Rationalist Philosophy of Mathematical Proof

Abstract

Mechanism entails the view that all mathematical proofs can be formulated within first-order logic, which is further characterised as analytic logic. The dispute between mechanists and rationalists is dialectical: both sides agree with the facts about Gödel's theorem, but disagree over the nature of mathematical proof. A rationalist theory of mathematical proof is presented in which it is argued that mathematical inference is not always analytic: number theory, set theory and analysis are exemplars of synthetic mathematical proof. Although a machine that passed the Turing test would pose problems for the rationalist, a more rigorous version of that test is proposed. Both mechanists and rationalists make conflicting predictions about the future. Difficulties arise for the Penrose position that is both non-mechanist and materialist.

The dialectic

The project of building a computer that will simulate human behaviour in its entirety and thus pass the Turing Test has alarmed the minority that still cling to the beliefs of a bygone age – the freedom of the will and the immateriality of the soul. John Lucas is one such person.

A mechanist is someone who believes that that the human mind is a computer – more specifically, the human mind is a machine that can be modelled by a Turing machine. Church's Thesis states that whatever is Turing computable is a recursive function and that Turing machines and the theory of recursive functions equally constitute a full analysis of what an algorithm is. The mechanist thesis is equivalent to the claim that all valid mathematical inference can be expressed in the language of recursive functions, which is a theory embedded in first-order logic. As Wolf puts it: ""Most logicians (though perhaps not most mathematicians) are convinced that all correct proofs in mathematics could, with enough effort, be translated into formal proofs of first-order logic."[1]

Lucas proposed that "Gödel's theorem shows that mathematical insight need not be algorithmic"[2], thereby claiming to refute the mechanist position that the human mind is a Turing machine. During the exchange of papers that followed, he made a very perspicuous remark stating that "The argument is a dialectical one."[3]

First-order logic is a system of inference that admits quantification only over individuals; in second-order logic there is quantification over properties. Therefore, the mechanist also believes that no second-order inference that cannot be reduced to a first-order inference is meaningful. If there exists a second-order or informal inference that was acknowledged to be (a) meaningful, (b) valid, and (c) not reducible to a first-order inference, then that would constitute a refutation of mechanism, since such an inference could not be recursive.

Let us define a "monster" to be a meaningful, valid argument that could not be formalised within a first-order language. Then Lucas believes that he has found such a monster in Gödel's theorem.

> Granted that no false formula can be proved in Elementary Number Theory, it follows that the Gödelian formula is both true and unprovable from Peano's axioms. I thought I could

> apply this to the mechanist hypothesis that the human mind was, or could at least be represented by a Turing machine … there would be a Gödelian formula which could not be proved in the formal system and could be seen to be true by a competent mathematician who understood Gödel's proof.[4]

(My underlining.) What demonstrates that this argument is dialectical is that there is general agreement between the two sides on the "facts":

(1) The proof of Gödel's theorem is first-order.
(2) The statement of the theorem is conditional: If Peano Arithmetic is consistent and the theory is "sufficiently strong", then Gödel's theorem is true.
(3) No machine that is John Lucas has been specified.
(4) If a *K* is a first-order theory for which the Gödelian formula *X* is true but not provable, then $K \cup X$ is also a first-order theory in which *X* is both true and provable.

Lucas argues that in addition the mind can "see" that if a mind be identified with any given machine, a contradiction ensues. The emphasis is on "seeing" – that is, a species of mathematical intuition, which the mechanist would never grant. The mechanist replies that the human mind is limited just in the same way any Turing machine is limited[5]. He glosses that human creativity is limited in the same way the creativity of computers is limited[6].

A mechanist must be an empiricist, for to allow for non-empirical knowledge is to grant the non-mechanist the very premise that is in dispute. Therefore, it is fitting to describe the non-mechanist as a rationalist – since this is the viewpoint that is most characteristically opposed to empiricism.

> Empiricism: All knowledge is derived from sense-experience.
> Rationalism: There exist concepts that are (a) sources of (infallible) knowledge, and (b) not derived from sense-experience.

Penrose represents a viewpoint that is materialist, empiricist and yet non-mechanist, but for heuristic clarity I shall present the non-mechanist as a rationalist – one who believes (a) that knowledge can be of non-material objects called properties, concepts or universals, and (b) that hence the mind is equipped with a non-material, transcendental faculty that traditionally has been called "reason". In its fundamental character the dialectical debate between the mechanist and non-mechanist is a manifestation of the dispute between empiricism and rationalism.

How, one may ask, can mental faculties subsist without being grounded in material processes? Regardless of whether to the mechanist/empiricist the doctrine seems strange or not, the rationalist may assert that there is no material basis to the mind. The rationalist proposes that the doctrine of the immateriality of the mind is supported by examination of its faculty for making inferences.

First-order logic provides a series of algorithmic procedures that recursively generate proofs. Any such proof can be encoded in a binary machine. Hence, first order logic is mechanical. The algorithmic procedures provide a syntax. First-order logic is complete: any statement of first-order logic that is true within-the-system, can be proven mechanically to be so. There exists a positive test for validity in first-order logic.

> Mechanism (Formalism): Syntax equals semantics
> Rationalism: Syntax is not equal to semantics; there exist proofs that are not encoded in a syntax.

