Wittgenstein, Turing and Gödel

Juliet Floyd

Boston University

Lichtenberg-Kolleg, Georg August Universität Göttingen

Japan Philosophy of Science Association Meeting, Tokyo, Japan
12 June 2010

Georg-August-Universität Göttingen

Lichtenberg-Kolleg - in der Historischen Sternwarte -

BOSTON UNIVERSITY

Wittgenstein on Turing (1946)

RPP I 1096. 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”. One has received the order "Go on in the same way" when this makes no sense, say because one has got into a circle. For that order makes sense only in certain positions. (Watson.)

Talk Outline:

- Wittgenstein's remarks on mathematics and logic
- Turing and Wittgenstein
- Gödel on Turing compared

Wittgenstein on Mathematics and Logic

- The most dismissed part of his writings {although not by Felix Mülhölzer – BGM III}
- Accounting for Wittgenstein's obsession with the intuitive (e.g. pictures, models, aspect perception)
- No principled finitism in Wittgenstein
- Detail the development of Wittgenstein's remarks against background of the mathematics of his day

Machine metaphors in Wittgenstein

- Proof *in* logic is a "mechanical" expedient
- Logical symbolisms/mathematical theories are "calculi" with "proof machinery"
- Proofs in mathematics (e.g. by induction) exhibit or *show* algorithms
- PR, PG, BB: "Can a machine think?"
- Language (thought) as a mechanism
- Pianola
- Reading Machines, the Machine as Symbolizing its own actions, "Is the human body a thinking machine?" is *not* an empirical question

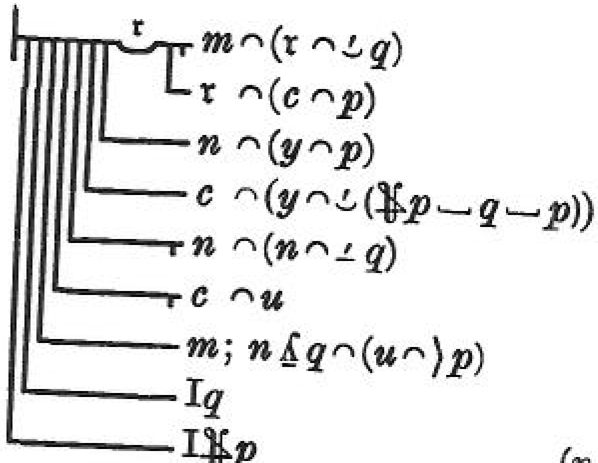
Turing Machines

Turing resolved Hilbert's *Entscheidungsproblem* (posed in 1928):

Find a definite method by which every statement of mathematics expressed formally in an axiomatic system can be determined to be true or false based on the axioms.

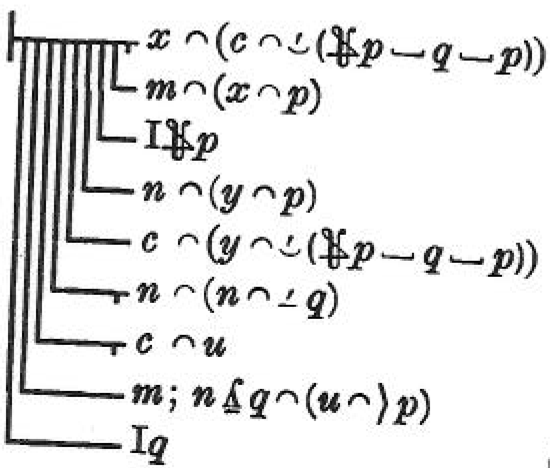
(The method need not generate a proof; it had only to be always correct. Mainly a logical problem.)

Turing (1936): There can be no such method.



