COMMENTS ON THE FOUNDATIONS OF MATHEMATICS [Penrose1.doc]

Introductory Remarks

These notes have been put together in support of my attempts to come to terms with that most recalcitrant and mysterious of the mind/brain's accomplishments -those of creativity, true learning and progressive self-organization. Explorations of the foundations of mathematics -in particular the classic papers of Gödel and Turing- clearly need to be carefully looked at in this connection. An open question is this: Should it be possible to make a firm case that creativity in mathematics is dependant upon a 'Platonic Pipeline' or its equivalent, is it to be inferred that more general forms of creative endeavour -in the arts, philosophy, engineering etc -and for that matter, the kind of moment-to-moment acts of understanding which would seem to be called for in casual conversations between friends- are of the same *formal* nature? To suggest -as does Penrose- that they are is to imply that the 'Calculus Ratiocinator' of Leibniz -brought into full flower in the logic of Frege-Russell- can be extended to capture thought executed within the 'soft formalism' of natural language. Leibniz at one time believed that it might be possible, dubbing the extension of rigorous logic into natural language the 'Characteristica Universalis' though later in life came to entertain doubts about the idea.

These topics follow these introductory remarks:

O First & Second Order Systems
O Undecidability (in its Various Meanings)
O Solvability
O Avenues of Approach to Problem Solving
O The Proof-Process in Practice
O Gödel's Discovery

    Preliminary Remarks.
    Outline of Gödel's Proof.
    The Turing Shortcut

Despite my best efforts, I never quite managed to get through Gödel's proof process -even in outline. However, I seem to have much less difficulty with Turing's 'Halting' version of the theorem.

First- & Second-Order Systems

First-order systems are those in which all true conjectures are *theorems*, hence can be derived from the axiomatic basis through the process of Peanoese induction. As a consequence, the proof of any conjecture within the system can be ascertained through an *automated* process; this is equivalent to what is meant by saying that the system is strictly formal. If the axiomatic basis is well-formed, i.e. the axioms are consistent and complete (and, for that matter, mutually independent) then undecidability never arises. The entire system is tautological in the sense that there is nothing in the theorems which isn't already latent within the axioms:

> "......The name tautology.....was introduced by one of the founding fathers of linguistic analysis- Ludwig Wittgenstein, who argued in the Tractatus Logico-philosophicus (1921) that all necessary propositions are tautologies and that there is therefore a sense in which all necessary propositions say the same thing -viz.nothing at all." Encyc Brit 11, p.501

> ".......By a formal 'proof' (or 'demonstration') we shall mean a finite sequence of formulas, each of which is either an axiom or can be derived from preceding formulas in the sequence of transformational rules (Nagel & Newman 1958 p.46) -that is, if all the axioms have the property, any formula properly derived from them by the Transformational Rules must also have it." (Ibid., p.51)

It would, however, be an overstatement to say that there can be no surprises in strictly formal (that is, first order) systems; most of us, in our first encounter with the Euclidean archetype of such systems, were charmed by the way in which simple but putatively unlikely truths could reside within complex figures.

If such a system were to be inconsistent, then it would be possible to demonstrate that *any* well-formed statement in the system were true, that is to say derivable from the axioms. In particular, one could establish that conjecture C and also its inverse Not-C were both true. Just as a liar can 'prove' anything, so unlimited falsehood can propagate upwards from a corrupt axiomatic substratum. In consequence, one has only to demonstrate that *one* conjectured truth cannot be derived from the axioms to prove their consistency. This contrasts with the inference to be drawn when dealing with second-order systems. In this latter case, a true statement not strictly demonstrable does not prove the authenticity of the axiom set, instead merely proclaiming itself to be undecidable -given only that the system be consistent..

