "an order only makes sense in certain positions". Recall Z §689: "Why is a contradiction to be more feared than a tautology"?

argument shows a "crossing of pictures" or concepts which yields something new. If one likes, it proves that there is a number which is not a number given on the list, for it shows how to construct a rule for a sequence of 0s and 1s which cannot be a rule on the list like the others. The argument would apply, moreover, in any context in which the rule-articulable ("computable") real numbers were asserted to be listed or enumerated in any way according to a rule – including, of course, any context in which, more controversially, one assumed that *only* rule-articulable real numbers *are* real numbers. But this particular assumption is not essential, either to Turing's or to Wittgenstein's arguments, which involve no such necessarily revisionary constructivist or finitistic implications or assumptions.

To recapitulate. Unlike the Halting Problem or the First Argument presented above, Wittgenstein's argument does not apply the law of the excluded middle, or any explicit contradiction or negation *by* the machine. It is not propositional, but in a sense purely conceptual or performative, turning on the idea of a coherently expressed command that turns out, upon reflection, to be empty, thereby generating a rule that we *see* cannot be applied in the same way as other rules are applied. There is of course no direct appeal to community-wide standards of agreement or any explicit stipulation used to drawn the conclusion, so, it is not a purely "conventional" argument, though we see that the order could not be followed by anyone. Oddly, because it turns on a tautology, its conclusion is "positive": it "constructs" a formulable rule that cannot be literally identified with any of the rule-commands on the list of rules supposed to be given. The diagonal then gives one a positive way of creating something new, i.e., a directive that cannot be sensibly followed.

Before commenting further on this version of the proof, I want to underscore that as I have construed it there is no *rejection* of the results of Turing or Cantor involved in accepting Wittgenstein's Diagonal Argument. To make this clear, I shall briefly rehearse an analogous argument.

2.4 The Positive Russell Paradox

Consider the binary array of 0's and 1's anew, but this time as a membership chart for an arbitrary set S.

$x_i \in x_j$?	1	2	3	4	…	
1	1	0	0	1	1	…
2	0	1	0	1	1	…
3	1	1	1	0	1	…
4	0	0	0	0	1	…
…						
						???

Let the array be a diagram of membership relations. At the point (i, j) if we see a "0", this indicates that $x_i \notin x_j$; if we see "1", it means $x_i \in x_j$.

Now let $S = \{x_i | x_i \in x_i\}$. This is the exact complement, so to speak, of the usual Russell set of all sets that are *not* members of themselves: I think of it as the *positive* Russell set. Whenever there is a "1" at a point (i, i) along the diagonal, this means that $x_i \in S$. In a certain sense, S "comes before" Russell's set, for there is no use of negation in its definition.

Is $S = x_j$ for some j? Well there is a difficulty here. For $x_j \in x_j$ iff $x_j \in S$. But $x_j \in S$ iff $x_j \in x_j$. So we are caught in a circle of the form "it is what it is". This cannot be implemented.

An apparently unproblematic way of thinking is applied here, but two different ways of thinking about S are involved. They are at first blush buried, just as in Russell's usual form of the paradox, but they are there, and they are separable, viz., there is the thinking of S as an object or element that is a member of other sets, and the thinking of S as a concept, or defining condition.

We have here what might be regarded, following Turing and Wittgenstein, as a kind of performative or empty rule. You are told to do something depending upon what the rule tells you to do, but you cannot do anything, because you get into a loop or tautological circle. This set membership question cannot be a question on the list which you can apply, because you cannot apply the set's defining condition at every point. (An analogous line of reasoning may be applied to, e.g., "autological" in the Grelling paradox. Without negation, one does not get a contradiction, but one may generate a question that may be sensibly answered with a either Yes or No question, i.e., with a question that is unanswerable *in that sense*.)

Is the Positive Russell argument "constructive"? In a sense Yes. It does not have to be seen to apply to actually infinite objects and name them directly, or invoke any axioms of set theory involving the infinite, though of course it might.[27] So, in this other sense, No. Its outcome is that there is an essential lack of uniformity marking the notion of a rule that can be applied. It involves no use of negation in the rule itself. So what is essentially constructive here is the implication: *If* you write the list as a totality, *then* you will be able to formulate a new rule. And *it* will yield a question one cannot answer without further ado, i.e., *that* rule will not be applicable in the same sense.