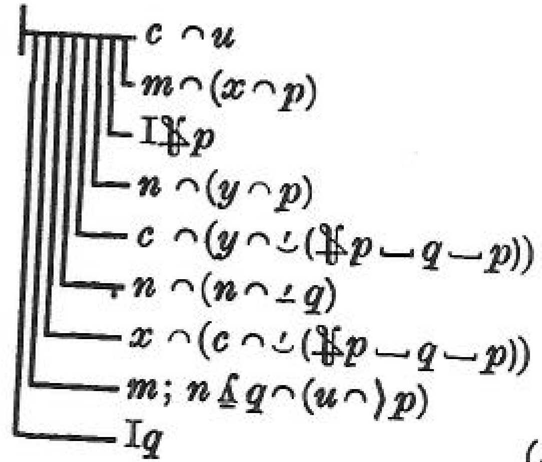
(η

(338): — — — — — — — — — — — —



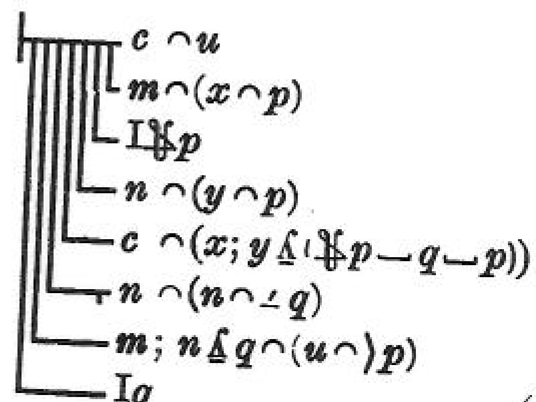
(ϑ

×



(ι

(269, 270):: = = = = = = = = = =



(κ

(18):: — — — — — — — — — — —

$*13{\cdot}192.\ \vdash :.\ (\exists c) : x = b\ .\equiv_x .\ x = c : \psi c : \equiv .\ \psi b$

Dem.

$\vdash . *4{\cdot}2 . *3{\cdot}2 . \supset \vdash :: \psi b . \supset :. x = b . \equiv_x . x = b : \psi b :.$

$[*10{\cdot}24] \qquad \supset :. (\exists c) : x = b . \equiv_x . x = c : \psi c \qquad (1)$

$\vdash . *10{\cdot}1 . \supset \vdash :. x = b . \equiv_x . x = c : \psi c : \supset : b = b . \equiv . b = c : \psi c :$

$[*5{\cdot}501 . *13{\cdot}15] \qquad \supset : b = c . \psi c :$

$[*13{\cdot}13] \qquad \supset : \psi b \qquad (2)$

$\vdash . (2) . *10{\cdot}11{\cdot}23 . \supset \vdash :. (\exists c) : x = b . \equiv_x . x = c : \psi c : \supset . \psi b \qquad (3)$

$\vdash . (1) . (3) . \supset \vdash .$ Prop

This proposition is useful in the theory of descriptions (*14).

$*13{\cdot}193.\ \vdash : \phi x . x = y . \equiv . \phi y . x = y$

Dem.

$\vdash .$ Simp $. \qquad \supset \vdash : \phi x . x = y . \supset . x = y \qquad (1)$

$\vdash . *13{\cdot}13 . \qquad \supset \vdash : \phi x . x = y . \supset . \phi y \qquad (2)$

$\vdash . (1) . (2) .$ Comp $. \supset \vdash : \phi x . x = y . \supset . \phi y . x = y \qquad (3)$

$\vdash . *13{\cdot}16 .$ Fact $. \qquad \supset \vdash : \phi y . x = y . \supset . \phi y . y = x .$

$\left[(3)\dfrac{y, x}{x, y}\right] \qquad \supset . \phi x . y = x .$

$[*13{\cdot}16 .$ Fact$] \qquad \supset . \phi x . x = y \qquad (4)$

$\vdash . (3) . (4) . \supset \vdash .$ Prop

This proposition is very often used.