Euclid's geometry, in its original form, was, in fact, lacking in consistency, as instanced by the way in which demonstrations can be contrived to 'prove that all triangles are isosceles, or that an obtuse angle is equal to a right angle. In such cases, the axioms do not protect you from making false assumptions in the construction of complex figures. Even the 'proof' of the famous theorem that two triangles are congruent if they share two sides and the included angle, is in fact on very shaky ground. Hence Bertrand Russell's famous aside that euclidean geometry consisted of a 'set of fallacious proofs for obvious truisms'. In his reformulation of the system's definition, Hilbert demonstrated that a substantial expansion of its axiom set was required, as shown below, if the system were to be cleansed of inconsistency:

7 Axioms of Connection
5 Axioms of Order
1 Axiom of Parallels
6 Axioms of Congruence
1.Axiom of Continuity

One of the axioms is specifically addressed to shoring up the two-sides-and-included-angle theorem.

It was David Hilbert's fond belief that properly grounded and conceived, the whole of mathematics could be shown to be first-order in the sense discussed above. It was not until the advent of Gödel, Church and Turing. that this was shown to be an impossible dream. All systems as complicated as arithmetic are now known to be incorrigibly incomplete. Most of the more interesting discoveries within arithmetic do not have the status of theorems, and hence must necessarily have had their truth established by methods which transcended Peanoese induction. In consequence, radical new discoveries and new proof-processes must advance hand in hand. Such new truths can be made theorems by fiat, that is to say, added to the original axiomatic set, just as newly invented proof procedures can supervene upon mathematical induction. Though it might be possible to reconstitute the axiomatic base by processes more subtle than the mere addition of 'undecidables', no amount of augmentation of truths or algorithms can serve to make the basis of the system complete and self-sufficient. Penrose (1995) makes these comments on the subject:

> "......One might imagine that it would be possible to list all possible 'obvious' steps of reasoning once and for all, so that from then on everything could be reduced to computation. -i.e. the mere manipulation of these obvious steps. What Gödel's argument shows....is that this is not possible. There is no way of eliminating the need for 'obvious' understandings. Thus, mathematical understanding cannot be reduced to blind computation." p.56
>
> "......Some readers may....be puzzled as to why they [the Peano mathematical specifications] do not define the natural numbers adequately.....While it is true that the Peano mathematical specifications that are given for the successor operator S do in fact characterize the ordering relationships that distinguish the natural numbers from any kind of 'supernatural' numbers e.g. π, those specifications are not captured in terms of the formal rules that these quantifiers $\forall$ and $\exists$ satisfy. In order to capture the meanings of the Peano mathematical specifications, we need to pass to what

> is known as 'second order logic' in which quantifiers like ∀and ∃ are introduced but where now their ranges are over (infinite) *sets* of natural numbers rather than just single natural numbers. In the 'first order logic' of Peano arithmetic, the quantifiers would simply range over single numbers, and we get a formal system in the ordinary sense. But second order logic does not provide us with a formal system. For a strict formal system, it has to be a purely *mechanical* (i.e. algorithmic) matter to decide whether or not the rules have been correctly applied...." p.109/110

Second-order systems -again, the quintessential example being that of arithmetic- unlike their junior first-order colleagues then, are top-down incomplete, and essentially so. Perhaps somewhat surprisingly, its pristine axiomatic grounding as supplied in modern form by Peano(1899)- is more parsimonious than that of Euclidean geometry (as reformulated by Hilbert):

1. Zero is a number
2. The immediate successor of a number is a number
3. Zero is not the immediate successor of a number
4. No two numbers have the same successor.
5. Any property belonging to zero, and also to the immediate successor of every number that has the property, belongs to all numbers [The Principle of Mathematical Induction]

Despite the vast undecidable domain hovering 'upstairs' beyond its grasp, this axiom set enjoys its own kind of completeness. It is the completeness of *ordinal* arithmetic, as the fifth axiom makes clear, whereas undecidable truths extend the discussion to include *cardinals*. In other words, while arithmetic is *ordinally* complete it is *cardinally* undecidable.