The Positive Russell argument refers to an extensional context, that of sets. But there is a creative, "positive" aspect of the argument that emerges, just as it does in Turing's and Wittgenstein's Pointerless Arguments. One must appreciate something or see something about what does *not* direct (any)one to do a particular thing, or assert the existence of a particular solution – rather than being forced to admit the existence of something. Cantor's diagonal argument is often presented as doing the latter, and not the former. But, as Turing and Wittgenstein's proofs make clear, Cantor's argumentation is actually furnishing the materials for more than one

[27] S is empty by the axiom of foundation. Quine worked with *Urelemente* of the form x={x}, sets whose only members are themselves. (Quine (1937), Reprinted in Quine (1953, 1980)).

kind of argument. Such, I suggest, is Wittgenstein's point in writing in the above-quoted remark of 1947 that Cantor did two different things. This is not to deny that Wittgenstein's argument is insufficient for Cantor's wider purposes, just as Turing's is, and for the same reason. These later "variants" of Cantor's argument are proofs with and about rules, not proofs utilizing or applying to actually infinite totalities. Nevertheless, we can distinguish Cantor's argumentation from his proof and from its applications, and regard what Turing and Wittgenstein do as "variants" of what Cantor did.

2.5 Interpreting Wittgenstein

The "pointerless" proofs I have considered are down-to-earth in the way Wittgenstein and Turing liked: the "entanglement" in the idea of an exhaustive listing of rules is exhibited in the form of a recipe for a further rule, and the diagonal argument is conceived as a kind of process of conceptualization that generates a new kind of rule. The reasoning in both cases, is, moreover, presented in a way unentangled with any expression in a particular formalism. This does not mean that the arguments are unformalizeable, of course: certainly they apply, as Turing taught us, to formal systems of a certain kind. And a Turing Machine may well be conceived of as a formal system, its activities encodable in, e.g., a system of equations. But Turing's Machines, being framed in a way that is unentangled with a specific formal system, also offer an analysis of the very notion of a formal system itself. This allows them to make general sense of the range of application of the incompleteness theorems, just as Gödel noted.[28]

Turing's and Wittgenstein's arguments from pointerless commands *evidently* do an end run around arguments over the application of the law of the excluded middle in infinite contexts, as other diagonal arguments do not. In this sense, they make logic (the question of a choice of logic) disappear. But I hope that my reconstruction of Wittgenstein's Diagonal Argument will go some distance toward in responding to the feeling some readers have had, namely, that Wittgenstein takes Cantor's proof to have no deductive content at all. It has been held that Wittgenstein took Cantor to provide only a picture or piece of applied mathematics warning against needless efforts to write down all the real numbers.[29] And it is true that Turing's and Wittgenstein's arguments require us to conceive of functions as presented through a collection of commands, rules, directives, in an *intensional* fashion. But they leave

[28] In a note added in 1963 to a reprinting of his famous 1931 incompleteness paper, Gödel called Turing's analysis "a precise and unquestionably adequate definition of the general notion of formal system", allowing a "completely general version" of his theorems to be proved. See Gödel (1986), p. 195. On the subject of "formalism freeness" in relation to Gödel see Kennedy (unpublished). Compare footnote 19.

[29] Hodges (1998).

open in what sense this notion, or the notion of a rule, is meant (i.e., the digits of 0s and 1s are a mere *façon de parler* in the way I have presented the arguments here). A critique of the idea that the extensionalist attitude is the *only* legitimate attitude is implied, though, as I have argued, no refutation of extensionalism, Cantor's Diagonal Proof, or set theory follows.

Of course, Wittgenstein's remarks criticizing extensionalism as an exclusively correct point of view are well known. So are his suggestions to look upon mathematical statements as commands. However, though I shall not argue the point here, it seems to me that taking Wittgenstein's Diagonal Argument seriously, at its word, should call into question the idea that he is either dogmatic or skeptical about the notion of following a rule and the "intensional" point of view – unless one means that the notion of a rule and the following of a rule in general are something to be *uniformly* understood in terms of a special kind of fact or intuitive insight. Neither Wittgenstein nor Turing believed this. Wittgenstein's Diagonal Argument serves, instead, to call into question forms of constructivism that take the notion of rule-following as clear or uniform. (I hope to discuss elsewhere the interpretations of Fogelin,[30] Kripke and Wright in light of the diagonal arguments I have discussed here.) His "everyday" version of the Argument from the Pointerless Machine, even more than Turing's, shows that there is a way of carrying out Cantor's argumentation that involves and applies to an "everyday" appeal to our sense of our ordinary activities when we compute or follow rules. In this sense, it makes the argumentation intelligible. One might want to say that it is more deeply or broadly anthropomorphic and intensional than Turing's. But that would be misleading. There is no scale involved here.