p	q	r	p ≡ q	p ≡ r	q ≡ r	(p ≡ q) v (p ≡ r) v (q ≡ r)
T	T	T	T	T	T	T
T	T	⊥	T	⊥	⊥	T
T	⊥	T	⊥	T	⊥	T
T	⊥	⊥	⊥	⊥	T	T
⊥	T	T	⊥	⊥	T	T
⊥	T	⊥	⊥	T	⊥	T
⊥	⊥	T	T	⊥	⊥	T
⊥	⊥	⊥	T	T	T	T

22. $FR(x) \equiv (n)\{0 < n \leq l(x) \rightarrow Elf(n\,Gl\,x) \vee$
$(Ep, q)[0 < p, q < n \,\&\, Op(n\,Gl\,x, p\,Gl\,x, q\,Gl\,x)]\} \,\&$
$l(x) > 0,$

x is a SEQUENCE OF FORMULAS, each term of which either is an ELEMENTARY FORMULA or results from the preceding FORMULAS through the operations of NEGATION, DISJUNCTION, or GENERALIZATION.

23. $\text{Form}(x) \equiv (En)\{n \leq (Pr([l(x)]^2))^{x \cdot [l(x)]^2} \,\&$
$FR(n) \,\&\, x = [l(n)]\,Gl\,n\},$[35]

x is a FORMULA (that is, the last term of a FORMULA SEQUENCE n).

24. $v\,Geb\,n, x \equiv \text{Var}(v) \,\&\, \text{Form}(x) \,\&$
$(Ea, b, c)[a, b, c \leq x \,\&\, x = a * (v\,Gen\,b) * c \,\&$
$\text{Form}(b) \,\&\, l(a) + 1 \leq n \leq l(a) + l(v\,Gen\,b)],$

the VARIABLE v is BOUND in x at the nth place.

25. $v\,Fr\,n, x \equiv \text{Var}(v) \,\&\, \text{Form}(x) \,\&\, v = n\,Gl\,x \,\&\, n \leq l(x) \,\&\, \overline{v\,Geb\,n, x},$

the VARIABLE v is FREE in x at the nth place.

26. $v\,Fr\,x \equiv (En)[n \leq l(x) \,\&\, v\,Fr\,n, x],$

v occurs as a FREE VARIABLE in x.

27. $Su\,x\binom{n}{y} \equiv \varepsilon z\{z \leq [Pr(l(x) + l(y))]^{x+y} \,\&\, (Eu, v)[u, v \leq x \,\&$
$x = u * R(n\,Gl\,x) * v \,\&\, z = u * y * v \,\&\, n = l(u) + 1]\},$

$Su\,x\binom{n}{y}$ results from x when we substitute y for the nth term of x (provided that $0 < n \leq l(x)$).

28. $0\,St\,v, x \equiv \varepsilon n\{n \leq l(x) \,\&\, v\,Fr\,n, x \,\&$

q_1	S_0	PS_1, R	q_2
q_2	S_0	PS_0, R	q_3
q_3	S_0	PS_2, R	q_4
q_4	S_0	PS_0, R	q_1

Other tables could be obtained by adding irrelevant lines such as

q_1	S_1	PS_1, R	q_2

Our first standard form would be

$$q_1 S_0 S_1 R q_2;\ q_2 S_0 S_0 R q_3;\ q_3 S_0 S_2 R q_4;\ q_4 S_0 S_0 R q_1;.$$

The standard description is

DADDCRDAA;DAADDRDAAA;