We are indeed free to add undecidable propositions as supplementary axioms to the grounding gang-of-five -or perhaps more appropriate would be the addition of the proof-procedures, stated in general form, which had to be devised to cut the mustard. But in so doing, we are, so to speak, adding apples to oranges. One might say that the distinction between the two is that between axioms which are *obligatory* and those which are *facultative*.

It is of interest to note that arithmetic can be made strictly formal if certain limitations are imposed upon the kinds of operations which are to be permitted, and whether or not sets i.e. propositional functions) are to be allowed. Two examples are Pressburger arithmetic (limited to (variable + variable) statements -no sets allowed), and S1S arithmetic which allow only (variable + numeral) statements, though sets *are* permitted. Evidently, then, the presence of sets doesn't, in itself, categorically rule out demonstrability within-the-system, provided only that the operations permitted be sufficiently circumscribed.

The Prime architect of the second-order predicate calculus was Gottlob Frege, his seminal work, the 'Begriffsschrift' coming out in 1879. Among the encomiums which Frege has received were those of Quine: "....Logic is an old subject, and since 1879 has been a great one" and De Long, naming Frege "...the greatest Logician of all time". Frege amalgamated propositional and predicate logics through his introduction of the propositional function, which allowed functions to be substituted for variables p,q,r. By making propositional functions variables, he raised symbolic logic up to a higher level of subtlety, seemingly, in effect bringing Leibniz's dream of a 'Calculus Ratiocinator' into existence (but not Leibniz's 'Charactersitica Universalis' aimed at a complete formalization of natural language). He was, however, not the author of the inter-formal mapping technique -going back to Descarte's unification of algebra & geometry, nor of metamathematics (first introduced by Hilbert in 1905 and elaborated in 1918) nor, of course, of the mapping of meta-mathematics back into object-level mathematics, so prominently featured by Gödel in the development of his renowned theorem.

Russell & Whitehead's 'Principia Mathematica' addressed the same domain from a different point of view, their aim being to reduce the whole of mathematics, and not just arithmetic to logic. In the process of so doing, Russell discovered that the second order predicate calculus, (probably) unlike arithmetic was not consistent. The dilemma surfaced in connection with self-referencing sets resulting in a number of antinomies some of which have become household words -e.g. the 'barber' paradox,(the barber shaves all who do not shave themselves -so who shaves the barber?); 'This Sentence is False'; or more generally:

O A class is normal if it does not contain itself as a member.

O Is the class of all normal classes (N) normal or non-normal?

1.If N is normal, it is, by definition, a member of itself
2.But if so, it is non-normal because containing itself as a member
3.Hence it is normal only if it is non-normal

Russell drew this to Frege's attention in a letter written in 1902, just at the point when Frege was about to publish his life's work on the foundations of mathematics; Frege makes a rueful reference to this in an appendix to the work:

> ".....Hardly anything more unwelcome can befall a scientific writer than that one of the foundations of his edifice be shaken after the work is finished. I have been placed in this position by a letter from Mr Bertrand Russell." Penrose 1994 p.138

The advance, then, from arithmetic to Frege-Russell set theory was bought at the cost of the loss of consistency. One solution -that resorted to by Russell- was that of radical surgery. Through his 'Theory of Types' (involving a nest of metalinguistic levels) he was able to banish self-referencing sets. This is a high price to pay when it be regarded that a great deal of self-reference is non-pathological -e.g. "This sentence is written in English".

Other attempts were made to reformulate set theory to repair this inconsistency, notably that of Zermelo-Fraenkel. Its axiomatic grounding appears surrealistic by comparison with that of Peano's arithmetic:

> ".....the standard axiomatic system for set theory -the Zermelo-Fraenkel system.......possesses infinitely many axioms expressed in terms of structures called 'axiom schemata'. By an appropriate reformulation, the ZF system can be re-expressed so that the number of actual axioms becomes finite.: Penrose (1995 p.135)

That is to say, the infinite set of axioms can be generated by a finite set of computational rules (ibid p.88)

It might seem that with the arrival of ZF set theory, our house of mathematics has finally been put in order. Unfortunately, this was not to be the case. Cohen (1966) demonstrated that the 'Axiom of Choice' and Cantor's 'Continuum Hypothesis' are independent of the system:

O The 'Axiom of Choice' asserts that for any given collection of non-empty sets, another set is to be found which contains precisely one element from each of the sets in question.