Thus it seems to me that one of the most important things to learn from Wittgenstein's argument is that the very idea of a single "intensional" approach is not clear off the bat – any more than are the ideas that perception, understanding, and/or thought are intensional. Wittgenstein's "game" argumentation involves, not merely the notion of a rule, recipe, representation or feasible procedure, but some kind of understanding of *us*, that is, those who are reading through the proof: we must *see* that we can do nothing with the rule that is formulated. Not all rules are alike, and we have to sometimes *look and see* how to operate or use a rule before we see it aright.

This last point is what Wittgenstein stressed just before the 1947 remarks I have discussed in this paper. He wrote,

> That we *calculate* with some concepts and with other do not, merely shows how different in kind conceptual tools are (how little reason we have ever to assume uniformity here). (RPP I §1095; cf. Z §347)

One of the most important themes in Wittgenstein's later philosophy starts from just this point. The difficulty in the grammar of the verb "to see" (or: "to follow a rule") is not so much disagreement (over a particular step, or a way of talking about *all* the steps), but instead that we often can get what we call "agreement" much

[30] Fogelin (1987).

too quickly, too easily. And thus we may be much too quickly inclined to think that we understand what is signified by (what we conceive of as) "agreement" and "disagreement" (or "rule of computation"). Quietism is one thing, unclear apparent agreement is another. Apparent agreement may well hide and mask the very basis and nature of that agreement itself, and an agreement may well turn out to rest upon a misunderstanding of what we share. Just as we may get someone much too quickly to agree that "Yes, of course the shape and colors are part of what I see", we may get someone much too quickly to agree that "Yes, of course it is not possible to list all the real numbers" (cf. RPP I §1107). The difficulty is not, in such a case, to decide on general grounds whether to revise the principles of logic or not, or whether to resolve an argument by taking sides Yes or No, e.g., with Hilbert or Brouwer. The difficulty is to probe wherein agreement does and does not lie, by drawing conceptual boundaries in a new way and paying attention to the details of a proof. Wittgenstein's and Turing's arguments as I have presented them here are neither revisionary nor anti-revisionary in a global way. What they do is to shift our understanding of what such global positions do and do not offer us.

References

Church, A. 1936. An unsolvable problem of elementary number theory. American Journal of Mathematics 58: 345–363.
Copeland, B.J. (ed.). 2004. The essential Turing: The ideas that gave birth to the computer age. Oxford: Clarendon Press.
Dreben, B., and J. Floyd. 1991. Tautology: How not to use a word. Synthese 87(1): 23–50.
Floyd, J. 2001. Prose versus proof: Wittgenstein on Gödel, Tarski and Truth. Philosophia Mathematica 3(9): 901–928.
Fogelin, R.J. 1987. Wittgenstein. London/New York: Routledge & K. Paul.
Gandy, R.O. 1988. The confluence of ideas in 1936. In: The universal Turing machine: A half-century survey, ed. R. Herken, 55–112. New York: Oxford University Press.
Gefwert, C. 1998. Wittgenstein on mathematics, minds and mental machines. Burlington: Ashgate Publishing.
Gödel, K. 1986. Kurt Gödel collected works. Volume I: Publications 1929–1936. New York: Oxford University Press.
Gödel, K. 1990. Kurt Gödel collected works. Volume II: Publications 1938–1974. New York: Oxford University Press.
Hodges, A. 1983. Alan Turing the enigma of intelligence. New York: Touchstone.
Hodges, W. 1998. An editor recalls some hopeless papers. Bulletin of Symbolic Logic 4(1): 1–16.
Kennedy, J. (unpublished). Gödel's quest for decidability: The method of formal systems; The method of informal rigor.
Kreisel, G. 1950. Note on arithmetic models for consistent formulae of the predicate calculus. Fundamenta Mathematicae 37: 265–285.
Kripke, S.A. 1982. Wittgenstein on rules and private language: An elementary exposition. Cambridge: Harvard University Press.
Marion, M. 2011. Wittgenstein on the surveyability of proofs. In The Oxford handbook to Wittgenstein, ed. M. McGinn. New York/Oxford: Oxford University Press.
Martin-Löf, P. 1984. Intuitionistic type theory. Napoli: Bibliopolis.

Martin-Löf, P. 1996. On the meanings of the logical constants and the justifications of the logical laws. Nordic Journal of Philosophical Logic 1(1): 11–60.