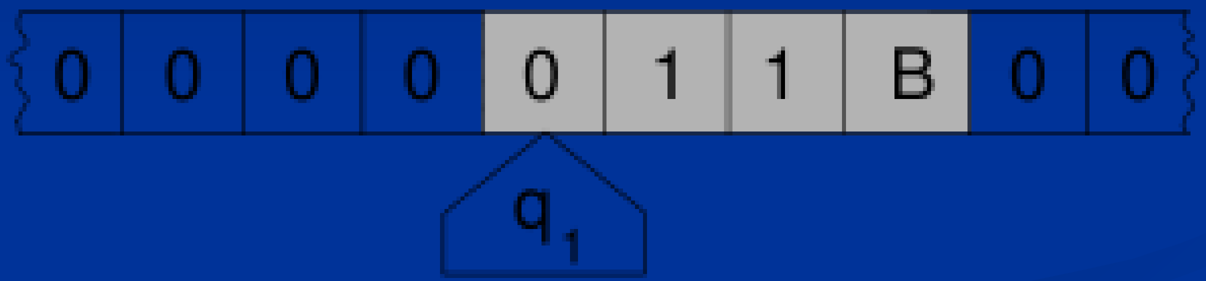
DAAADDCCRDAAAA;DAAAADDRD

A description number is

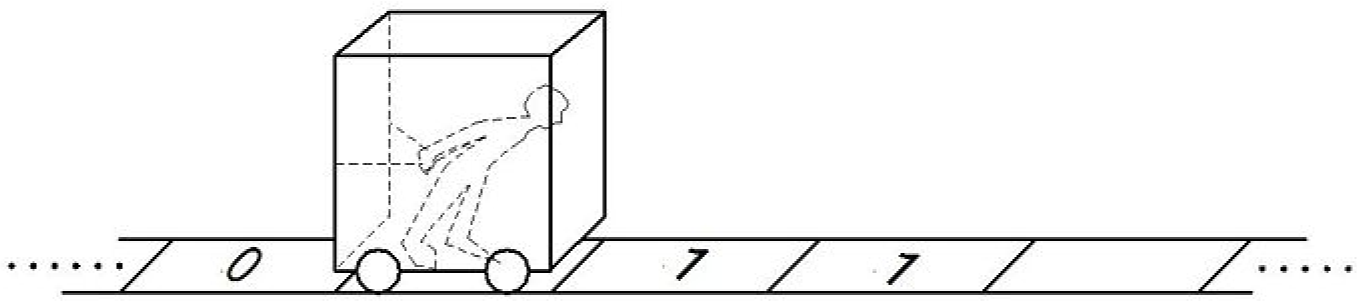
31332531173113353111731113322531111731111335317

and so is

3133253117311335311173111332253111173111133531731323253117


0 0 0 0 0 1 1 B 0 0
q1

‘Poor mug in a box’ visualization (Boolos, Jeffrey and Burgess):



Carnap and Turing's analysis: Carnap never mentions Turing

- Choice of logic not at issue
- Choice of linguistic framework not at issue
- Internal coherence of a framework not at issue
- Internal coherence/strength of a metastance not at issue (Principle of Tolerance)
- Formalization doesn't settle abstract analysis, but not for general reasons, for specific reasons having to do with a special problem context

Wittgenstein and Turing: Some Approaches

- Wittgenstein a romantic humanist, Turing an arch mechanist.
- Wittgenstein essentially hostile to science and mathematical logic: overly negative about idealization, ignorant, sloppy, insignificant, obsessed with style, dyslexic, propagandistic, paraconsistent (discussions on contradictions with Turing, 1939 Cambridge lectures)

Judson Webb, *Mechanism, Mentalism, and MetaMathematics* (1980); 'Introductory Note to Gödel 1972a'

- (Feferman): Basic new feature of Turing's machines is the *deterministic* character of their computations which automatically ensures the *consistency* of Turing's definition of computability; Webb: but we must dig deeper
- Formalism not refuted by Gödel 1931, Turing
- Philosophy of mind, finitism, mechanism, but not physicalism at issue: Fallacy to think that a mind that understands infinity requires an infinite number of states (Wittgenstein.)

Rejection of Cognitive Science

Proudfoot and Copeland, "Turing, Wittgenstein and the Science of the Mind", *Austr. J. of Phil.* (1994)

- Wittgenstein offers a "wholesale dismissal" of cognitive scientist's picture, not just Cartesianism.
- Rejects a wide conception of a thinking thing.
- Rejects arguments justifying existence of other minds.
- Wittgenstein denies perception is a mediated process
- Intentionality and normativity in language are not derivative for Wittgenstein

Wittgenstein is a Helpful Supplement to the Cognitivist Picture

Justin Leiber, *An Invitation to Cognitive Science* (1991)

Dan Bullock, "Social Interaction, Language Games, and Cognitive Convergence Rate", *Intellectica* (1998).