O The 'Continuum Hypothesis' asserts that the totality of all possible subsets of natural numbers is the next largest infinity after that of the natural numbers themselves.

In other words, Set Theory has been cleft into two formal domains -into Cantorian (that upholding the continuum hypothesis) and non-Cantorian Set Theory. This 'undecidability' disclosed within the ZM Theory is quite different to the 'top down' variety encountered in arithmetic, being more like what may be envisaged with respect to the fifth postulate in Euclidean geometry. In its absence, geometry sunders into three branches -we can select elliptical flat or hyperbolic by adding the corresponding completion to the other four axioms. In other words, the Axiom of Choice and the Continuum Hypothesis are neither necessary nor impossible; formally speaking they stand, as accidents which may be included or excluded by fiat.

An interesting question is -where do we go from here? We can 'fix' a 4-axiom geometry by adding the parallel line postulate, but the gap presently under discussion is not to be filled in such cavalier fashion. Mathematicians are not of one mind over the matter. For those among them who believe ZM set theory contains the whole of mathematics, nothing, by definition, is left outside to be conceived or discovered. From this point of view, the undecidability in question acquires an absolute status, that is to say, the continuum hypothesis and the axiom of choice are, in themselves, neither true nor false. It would seem to me that to take such a position is not absolutely to deny Platonism, but to encourage such a stance by implying that it is unable to uphold the whole of mathematics, running out, so to speak, when we reach ZF set theory. Certainly for those who wish to take the position that the whole of mathematics is but a game which is concerned, not with 'truth', but only with what follows from what, ZF incompleteness offers the best place for them to get their foot in the door. One can, they insist, change the rules in any way which one pleases, e.g. change the Peano foundation of arithmetic to obtain a mutated discipline from which quite different 'truths' will emerge, but which we must accord the same status as those given us by Peano arithmetic. What this line of reasoning fails to account for is why the 'arithmetic' of common knowledge has such extra-ordinary resonance aesthetic depth -which latter is such as to commend itself even to the non-mathematicians such as myself. Of course, no arguments can really drive the matter home; we can only say that its ultimate a question of taste; every argument can be countered by one inclining in the reverse direction. Thus the mathematical 'relativist' can respond to the 'aesthetic' rejoinder above by asserting that the Darwinian process automatically selects our sensitivities in this regard, so that we may successfully engage those formalisms upon which the Real World appears to be grounded.

For those who accept the Platonic position that set theory is either Cantorian or it is not, there still remains the open question of the possibility of resolution. If ZF set theory is truly the whole of mathematics, then it is difficult to see where one could find the higher ground from which superior viewpoint one might scrutinize the ZF system and hence, hopefully, reconstitute it so that the dilemma vanishes. I think that the Platonist would therefore believe that such an advance of understanding is possible in principle, though he could legitimately doubt that human minds are actually up to it, and we may have to wait for the emergence of supermen or their equivalents. I suspect that the typical Platonist might hesitate to acknowledge such human limitations, perhaps on the intuitive grounds that mathematics 'hangs together' so that nothing is quite separate from anything else, and that *something* which lies beyond *articulate* grasp may yet be *hinted at*. Cohen's views on the subject are of great interest:

> ".....In the final chapter of his 1966 book, Cohen makes the point that although he has shown that the continuum hypothesis is undecidable according to the procedures of ZF, he has left untouched the question of whether or not it is actually *true* -and he discusses how one might go about *deciding* this question! This makes clear that he is *not* taking the view that it is an entirely arbitrary matter whether one accepts the continuum hypothesis or not.....By these remarks, Cohen reveals himself to be, like Gödel, a true Platonist" Penrose 1995 p 116

The following schema summarises all of the above levels formalism which have been discussed above: 'Axiomatic Grounding', once more, denotes those axioms taken to be essential and is not to be taken to imply completeness in the wider sense.