McGuinness, B. (ed.). 2008. Wittgenstein in Cambridge: Letters and documents, 1911–1951. Malden/Oxford: Blackwell.

Mühlhölzer, F. 2010. Braucht die Mathematik eine Grundlegung? Ein Kommentar des Teils III von Wittgensteins Bemerkungen über die Grundlagen der Mathematik. Frankfurt am Main: Vittorio Klostermann.

Petzold, C. 2008. The annotated Turing: A guided tour through Alan Turing's historic paper on computability and the Turing machine. Indianapolis: Wiley Publishing, Inc.

Quine, W.V. 1937. New foundations for mathematical logic. American Mathematical Monthly 44: 70–80.

Quine, W.V. 1953, 1980. From a logical point of view. Cambridge: Harvard University Press.

Shanker, S.G. 1987. Wittgenstein versus Turing on the nature of Church's thesis. Notre Dame Journal of Formal Logic 28(4): 615–649.

Shanker, S.G. 1998. Wittgenstein's remarks on the foundations of AI. New York: Routledge.

Sieg, W. 1994. Mechanical procedures and mathematical experience. In Mathematics and mind, ed. A. George, 91–117. New York/Oxford: Oxford University Press.

Sieg, W. 2006a. Gödel on computability. Philosophia Mathematica 14(2): 189–207.

Sieg, W. 2006b. Step by recursive step: Church's analysis of effective calculability. In Church's thesis after 70 years, ed. A. Olszewski, J. Wolenski and R. Janusz, 456–485. Frankfurt/Paris/Ebikon/Lancaster/New Brunswick: Ontos Verlag.

Sieg, W. 2008. On computability. In Handbook of the philosophy of science: Philosophy of mathematics, ed. A. Irvine. Amsterdam: Elsevier BV.

Stenius, E. 1970. Semantic antinomies and the theory of well-formed rules. Theoria 35–36(36): 142–160.

Turing, A.M. 1937a. On computable numbers, with an application to the Entscheidungsproblem. Proceedings of the London Mathematical Society 2(42): 230–265.

Turing, A.M. 1937b. On computable numbers, with an application to the Entscheidungsproblem. A correction. Proceedings of the London Mathematical Society 43(Part 7, 2nd Series): 544–546.

Turing, A.M. 1937c. Letter to Ethel Sarah Turing. Cambridge, U.K.: King's College Archives, K/1/54, February 11, 1937.

Turing, A.M. 1937d. Correspondence with Paul Bernays. Zürich: Eidgenössische Technische Hochschule Zürich/Swiss Federal Institute of Technology Zürich, Bibliothek.

Turing, A.M. 1950. Computing machinery and intelligence. Mind 59: 433–460.

Turing, A.M. 1954. Solvable and unsolvable problems. Science News 31: 7–23.

Watson, A.G.D. 1938. Mathematics and its foundations. Mind 47(188): 440–451.

Webb, Judson C. 1990 Remark 3, introductory note to Gödel (1972a), in Kurt Gödel collected works. Volume II: Publications 1938–1974, eds. S. Feferman, et al., 281–304. New York: Oxford University Press.

Wittgenstein, L. 1970. Zettel [Z]. Berkeley: University of California Press.

Wittgenstein, L. 1978. Remarks on the foundations of mathematics. Cambridge: MIT Press.

Wittgenstein, L. 1980. Wittgenstein's lectures, Cambridge 1930–32, from the notes of John King and Desmond Lee [DL]. Oxford: Blackwell.

Wittgenstein, L. 1999. The published works of Ludwig Wittgenstein [CD-Rom], Charlottesville, VA/Oxford: Intelex Corporation. Oxford University Press.

Wittgenstein, L. 2004. Ludwig Wittgenstein: Briefwechsel (CD-Rom, Innsbrucker elektronische Ausgabe), ed. M. Seekircher, B. McGuinness, A. Unterk Charlottesville, VA: Intelex Corporation.

Wittgenstein, L., and G.H.V. Wright et al. 1980. Remarks on the philosophy of psychology, vol. I [RPP I]. Chicago/Oxford: University of Chicago Press/B. Blackwell.

Wittgenstein, L. 1989. Wittgenstein's lectures on the foundations of mathematics: Cambridge, 1939, ed. C. Diamond. Chicago: University of Chicago Press.

Wright, C. 2001. Rails to infinity: Essays on themes from Wittgenstein's philosophical investigations. Cambridge: Harvard University Press.

Square brackets, as in line 684?