Gödel's theorem appears to offer the rationalist a single clear mathematical proof that is not encoded in a syntax. However, since the proof itself is a first-order proof, the rationalist claim must go outside the proof, and appeal to mathematical intuition ("We just see it"), which the empiricist denies; hence the argument is dialectical.

Ways to the resolution of a dialectical debate: (1) finding a premise on which both parties agree; (2) higher resolution as in Kant – a synthesis of thesis and antithesis; (3) conversion based on accumulation of evidence and insight; (4) conversion based on psychological investigation of motives for holding a viewpoint discovered upon self-examination to be held in bad faith.

Kant's *Critique of Pure Reason* was an attempt to reach a higher-level synthesis of empiricism and rationalism, exemplified in the four antinomies. In each of these, the apparent conflict between the empiricist thesis and the rationalist anti-thesis is resolved by appeal to the distinction between empirical and transcendental reality; the rationalist claim to have direct access to noumena in transcendental reality is rejected, but the empiricist claim that all knowledge comes from sense data is also shown to be false, since there is knowledge that encodes our ability to understand the world of phenomena, and this is said to be synthetic a priori. Kant's idealism historically formed the framework of C19th philosophy.

In the C20th the Kantian synthesis was overthrown by the empiricists. Landmarks of this cultural development include the influence of the Vienna circle, the essay *Two Dogmas of Empiricism* by Quine in which the notion of the synthetic a priori is rejected, Moore's *Proof of an External World*, the emergence of the American New School of Realism, the rejection of the logic of intensions propounded by Bosanquet, Bradley and Husserl in favour of the logic of extensions, advanced by Russell; the influence of Wittgenstein.

Since the empiricists rejected the Kantian synthesis, they restored the dialectic of the pre-Kantian period. Then, two worldwide schools of philosophy might have arisen, one representing the modern update of the C18th empiricism advanced by Hume, and the other the modern update of the rationalism of the C17th found in Descartes or Leibniz. This did not happen. Rationalism was overwhelmed and came scarcely to be represented in academic circles; the empiricists triumphed right across the board.

By rejecting the Kantian synthesis, the empiricists also rejected the Kantian solution to the problem of freewill, and thus reinstated that dialectic[7].

> Mechanism: that the mind is determined.
> Rationalism: that the mind (reason) is possessed of freewill, a causality not otherwise determined by events in time.

The development of logic lies at the centre of the mechanist/empiricist movement. If logic is the science of inference based on intensions mediated by phenomenological enquiry as a theory of judgement, then the empiricist case collapses immediately. Logic, therefore, came to be treated only as a science of extensions, in which first-order logic deals unambiguously.

Regarding the philosophy of mathematics, with the collapse of the neo-Kantian solution there initially emerged: (1) logicism, (2) formalism[8], (3) intuitionism, (4) Hibertism[9]. All four theories have coalesced into formalism, which is the dominant theory of the contemporary period, and the empiricist solution. No rationalist philosophy of mathematics has been forthcoming.

Hence, from a socio-cultural point-of-view, no attempt to use a single result (Gödel's theorem) could possibly succeed in over-turning the spirit of the age, which is empiricist and mechanist: for such a program to succeed, the rationalist must obtain the consent of empiricists to a premise that is equivalent to denying their empiricism. A single rationalist argument out of context of a systematic rationalist philosophy could not succeed.

Rationalist philosophy of proof

Now I shall construct a rationalist philosophy of proof. I begin by examining examples of how rationalists would see routine operations in mathematics, even those whose formal equivalence to algorithms is not in dispute, before presenting those further issues that expose to scrutiny the mechanist assumption that all proof is algorithmic.

1. Word problems in elementary mathematics are typically converted to diagrams, and together formulated into algebraic symbols, which are then solved and reinterpreted.

The mechanist can maintain that the processes involved in human problem-solving in such cases have an underlying mechanical basis in the brain, that perception, encoding of words problems into algebra, and mathematical insight are mechanical at the material level, but in the absence of a specific proposed mechanism at that level to do all of this, the rationalist is not obliged to give up his way of seeing things, for to him, concepts and insights are not reducible to mechanical procedures.

2. Problems in which two answers are produced, where one of the answers must be discarded as having no meaning in the context, or being inconsistent. The human approach is to solve a problem by using an inconsistent system, and then removing the extraneous solution. It is equivalent to adding an additional hidden premise. A computer algorithm that proceeded in this way, would be inconsistent.

This example illustrates how human reasoning can make use of inconsistent arguments to derive consistent conclusions. However, it is probable that any such inconsistent procedure could be replaced by a consistent one that was algorithmic. Notwithstanding, for the rationalist the difference in the human and the mechanical procedure highlights the metaphysical distinctions between them.

3. Problems in which the fixing of an origin or axis introduces a concept that has no physical meaning, but can be dealt with by ad hoc interpretations.[10]

This represents mathematics as a tool for the mind in solving its problems; as a tool, the introduction of an arbitrary line of reference creates merely a problem of interpretation, but does not prima facie suggest to the rationalist that human intelligence is anything like a mechanical procedure.