	Axiomatic Grounding		True
System	Complete?	Consistent?	Undecidability?
O Euclidean Geometry	Yes	Yes	No
O Ditto, Minus 5th Postulate	No	Yes	No
O 'Ordinal' Arithmetics	Yes	Yes	No
O Cardinal Arithmetic	Yes	Yes	Yes
O Frege-Russell Set Theory	Yes	No	Yes

O Zermelo-Fraenkel Set Theory	No	Yes	Yes
O [Reformulated ZF Set Theory?]	[Yes?]	[Yes?]	[Yes?]

I'm not sure, once again, that the Platonist would expect reformulated ZF to end the matter for all time; fresh antinomies would be expected to arise, to be fixed in their turn -without end. I suspect that the platonist would take the PLatonic realm to be bottomless.

<u>Undecidabily</u>

'Undecidable' as used in mathematics, has become someone slippery term whose meaning wanders at times from that originally given to it by Gödel and Turing.

In its classical meaning, it refers to those true propositions within a system which cannot be derived from the set of axioms upon which it is grounded. It only arises in systems at least as complex as arithmetic, when conjectures which refer to sets as well as variables be allowed -e.g. the set of prime numbers. True propositions within arithmetic are overwhelmingly undecidable, i.e. not demonstrable through Peanoese mathematical induction. However, such newly discovered truths can, by fiat, be added to the system, thus upgrading their status to that of theorems. This being done, its formal power will be raised, because an indefinite number of newly proposed conjectures will thereby be brought within reach of demonstration through appropriately constituted algoriths.

As we now know from Gödel's theorem, however, all such systems are incorrigibly incomplete because no finite expansion of the axiomatic basis can serve to achieve closure.

Systems such as Euclidean Geometry and first-order predicate calculus are complete, which means that all truths within the system are either axioms or theorems. Questions of undecidability, therefore, simply do not arise. However, it can arise in such first order systems if the axiomatic basis is malformed or insufficiently defined. Thus, consider Euclidean geometry minus the firth axiom. Let us start with a Euclidean geometry in which the fifth postulate had been omitted. What, in fact, would be the effects of its omission? An immediate consequence would be a softening and loss of definition whereby many theorems of Euclid would simply go by the board. For example, it would no longer follow that the three angles of a triangle would add up to 2 right angles, but could be anything from (asymptotically) zero to four right angles. What this would imply that the geometry was no longer necessarily flat, but could be positively or negatively curved, giving rise, respectively, to spherical and hyperbolic geometries; ordinary flat Euclidean geometry would then emerge as a kind of singular or degenerate case hovering over the point of division of oppositely intended curvatures.

[It is of some interest to speculate upon why the concept of space-curvature eluded the Greeks. The possibility of curvature, as normally conceived and encountered arises whenever the space of objects is of lower dimensionality than the space of dispositions. with respect to circles, the dimensional offset is two, and to spheres, is one. But the space of dispositions -in which we are living- is 'obviously' locally flat, and the notion that in the very large or the very small it might cease to be the case somehow never came up for consideration. Perhaps the notion of flatness appealed to the Greek sense of perfection, just as the sphere was taken to be a perfect solid. One imagines that the Greeks must have had convex mirrors, but apparently the spectacle of the world which they reveal failed to evoke the notion of hyperbolic space. I can appreciate Greek limitations in that I find it hard to conceive of space curvature without the addition of an extra spatial dimension within which such curvature may be expressed.]