Rohit Parikh, Social Software (2001)

P.M.S. Hacker, *Wittgenstein's Place in Twentieth Century Analytic Philosophy* (1996)

Turing's metaphysics of mental states and persons is mechanistic; Turing was in this respect fundamentally opposed to Wittgenstein, implicitly or explicitly a focus of his criticisms

Stewart Shanker, "Wittgenstein vs. Turing on the Nature of Church's Thesis" (1987); *Wittgenstein and the Foundations of AI* (1998):

- LW rejects Church's thesis, that all humanly computable functions are Turing computable, and the behaviorism buried in the AI program. Rule-following is normative.
- Wittgenstein's criticisms of Turing form part of the story of inadequacies in Frege's form of anti-psychologism, which allowed the artificial intelligence program in psychology to take root.

A different approach

- Turing exactly the kind of logician/mathematician Wittgenstein approved of. (Phil. of mind not the main issue).
- Document reactions we have Wittgenstein to Turing and Turing to Wittgenstein
- Wittgenstein and Turing have strikingly overlapping interests (in incompleteness phenomena, machine metaphors, decidability, the continuum and the reals, the human/machine interplay, truth, realism, scepticism about 'foundations')
- Wittgenstein investigates the idea that all of mathematics is "experiment" after talking with Turing.

Turing and Wittgenstein

- Turing to Moral Sciences club, 1933:
  purely logistic view of mathematics is inadequate; mathematical propositions possess a variety of interpretations, of which the logistic is merely one.

- Turing 1936: "On Computable Numbers"; universal machine, unsolvability of the halting problem, irreality of uncomputable reals, extrusion of axiomatic approach; some discussion of finitude of human memory in computation.

A. Watson's *Mind* paper, 1938

- Reflects summer 1937 discussions; strikingly Wittgensteinsteinian in tone
- Von Wright reports Wittgenstein admired the paper
- A quite competent presentation of recent results
- Credits presentation of Gödel 's incompleteness theorem to "discussions with Turing and Wittgenstein"
- Reiterates "Poincaréan" objection to Russell-Frege analysis of number; criticisms of Dedekind cuts as essential definitions
- Presents Turing 1936, halting problem, Church 1936, Cantor's diagonal proof, continuum in terms of games
- Hardy on proof of irrationality of square root of 2

Wittgenstein and Turing

- 1939 Cambridge Lectures on the Foundations of Mathematics, ed. C. Diamond

  Wittgenstein mentions incompleteness, but Turing doesn't respond

  Focus on the status of contradictions

- 1939 appears Turing's "systems of logic based on ordinals"
- 1940(44-5) Turing's paper "Reform of Mathematical Notation" says Wittgenstein's lectures suggested "the statement of the type principle" in his reformulation of type theory.

Turing 1940 (44-5)

We should conduct an extensive examination of current mathematical, physical and engineering books and papers with a view toward listing all commonly used forms of notation and examine them to see what they really mean. (p.2, AMT/C12)

Turing 1940 (44-45)

This will usually involve statements of various implicit understandings as between writer and reader. But the laying down of a code of minimum requirements for possible notations should be exceedingly mild, avoiding the straightjacket of a logical notation.

Turing 1940 (44-45)

It is not difficult to put the theory of types into a form in which it can be used by the mathematician-in-the-street without having to study symbolic logic, much less use it. The statement of the type principle given below was suggested by lectures of Wittgenstein, but its shortcomings should not be laid at his door.

(p. 6, AMT/C12; highlighted by Gandy)

Wittgenstein on Turing 1950

- Wittgenstein to Malcolm 1950 on the Turing paper on “Computing Machinery and Intelligence”: “I haven’t read it but it is probably no leg pull”.