4. The problem of interpreting the introduction of ideal elements. For example, the introduction of the number $i = \sqrt{-1}$. Imaginary numbers are related to the properties of quadratic functions. The symmetry of the function $y = x^2 - 1$ is expressed in the fact

that a reflection in the *y*-axis swaps its roots; the same symmetry property is expressed for the function $y = x^2 + 1$ in that this reflection swaps its imaginary roots, $+i, -i$; thus imaginary numbers are grounded in the symmetry of the quadratic function.

The rationalist perceives the connection between the formal definition of an imaginary number $i = \sqrt{-1}$ and the reflection symmetry of the quadratic function as a synthetic one presented to mathematical intuition. The mechanist must either reject such an interpretation or perceive it as just another instance of a mental epiphenomenon produced by a process that at the brain/machine level is an algorithm.

Logic in general is the science of deduction. That is, the science that wishes to expose the valid means by which we may advance from a proposition A to another proposition B.

A → B (A and B might also be collections of statements.)

First-order logic narrowly circumscribes what the arrow in this diagram could mean. It limits valid movements from A to B to those given in the rules and axioms of first-order logic. For example, modus ponens is formulated as a detachment rule for material implication. When we come to mathematics, it is assumed that any valid mathematical argument can be formulated in first-order set-theory (*ZFC*). The assumption is that proof is first-order, and that there is no valid movement from A to B that is other than what is given in first-order logic, or reducible to it.

> "*ZFC* is a remarkable first-order theory. All of the results of contemporary mathematics can be expressed and proved within *ZFC*, with at most a handful of esoteric exceptions. Thus it provides the main support for the formalist position regarding the formalizability of mathematics. In fact, logicians tend to think of *ZFC* and mathematics as practically synonymous."[11]

No mechanist can consider any proof that is not first-order axiomatisable as a candidate for a proof. But is it true that all meaningful logic is first-order?

First-order logic is analytic. By this I mean that there is no inference in first-order logic in which the conclusions contain more information than the premises. Inference in first-order logic may be likened to the notion of spatial containment. It is always valid to infer from a given space what is contained in that space, but not vice-versa. Self-reference in first-order logic is generally not possible, or can only be provided subject to certain restrictions, because it is not possible for a space to contain itself. First order logic is the complete analytic logic: there is no larger analytic logic in which more inferences are exposed to be valid.

The analytic character of first-order logic raises the question: is there a synthetic logic? Are there inferences in which the conclusions contain more information than the premises?

5. Mathematical induction, also called complete induction, is the prime candidate for such a synthetic principle of inference. In it we pass from finite premises, $P(0),\ P(k) \to P(k+1)$ to a general statement, for all $n,\ P(n)$, that applies to a potentially infinite collection, and hence contains more information than was given in the premises. Thus, number theory, which is grounded in mathematical induction, is a prima facie candidate for a synthetic logic.

The passage from the finite to the non-finite is mediated in the human mind by the concept of potential infinity. A rationalist would interpret that there is no mechanical basis in the human mind to the concept of infinity.

6. Aristotle in his reply to Zeno, introduced the distinction between the potential and actual infinity. In mathematical induction, the notion of infinity that is encountered is that of the potential infinite. It is the idea of going on and on *ad infinitum*. But the notion of the potential infinite is not encompassed by first-order logic. In first-order set theory, a theory of collections called sets that is embedded in the language of first-order logic, we encounter actually infinite collections. In mathematics, the notion of the potential infinite is represented by the symbol ∞, whereas in set theory the notion of an actually complete infinity is represented by ω.[12]

In what language is the distinction between the potential and actual infinite formulated? A rationalist would not accept that the human ability to conceive of the difference between the two concepts is grounded in some mechanical procedure of the brain that may be represented by a first-order theory.

7. We note that in analysis the expression $\lim_{n\to\infty} u_n = l$ uses both concepts of the potential and actual infinite, since on the left we have a sequence that continues indefinitely, *ad infinitum*, whereas on the right we presume the actual completion of this process and make it into a new object: a real number.

Analysis is that branch of mathematics that is founded upon the theory of limits. The rationalist affirms that analysis is not a first-order theory.

8. In set theory, the actual infinite is introduced in such a way that contradicts results obtained in number theory. The Archimedean principle states that no collection of numbers is bounded above[13]; but in set theory any limit ordinal is constructed precisely upon this basis[14]. Wolf writes: "... there is a least limit ordinal, which is called ω ("omega"). The members of ω are called **finite** ordinals or **natural numbers**. In other words, to a set theorist $\omega = \mathbb{N}$." [15]

The rationalist denies that $\omega = \mathbb{N}$ ("The set of all finite ordinals is identical to the collection of all natural numbers.") This contention extends to analysis, since the Archimedean principle is a consequence of the completeness property[16]. Thus, it seems that we have two systems rather than just one. The first is the logical system represented by first-order set theory. The second is number theory with its own axioms and the principle of mathematical induction. The first system is putatively analytic, whereas the second system is not. The rationalist proposes that there may be theorems provable in number theory that are not provable in first-order set theory.

9. First-order set theory is a more "powerful" theory than first-order logic alone; set theory is a candidate for a synthetic logic.

First-order set theory has in respect of its abstract collections Platonist tendencies that are altogether more fitting for a rationalist than for a mechanist or formalist.