As a second example, consider a 'constructional' geometry consisting of Euclid's five postulates plus two axiomatic tools -a compass and a straight-edge. Almost any figure can be constructed satisfactorily. However, three operations which cannot be performed are (1)the trisection of the angle, (2)the doubling of the cube, and (3) The squaring of the circle. The reason is that these operations involve irrational numbers which would require an infinity of iterations to be contended with. However, as with the addition of the fifth postulate, this problem may be made to vanish if we add a 'postulate' of a mark somewhere on the straight edge indicating a known distance. (Gerver (1997) suspects that something other than a mark would be necessary in squaring the circle in view of the appearance of the transcendental π.

In these examples, a triangle with angles summing to 2 right angles, a trisected angle, a doubled cube, and a squared circle are neither necessary nor impossible. While not derivable from the axioms, they can yet exist as 'accidents'. There is really no absolute distinction here between those 'accidents' already resident within well-formed euclidean geometry. For example, size enters nowhere into the axiomatic grounding so that the scale of figures is arbitrary and open -as is also the proportional shaping and scaling of triangles, polygons, etc. Indeed, such parameters simply provide the legitimate envelope of expression which it is the raison d'être of the system to supply. The effect of relaxing postulates is to enlarge the envelope of expression; whether this is good or bad depends upon the matter at hand. With respect to Euclid, such is obviously bad to the extent that on parochial scales, the space 'out there' appears to be both flat and uniform.

Notice that if, instead of discarding the fifth postulate, we make of it a kind of 'axiomatic parameter' whose value can be set to obtain both an elliptic and hyperbolic space of any desired degree of curvature, then we get spaces of equal coherence and firmness of definition as the euclidean, further, with it being possible to map every theorem in Euclid into a corresponding theorem in the curved spaces. An interesting feature of both spaces is the way in which the proportions of figures are no longer independent of their absolute size. In both cases, sufficiently scaled-down figures become correspondingly indistinguishable from their euclidean counterparts.

A further instance of 'first-order' incompleteness -that encountered within Zermelo-Fraenkel set theory- will be discussed [further below]

Gerver (1997) makes this distinction between the two kinds of incompleteness -those systems which can be fixed by the addition of a finite number of axioms, and those, like arithmetic which are incorrigibly incomplete.

We may also speak of 'absolute undecidability' -a matter which will be examined in the paragraphs which follow.

Provability

For a start, everyone believes that arithmetic is consistent, hence there can be no question of any conjecture being both provable true and provably false.

This having been said, there are degrees of provability, as the following sequence summarizes:

O No Proof Possible -even in principle
O Proof possible in principle:

    1.But only open to enumerative search
    2.Analytical solutions also possible in principle

        a) Demonstrable from the founding axioms
        b) Provable within the augmented system
        c) Proof Process demands novel insight.

    3.Analytical Proof possible in fact.

A conjecture which is in principle beyond the reach of any conceivable proof-process might be said to be "absolutely undecidable." *Are* there conjectures which are so disposed? No one knows for sure; Penrose speaks of it as a matter of taste or intuitive appeal:

> "....Mostowski makes it clear......that arguments such as Gödel's have no bearing on the question of whether there might exist *absolutely* undecidable mathematical questions. The issue must be

> regarded as being completely open, as of now, as far as what can be proved or disproved. This question....remains a matter of faith." Penrose (1995) p.421

De Long (1971 p.60) suggests that the conjecture 'x is the divisor of a Perfect Number' [i.e., one which is the sum of its proper divisors] might be recursively non-solvable. [Is the sticking point here that whether or not there *is* a highest perfect number is itself unknown?]

Presumably, if one were Penrose's sort of Platonist, one would hesitate to put any mathematical truth beyond the reach of mind, though perhaps demanding minds more powerful than have yet emerged. Implied here is a belief that mathematical truth in its Platonic entirety 'hangs together' so that there is no absolute demarcation between truths which are mind-reachable and those which are not. That is to say, no matter how recondite any topic might seem to be, reflections on the matter should be productive of useful 'hints' -no matter how poorly focussed these may be.