Gödel 1964

The precise and unquestionably adequate definition of the general concept of formal system [made possible by Turing's work allows the incompleteness theorems to be] proved rigorously for every consistent formal system containing a certain amount of finitary number theory.

Gödel (postscript 1972): rejects Turing's work as an analysis of "humanly effective" procedure

A philosophical error in Turing's work…

Turing in his 1937…gives an argument which is supposed to show that mental procedures cannot go beyond mechanical procedures. However, this argument is inconclusive. What Turing disregards completely is the fact that *mind, in its use, is not static, but constantly developing*, i.e., that we understand abstract terms more and more precisely as we go on using them, and that more and more abstract terms enter the sphere of this development.

Gödel on Turing, continued

Note that something like this indeed seems to happen in the process of forming stronger and stronger axioms of infinity in set theory. This process, however, today is far from being sufficiently understood to form a well-defined procedure.

Gödel on Turing, continued:

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.

Turing on the significance of incompleteness results:

(1947) If a machine is expected to be infallible, it cannot also be intelligent. There are several mathematical theorems which say almost exactly that. But these theorems say nothing about how much intelligence may be displayed if a machine makes no pretence at infallibility.

(1948) The argument from Gödel's and other theorems rests essentially on the condition that the machine must not make mistakes. But this is not a requirement for intelligence.

Turing, “Solvable and Unsolvable Problems (1954)

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, "Intelligent Machinery" (1948)

To convert a brain or machine into a universal machine is the extremest form of discipline. Without something of this kind one cannot set up proper communication. 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.

Sieg, "Gödel on Computability":

In a deep sense neither Church nor Gödel recognized the genuinely distinctive character of Turing's analysis, i.e., the move from arithmetically motivated calculations to general symbolic processes that underlie them. Most importantly in the given intellectual context, these processes have to be carried out programmatically by human beings: the *Entscheidungsproblem* had to be solved by us in a mechanical way; it was the normative demand of radical intersubjectivity between humans that motivated the step from axiomatic to formal systems.

Sieg, “Gödel on Computability”:

It is for this very reason that Turing most appropriately brings in human computers in a crucial way and exploits the limitations of their processing capacities, when proceeding mechanically.

Wittgenstein on Turing (1946)

RPP I 1096. 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". One has received the order "Go on in the same way" when this makes no sense, say because one has got into a circle. For any order makes sense only in certain positions. (Watson.)

Wittgenstein on Turing (1946)

BPP I 1096. Turings 'Maschinen'. Diese Maschinen sind ja die Menschen, welche kalkulieren. Und man könnte, was er sagt, auch in Form von Spielen ausdrücken. Und zwar wären die interessanten Spiele solche, bei denen man gewissen Regeln gemäß zu unsinnigen Anweisungen gelangt. Ich denke an Spiele ähnlich dem "Wettrennspiel". Man erhielte etwa den Befehl "Setze auf die gleiche Art fort", wenn dies keinen Sinn ergibt, etwa, weil man in einen Zirkel gerät; denn jener Befehl hat eben nur angewissen Stellen Sinn. (Watson.)

1097. 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. The rule of the game runs "Do the same as..."—and in the special case it becomes "Do the same as you are doing".

Wittgenstein’s Variant of Cantor’s Diagonal argument: Does F’ (n) = K (100,100)?

n ⇒

K ⇓

1	1	2	3	4	5	...
2	0	1	0	1	1	...
3	1	1	1	0	1	...
4	0	0	0	0	1	...
5	1	1	1	1	1	...
...	...					
100	...	...	...	...	...	????

This is F'(n)

The calculation/experiment distinction

RPP 1095. That we calculate with some concepts and with others do not, merely shows how different in kind conceptual tools are (how little reason we have ever to assume uniformity here).

…

Everywhere a shifting of concepts.

Rule-Following

2,4,6,8...