In a series of papers in the early 1920s Henri Poincaré strove to refute the claims being made by Russell and Peano for the expository power of logic to define number[17]. In addition to claiming that their definitions involve circular reasoning, by defining number in terms of expressions that can only be understood by mediation of the concept of number, he argued that the distinctive way in which mathematicians prove things is by mathematical induction (complete induction). While Poincaré's claim was passed over, there developed within the philosophical community the assumption that mathematical induction is encompassed by first-order set-theory. The rationalist would maintain that the two systems:

Number Theory	First-order set theory
Potential infinity	Complete collections; finite and actually infinite
Complete induction	Rules of inference of first-order logic

are different theories, and that the former (number theory) is synthetic, while the latter is analytic in respect of its underlying logic, but also synthetic in respect of its use of the actual infinite. To this debate, it is pertinent to add:

(1) That the device by which infinity is added within set theory is by means of an axiom of infinity: that there exists an infinite set. Such a principle is In Kantian terms synthetic; for the empiricist, it is equivalent to an empirical hypothesis.
(2) That complete induction is only fully characterised by a statement of second-order logic.

The rationalist would eschew the device whereby the second-order axiom of complete induction is replaced by a first-order axiom schema.

10. Stewart Shapiro demonstrates that the Schröder-Bernstein and Cantor's Theorem are second-order theorems[18]. The first is used to define the notion of equipollent sets (sets of the same size) and establishes the rule that that if $|A| \le |B|$ and $|B| \le |A|$ then $|A| = |B|$; the second establishes that the power-set of any set is larger than that set.

This result establishes that significant and essential results required for the development of set theory are second-order. Since first-order theory represents the upper limit of what could be an analytic logic, this suggests that the language in which these two theorems are constructed is synthetic.

To derive proofs of both theorems Shapiro uses a substitution rule of inference. This raises the question as to whether second-order logic really is anything different from first-order logic. Implicitly, first-order logicians seek to constrain second-order logic in such a way that in principle no second-order inference could not be reformulated in first-order logic; hence, the mechanist would argue that no fundamental breach of the computational foundation to the human mind has been made.

This device creates a confusing picture of the properties of second-order logic. Whereas first-order logic is complete, recursively axiomatisable, and has a positive decision procedure for generating all valid theorems, second-order logic is incomplete, not recursively axiomatisable and has no such positive decision procedure. Furthermore, in first-order theory it is not possible to characterise the cardinality of the domain of discourse, there exist, for example, non-standard models of first-order number theory; and every first-order theory obeys the Lowenheim-Skolem "paradox" that states that a theory with a non-denumerable model has denumerable models, and vice-versa. However, even these statements are meta-theorems written in a second-order language. Diagonalization, by which the different cardinalities of sets is established, is not a first-order argument, and hence, a first-order logician could throw out the whole theory, claiming it was nonsense after all. Though not a logician, that was the instinct of the techno-realist Hamming in his well-known paper[19].

Second-order logic, via the Quinian "to be is to be the value of a variable" may be ascribed a Platonist/rationalist ontology; its very use is the prima facie exemplar of monsters in mathematics. Gödel's theorem, far from being an isolated case of a non-algorithmic inference, is just one among a collection of results, whose proofs could not be algorithms. Notwithstanding this, the formalist response to second-order logic has been to treat it as just another formal system. According to the approach of Henkin we give second-order logic a "semantics" that is built upon the same foundation as the semantics of first-order logic; and hence indicates that second-order logic really is just another form of first-order logic.[20]

Since second-order logic is non-axiomatisable, it is fundamentally different from first-order logic. The implicit assumption is that the semantics of second-order logic must be of a piece with the semantics of first-order logic; that the collections over which one quantifies in second-order logic are reducible to the collections of first-order logic. This assumption can be challenged: the properties that are the subject of second-order logic may be nothing like first-order collections. Certainly, the rationalist will have nothing to do with that assumption; for him, second-order logic is per se a form of concept writing; and its concepts are not disguised extensions, though they may have extensions.

There exists a limited second-order logic which is an analytic extension of first-order logic, and it is this logic that Henkin and others have examined, and provided a semantics. Notwithstanding, second-order logic in the fullest sense is a synthetic logic, and does not have a semantics in the formal sense that Henkin supplies. The underlying assumption is that the meta-language, in which most real mathematics is conducted, must be a form of the first-order object language in which certain abstract structures are given.

> "… Nothing stated above implies that in our opinion there is any fundamental difference between metamathematics and mathematics "proper". Quite the contrary: we believe that, from every reasonable point of view, metamathematics is an integral part of mathematics."[21]

Given an expression in a mathematical discourse, "If A, then B", what is there to demonstrate that the implication is necessarily interpreted in the manner of first-order logic – by material implication? Gentzen was said to have formulated his natural deduction system upon the study of the real arguments that mathematicians use. He could not have looked very far for examples, since his system is equivalent to first-order logic, and differs only in that it has more rules and no axioms.

11. Analysis is founded upon the completeness axiom, which is universally acknowledged to be a second-order principle[22]. Thus, the first-order logician must either throw out analysis or reformulate it.

Errett Bishop endeavoured to provide a theory of analysis consistent with mechanism, but he began by acknowledging that the completeness axiom is non-constructive[23]. To what extent his construction is successful is a debating point of the dialectic, but he does reject the completeness axiom in its Dedekind or equivalent formulation.[24]

12. The Dedekind Completeness Axiom is equivalent to (a) The Cantor Nested Interval Theorem, (b) the Heine-Borel Theorem, (c) the Cauchy Convergence Criterion, and (d) the Bolzano-Weierstrass theorem.[25] In what logic are these theorems of equivalence conducted?