Next, no analytical proof may be possible in principle, but the conjecture is open to an enumerative search. Hilbert's 'tenth problem' would appear to be an instance of omega-undecidability [??]. The problem is that of finding an analytical proof of determining whether a set of diophantine equations have a common solution. This problem, one of a list proposed by Hilbert in 1900 for solution during the coming century had to wait until 1970 for a definitive solution -that there is no way of deciding which sets of diophantine equations have a common solution. Likewise, it is not possible to determine in general whether a set of polygonal shapes will tile the plane, that is, that no gaps or overlaps will ever turn up. Penrose (1994) p.31

Finally, a proof process which is possible in principle might be ruled out in practice because of exorbitant computer demands. De Long (1971) notes:

> "......It might be that Goldbach's conjecture [that every even number greater than 2 is the sum of two primes] is true, but that the shortest proof in a given system would require more pounds of ink than there are atoms in the universe. Or, if the conjecture is false, the first even number which is not the sum of two primes may be so large that the chance of discovering it is virtually nil." Scientific American March 1971

The Solution Process

The prime differentiation here is between proof-processes which are algorizable -i.e. can be automated- and those which are not, requiring the intrusion of novel insights, in which the mathematician must come up with one or more 'bright ideas'.

Automation is always possible -at least in principle- for all conjectures within first order systems, and for those which *are* reachable from the axiomatic basis of second-order systems. Only a small fraction of these latter will be Peanoese-provable; rather, most will depend upon a whole armamentarium of algorithms which have been devised and discovered over the ages.

As will be explored further below, to say that automation is possible, is not at all to say that it is attractive or efficient, nor should it be implied that inductive proofs are simpler than those which are not.

Those conjectures remaining outside and beyond the (augmented) axiomatic floor of all second-order systems will be incapable of any automatable solution, i.e. will be non-algorithmic with respect to the resources at hand. Some new insights or dodges will be needed which must seemingly be derived de novo. On the surface, such a conclusion would seem to follow from the Gödel-Church-Turing theses. Something needs, willy nilly, to be produced from *nowhere* -which, of course, is one of the circumstances inclining most mathematicians towards Platonism; the 'nothing' in question is too damned fecund, and there is something suspiciously coincidental about the way in which the same mathematical discoveries may be made by persons mutually incommunicado at the time. Whatever that process may be, it is evident that at least:

> "......Human mathematicians are not using a knowably sound algorithm in order to ascertain mathematical truth........ Conscious mathematical understanding cannot be modelled at all in computational terms -whether top-down, bottom-up or any combination of the two". Penrose 1995 p.76

Some for whom such a non-reductionist appeal goes strongly against the grain, have sought ways of dodging what, to them is an unwelcome inference. Penrose gives a detailed account and analysis of the many loopholes which have been suggested and explored, and claims to have successfully fielded them all. Here are a few special interest:

- O An Appeal to Chaos (Particularly to the 'Edge of Chaos'); does it point to something deeper?

- O The Postulation of a horrendously complicated but 'unknowable' Algorithm of Darwinian (or even Divine) Origin, which can *simulate* (or, more correctly, perhaps *mimic*) mathematical intuition.

- O Taking an appropriately complex formal system as a point of departure, could it not be that a loop, or front end, attached to the formal substratum of the system could indefinitely enrich it by recursively adding itself as a new axiom? Continued indefinitely, wouldn't such an automatic process bring all conjectured truths (asymptotically) within reach of algorizable exploration? ..

- O The 'Monkeys at Typewriters' argument; given sufficient time, a purely random process could 'prove' even Gödel's Theorem itself..

- O Can the Theorem itself be trusted, given the shakiness of the foundations of mathematics? There is no agreement about the status of the ZF system which has been accepted, by default -for want of a better, so to speak- as the basement of the mansion of formalism, that is, the whole of logic/mathematics.