Can Shapiro derive the Completeness Axiom and all its equivalences from his substitution rule alone? Similarly, it is a challenge for a first-order theorist to formalise within that language the entire theory of either the Riemann or Lebesgue Integral.

As already indicated, the response of constructivists such as Bishop is to reject classical analysis. But this approach has the character of so defining logical inference that mechanism shall be true by definition. The issue is dialectical; for nothing can "force" the mechanist to give up his mechanism; but the point focuses the debate on a matter of conflicting convictions.

Imagine that you are a mechanist working at a mathematical research institute. In the room next to you, there are mathematicians working on the Lebesgue integral. Upon examination, you discover that what they are doing cannot be formalised in a first-order language. You conclude that their activity is meaningless, because not first-order, and that they are engaged in an act of self-deception, for they think they understand what they could not possibly understand. You are entitled to your position, for no argument of mine could force you to abandon it, but, nonetheless, is this not the very paradigm of an act of faith? All one can say to those mathematicians who continue to work in analysis is: I know you to be wrong, to be deceiving yourself into believing that your meaningless symbols are meaningful. But your colleagues could say something similar of you.[26] A dialectical problem has the character of two opposed faiths pitted against one another.

13. Set theory was invented by Cantor to construct the continuum and solve the question of when Fourier series were convergent. But could set theory alone constitute a theory of the continuum? The objects of which the continuum is said to be comprised are points of zero measure; the objects of which pure set theory is comprised are the null set and all iterative instances of the null set by applications of the axioms of set theory. Firstly, how possibly could I come across the notion of a point from the examination alone of a collection of the iterative hierarchy of sets? It is like trying to deduce the experience of snow by one who has never encountered snow. However, the difficulty for set theory is not this alone. For it is reasonable to suppose that the continuum requires not one but two primitives. Imagine the continuum as a line of no width joining two points, *A* and *B*. Let us remove a point from this line. Has the extension (measure) of the line, the distance between *A* and *B* been in any way diminished? No. Could the removal of any denumerably actually infinite collection of such points diminish the measure of the line? No, any such denumerable collection must be equipollent to the set of rationals, and the removal of all the rationals leaves the entire real line behind. Could the removal of a set of points of cardinality continuum diminish the measure of the line? No, this also has been refuted: the Cantor set is a set of zero measure. Hence, there must be two primitives on the real line: points, and extensions. And set theory is not a complete theory of the continuum.

Consider the following argument:

(1) A point has zero measure.
(2) No infinite collection of points, whether denumerable or non-denumerable, has a positive measure.
(3) Therefore, extension is a primitive of the theory of the continuum.

Whether this argument be valid or not, in what language is this inference evaluated, and can the implication involved be modelled by material implication?

14. The Dedekind pigeon-hole principle is the principle that $n+1$ items cannot be fitted into *n* boxes. The rationalist contends that it is another principle of synthetic reasoning, independent of both analysis and number-theory.
15. Just as not every implication is an instance of material implication, so not every use of proof by contradiction is an instance of a first-order proof by contradiction. Proof by contradiction is another principle of synthetic reasoning, sometimes modelled by an analytic argument, more generally not.

We reach a conclusion. Rather than being a single instance of a theorem whose proof or consequences are non-algorithmic, Gödel's Theorem is just one among a myriad of examples. The illusion that all mathematics is first-order set-theory is punctured. That does not leave the mechanist without recourse, for he can argue that all mathematics that is not first-order or reducible to a first-order theory is meaningless, but that is the recourse of prior conviction.

Rather than being devoid of resources, the rationalist has a thorough-going self-consistent philosophy of proof. He perceives mathematics as fundamentally a mental activity involving the use of concepts to know and understand the world. He acknowledges the role of analytic, first-order logic in progressing in proof from *A* to *B*, but claims that there are many other forms of proof, and that the logical progression in deduction from *A* to *B* is frequently synthetic, as it is elsewhere in human discourse. Implications are generally not material ones. He sees number theory as based on the principle of complete induction, and hence as a theory fundamentally synthetic in character; he understands the distinction between the potential and actual infinite, and sees it as a paradigm of the mental power to grasp concepts. He founds his study of analysis on the second-order axiom of completeness and grasps that arguments in analysis use inferences that cannot be formalised in first-order logic; he revels in the incompleteness of second-order logic, for it opens up the endless possibility of mathematical discoveries, never complete, and hence a potentially infinite source for his creativity and delight; he considers that mathematical discourse is akin to philosophical discourse, and considers the possibility that no single theory of collections could embrace the reality that he studies; in particular, he considers that the continuum, with its fundamental property of extension, is itself a primitive.

And in all this, he pays scant regard to whether there is a material basis to his activity, or whether human consciousness is the product of material forces; for the rationalist, the mind is the equal partner of prima materia in the relation of mind to reality, and the latter cannot be comprehended without the mediation of the concepts provided by the former.

The Turing Test

The Turing test arises in the context of the epistemology of other minds – for how do I know that any other given person is conscious? An empiricist will say, by behaviour alone – for there is nothing else for an empiricist to "observe" than behaviour. But a rationalist is not constrained to answer this question in the manner of an empiricist. He might adduce other considerations, even direct intuition. Be that as it may, even for a rationalist the production of a machine that could behave in such a manner that it could fool another person into thinking it was human, would be a considerable advance for the mechanist position.

It is now nearly seventy years since Turing first made his prediction about computers:

> I believe that in about fifty years' time it will be possible to program computers, with a storage capacity of about 10^9, to make them play the imitation game so well that an average interrogator will not have more than 70 per cent chance of making the right identification after five minutes of questioning. The original question, "Can machines think?" I believe to be too meaningless to deserve discussion. Nevertheless, I believe that at the end of the century the use of words and general educated opinion will have altered so much that one will be able to speak of machines thinking without expecting to be contradicted. I believe that no useful purpose is served by concealing these beliefs.[27]

Yet still we have no machine capable of fooling a living person. We receive frequent reports from the media of the annual attempt for a machine to breach this barrier, and promises that, with the exponential increase in computing capacity, this vital barrier will be breached in the imminent future. Certain successes are lauded, such as the ability of a computer to out-play Kasparov, break-throughs in playing the game Go, and successes in quiz programs. Algorithms to imitate conversations undergo development.

Will it be possible soon that some of these algorithms may be able to "hold conversations" for some minutes at least with human interlocuters? If so, what would a non-mechanist think? Some observations about the Turing Test need to be added.

1. The machine is practising an act of deception. This is akin to any other act of deception. If I go to my Bank Manager and he deceives me into making imprudent investments, then he is a liar; and my being deceived by him, does not make him any less a liar. If a person flatters another person with a view to gaining sexual favours, then he is a liar. Thus, the Turing Test must be placed into the context of deception in general. So, for example, I might hold a conversation via the telephone with an agent, and be convinced for half an hour, a day, a month, or even years that that agent is a living and conscious human being, only to discover later that it was a machine after all. Deception can take place over many years. A one-off success in the Turing Test is not decisive. To be rigorous, the computer must deceive all the people all the time; even partial success over many years will not be sufficient to compel a non-mechanist to reconsider his position.

2. The rigour of the test must also be considered in another respect. Simple conversations will not be sufficient; ordering a cup of coffee in a restaurant according to the "restaurant script" is not a rigorous enough test, for it is only in the deepest conversations and the profound acts of creation that the human mind is fully expressed. We must not set the bar too low. In

> other words, to "pass" the Turing Test, the machine must write a work equivalent to one of Shakespeare, paint a painting like Rembrandt, compose music like Bach and/or produce mathematical theorems and insights like Lebesgue, or equivalent. And this is, after all, the myth that is being propagated by the science-fiction of our contemporary culture – for what do all these stories tell us? – but that, at some time soon, artificial intelligence will surpass that of human intelligence? If that be so, then let the machines compose the novels, the works of art, the music and the science.

The rationalist perspective that I have sketched here in no way suggests that human intelligence is of such a nature that machines can copy it and exceed it. Human thinking in mathematics does not in its fullest sense even remotely appear to be like that of first-order logic, which is an analytic fragment of all inference whatsoever, and yet the only reasoning that a computer is capable of. A rationalist has no reason at all to suppose that artificial intelligence will be knocking out solutions, for instance, to the continuum problem, or proposing new fixed point theorems, or advancing new definitions of the integral – it is simply of an order beyond them.

So, what of the successes? The feature of those limited occasions wherein the computer out-performs the human, are that those systems are finite. Chess is a finite problem; so too is the playing of Go; if the knowledge delimited in a quiz program is finite, then an algorithm may beat a human contestant. What of it?

On this basis, the rationalist has no reason for believing a computer will pass the Turing Test, not once the full rigour of the test is instantiated.

What then, of the belief so strongly and widely held that the victory of artificial intelligence is just around the corner? What of the claim, for instance, that the whole universe is a computer simulation?

Logically, the statement, "A machine will pass the Turing Test" is in the same class as the claim, "The world will end soon." Shall we say that there is evidence for the claims about the Turing Test? The evidence for the forthcoming success of the Turing Test is based on the extrapolation of the capacities of computers. Extrapolations are notoriously fallible. Exponential growth has been known to hit a ceiling. In the case of artificial intelligence, the rationalist perspective on the capacities of the mind suggest to him at least that there is an insuperable barrier between the functions of a machine and the faculties of human intelligence – a barrier that no increase in the capacity of a computer could possibly breach – the extrapolation is not justified.

> Mechanist: A machine will be built that will pass the Turing Test, even the rigorous one.
> Rationalist: No such machine will ever be built.

Both are statements about the future, while grounded in reasons acknowledged as compelling by each side respectively, none of these reasons are sufficient to overpower the opponent. To the rationalist, it is not evidence for artificial intelligence now, to say that artificial intelligence in the future shall be better.

An intermediate position?

Is the position of Penrose, who is both a materialist and opposed to mechanism, tenable? That one could be an empiricist and non-mechanist seems to me to be entirely possible. This is because a non-materialist philosophy of science is tenable; it may be grounded in Kant. Such a view argues that empirical science identifies the regularities that exist within phenomena in empirical reality, but eschews the realist assumption that these regularities may be projected onto a reality of matter; hence, for the instrumentalist, there are only the regularities.

Rationalism does not preclude an empirical science of the mind, and parts of this science may involve mechanical models of the psyche, cybernetics and such like. For the rationalist, this would be the empirical reflection of the mind, on a principle akin to the relation between the Ego and Self in Kant's philosophy. For the rationalist, science is neutral as to metaphysics, and every empirical law that can be said to cohere with mechanistic materialism, can also be made to cohere with rationalism.

The question for Penrose is what role he assigns to the mind and consciousness in his philosophy. Is the mind the equal partner in the enterprise of understanding phenomena, or is it the product of material forces? As a materialist, he would seem to affirm the latter. As a physicist, he still subscribes to the principle of Galileo that mathematics is the language of reality. He seeks equations that will describe all phenomena, inclusive of the phenomena of the mind. If those equations are written in first-order language and use material implication, then his theory is a mechanical one. The problem arises when his equations are not written in that language, for he seems to be committed to some such view, when he advocates a variant of the Gödel argument. If the language he uses is second-order, then, with respect to the Quinian "to be is to be the value of a variable", he ascribes independent existence to concepts (Platonic realism), and hence brings the mind into equal partnership with matter. Such a theory is not materialist; it is dualist as to explanation.

The future?

Our eyes turn towards the future. As one sympathetic to rationalism, I acknowledge that should a machine be built that passes a rigorous Turing test, my understanding of human nature must be modified. But for the present my beliefs are not shaken, for I have as many reasons for believing no such machine will be built, as the mechanist can adduce for believing it will.

8,144 words in total, including references and footnotes.
7,884 words without references and footnotes.

References

Benacerraf, Paul [1967]: God, the Devil and Gödel, Monist (1967) 51 (1): 9-32.

Benacerraf, Paul and Putnam, Hilary, Eds. [1987], Philosophy of Mathematics, Selected readings. Second edition. Cambridge University Press, Cambridge.

Bishop, Errett [1967], Foundations of Constructive Analysis, McGraw-Hill, New York

Burton, David M [1976], Elementary Number Theory. Allyn and Bacon Inc. Boston.

Copeland, Jack [2008], The Mathematical Objection: Turing, Gödel, and Penrose on the Mind, July 2008

Curry, Haskell B [1954], Remarks on the definition and nature of mathematics. Original publication in

Henkin, Leon [1971], Monk, J. Donald, Tarski, Alfred: Cylindric Algebras – I, North-Holland

Hamming, R. W. [1998], Mathematics on a Distant Planet, American Mathematical Monthly. Vol. 105. No.7

Kant, Immanuel [1781], Critique of Pure Reason. Translated Smith, Normal Kent [1929]. MacMillan.

Lucas, J. R. [1998], The Implications of Gödel's Theorem, Talk given to the Sigma Club

Lucas, J. R. [1968], Satan Stultified: A Rejoinder to Paul Benacerraf, Monist (1968) 52 (1): 145-158, Dialectica, 8, 1954, 228 – 33. Reprinted in Benacerraf and Putnam [1987] p 202 – 206.

Lucas, J. R. [1996], The Gödelian Argument, http://www.leaderu.com/truth/2truth08.html.

Poincaré, Henri [1996], Science and Method. Trans. Andrew Pyle. Routledge, London 1996.

Potter, Michael [2004], Set Theory and its Philosophy – A Critical Introduction. Oxford University Press.

Shapiro, Stewart [2000], Foundations without Foundationalism, A Case for Second-Order Logic. Oxford University Press.

Turing, Alan [1950], Computing Machinery and Intelligence. Mind LXI (236), p. 433 – 460.

Väänänen [2001]. Second-order Logic and Foundations of Mathematics, Bulletin of Symbolic Logic. 7 (4): 504–520.

Wolf, Robert S [2005], A Tour through Mathematical Logic. The Mathematical Association of America.

[1] Wolf [2005] p.29

[2] Lucas [1998]

[3] Lucas [1968]

[4] Lucas [1996]

[5] For example, "… what is it that Gödel I precludes the machine (let's call her Maud) from doing? Evidently, it is to prove H (her Gödel formula) from her axioms according to her rules. But can Lucas do that? Just as evidently not." – Paul Benacerraf [1967]

[6] For example, "The learning mind successively mutates from one theorem-proving Turing machine into another." Jack Copeland [2008]

[7] Kant's Third Antinomy of The Critique of Pure Reason. A dialectical conflict may manifest itself as an antinomy. Kant's resolution, that determinism applies to phenomena that are experienced in time, whereas freewill is a property of the transcendental self, which is timeless, stands outside the time order, and thereby not a product of any successive event taking place in the time order, was passed over by C20th empiricists.

[8] "According to formalism the central concept in mathematics is that of a formal system. Such a system is defined by a set of conventions ... we start with a list of elementary propositions, called axioms, which are true by definition, and then give rules of procedure by means of which further elementary theorems are derived. The proof of an elementary proposition then consists simply in showing that it satisfies the recursive definition of elementary theorem." Curry [1954], p. 203

[9] I distinguish Hilbert's program from that of the formalism represented by Curry. Hilbert sought finite consistency proofs of ideal postulates in mathematics, but he sought then in order to ground mathematics synthetically; hence, he was not a formalist in the sense of Curry.

[10] For example, the introduction of axes in integration creates the notion of a negative area. No area in physical reality could be negative, but the signed area in an integral means one that lies above or below the given axis.

[11] Wolf [2005] p.36.

[12] Set theory does not have a direct representative for the potential infinite, though the symbol $< \omega$ is seen and used. We see that the collection: $\{n : n < \omega\}$ is conceptually different from the collection ω; yet that is not a distinction made in the object-language of set theory itself.

[13] **Archimedean property**: If *a* and *b* are any integers, then there exists a positive integer *n* such that $na \geq b$. (Burton [1976] p.2) This implies that $\mathbb{N}$ is not bounded above.

[14] For ω the following statements are equivalent: (1) ω is a limit ordinal; (2) $(\forall n)(n < \omega \supset n + 1 < \omega)$; (3) $\omega = \sup_{n<\omega} n$. Proof in Potter [2004] p. 181.

[15] Wolf [2005] p. 83.
[16] Analytic proof of the Archimedean property from the Completeness Axiom: Suppose $\mathbb{N}$ is bounded above. Then by the completeness axiom there exists a unique real number *u*, such that $u = \sup \mathbb{N}$. For any number $n \in \mathbb{N}$ the number $n+1 \in \mathbb{N}$, hence $n+1 \leq u$ and $n \leq u-1$. This is true for all $n \in \mathbb{N}$, hence $u-1$ is an upper bound for $\mathbb{N}$. This contradicts the uniqueness of *u*, so $\mathbb{N}$ cannot be bounded above.
[17] In his essay, *Mathematics and Logic,* Poincaré states that the principle of complete induction "appeared to me at once necessary to the mathematician and irreducible to logic." (Poincaré [1996] p. 148
[18] Shapiro [2000].
[19] "The Cantor diagonalization, which does some alteration in the objects being discussed, supposes that we can cope with the actual infinity, and uses a much weaker degree of self-reference than Turing, hence falls in between the two, and I am ambiguous as to my belief in the safety of relying on the result." R.W. Hamming [1998]. This is a polite statement of his doubts; earlier in the paper he writes, "I am inclined to believe that they would have confined Cantor in his old age to an insane asylum."
[20] Väänänen [2001] is a summary of this situation. The abstract of this paper states his conclusion: "… if second-order logic is understood in its full semantics capable of characterizing categorically central mathematical concepts, it relies entirely on informal reasoning. On the other hand, if it is given a weak semantics, it loses its power in expressing concepts categorically." By weak semantics he indicates a second-order logic with the Henkin semantics. In view of this paper I argue that second-order logic with full semantics, where the quantifiers range over all sets, is a larger system, not equivalent to first-order logic. However, a rationalist might also not concede that second-order quantifiers have a semantics in either sense; while some second-order statements range over sets, it is contestable whether the meaning of second-order statements is so given, because semantics for a rationalist is not always reducible to syntax, and the implications are not always material implications.
[21] Concluding remarks of the Preface to *Cylindric Algebras I* by Henkin, Monk and Tarski. Henkin [1971].
[22] "The completeness property – that every set of reals with an upper bound has a least upper bound – is unavoidably second-order." Wolf [2005] p.43.
[23] Contemporary constructivists oppose classical first-order logic by asserting that the law of excluded middle as a general principle is invalid; from a technical point-of-view, constructivist models are sub-lattices of those of classical logic, distributive lattices in general as opposed to complete, Boolean lattices. Bishop's constructivism is consistent with mechanism. Furthermore, the second-order axiom of completeness cannot be framed in classical first-order language either. Bishop's work exemplifies a problem for mechanists in general. His response is to acknowledge that the historical analysis of the C19th is non-mechanical, and to attempt to reconstruct analysis so that it is consistent with mechanism.
[24] "Unless there exists a general method M that produces such a computer program corresponding to each bounded constructively given sequence of rational numbers, we are not justified, by constructive standards, in asserting that each of the sequences has a least upper bound." Bishop [1967] p.4.
[25] The following are equivalent statements of the Completeness Axiom: (1) **Bolzano-Weierstrass theorem:** Every infinite bounded subset has a limit point in the set. In its original formulation, this was expressed as: Every bounded sequence in Euclidean space $\mathbb{R}^n$ has a convergent subsequence. (2) **Cauchy convergence criterion:** Let *S* by a non-empty subset of $\mathbb{R}$. Every Cauchy sequence on *S* converges to a real point in *S*. (3) **Dedekind completeness axiom:** Any non-empty subset of $\mathbb{R}$ which is bounded above has a least upper bound in the set. (4) **Cantor's nested interval principle:** Given any nested sequence of closed intervals in $\mathbb{R}$, $[a_1, b_1] \supseteq [a_2, b_2] \supseteq \ldots \supseteq [a_n, b_n] \supseteq \ldots$ there is at least one real number contained in all these intervals, $\bigcap_{n=1}^{\infty} [a_n, b_n] \neq \varnothing$ bound in the set. (5) **Heine-Borel theorem**: Let *X* be a closed, bounded set on the real line $\mathbb{R}$. Then every collection of open subsets of $\mathbb{R}$ whose union contains *X* has a finite subclass whose union also contains *X*.
[26] This is the predicament of Hamming, who lampoons his colleagues, whom he calls Platonists. See Hamming [1998].
[27] From Turing [1950].