As explored at greater length under [add heading], Chaos is much in vogue, and is a concept beset by many misunderstandings. Although we are often surprised by the behaviour which appropriately put together systems may exhibit, the phenomena in question are strictly deterministic and also algorithmic -this latter in principle, though often being very far from being so, in fact. Attempts have been made to argue, nevertheless, that such systems are not fully understood, and that when sufficiently complex may exhibit unusual, 'para-algorithmic' behaviour. Most mathematicians, as opposed to AI votaries are, I believe, sceptical about this;

> ".....Can it be that it is *chaos* that provides the needed answer to the mystery of mentality? For this to be the case, there would have to be something completely new to be understood about the way in which chaotic systems can behave in appropriate situations. It would have to be the case that, in such situations, a chaotic system can closely approximate *non-computational behaviour* in some asymptotic limit -or something of this nature. No such demonstration has, to my knowledge, ever been provided. Yet it remains an interesting possibility ......
>
> Irrespective of this possibility, however, chaos would provide only a very doubtful loophole to the conclusion that we have arrived at...Penrose (1995) p.178

In other words, even though it might eventually be demonstrated that chaotic systems can behave in even odder ways than have so far shown up, there is little hope that such oddness would be of the *kind* needed to deliver the goods. If it does indeed break out of narrowly formal constraints, it gives little appearance of being 'usefully non-algorithmic' (penrose 1990)

The invocation of Chaos is but the most recent instance of an appeal to the emergence of novel possibilities, within a formal system, that is, of novel properties which are threshold-complexity dependent, that is, exhibit modes of behaviour which lie *absolutely* beyond the reach of less complex configurations within the system lying below the 'critical mass' in question. Perlis (Penrose 1990) calls this the 'extrapolation fallacy'.

> "......Niels Bohr and Max Delbruck said they felt that the explanation of living things would require the discovery of new physical principles about the organization of matter. Life turned out to be a matter of fantastic and unsuspected features of suitably complex chemical tinkertoys, however, explainable in terms of existing physics. The extrapolation from simple to complex molecules was false: simple ones do not reproduce or metabolize, but complex ones can. The same sort of phenomena may hold for mind, in ways we do not yet see." Perlis (1990)

It may be noted in passing that what molecular biology has actually demonstrated is *what* is going on at the molecular level, without having so far explained how it discharges its operations so speedily, efficiently, and largely without error, given only the contact forces which physics currently allows.

Lucas [ ] concedes that there just may be truly emergent novelty which is ‘crtical mass dependent, but if so, then other principles are at work whichmust imply a breakout from the reductionist straight-jacket constraining mainstream thought:

> ".......When we increase the complexity of our machines there may, perhaps, be surprises in store for us. (Turing) draws a parallel with a fission pile. Below a certain 'critical' size, nothing much happens: but above the critical size, the sparks begin to fly. So too, perhaps, with brains and machines. Most brains and all machines are at present 'sub critical' --they react to incoming stimuli in a stodgy and uninteresting way, have no ideas of their own, can produce only stock responses-but a few brains at present, and possibly some machines in the future, are super-critical and scintillate on their own account. Turing is suggesting that it is only a matter of complexity, and that above a cetain level of complexity a qualitative difference appears, so that 'super-critical' machines will be quite unlike the simple ones hitherto envisaged. “
>
> "This may be so. Complexity often does introduce qualitative differences. Although it sounds implausible, it might turn out that above a certain level of complexity, a machine ceased to be predictable, even in principle, and started doing things on its own account, or, to use a very revealing phrase, it might begin to have a mind of its own. It would begin to have a mind of its own when it was no longer entirely predictable and entirely docile, but was capable of doing things which we recognized as intelligent, and not just mistakes or random shots, but which we had not programmed into it. But then it would cease to be a machine, within the meaning of the act. What is at stake in the mechanist debate is not how minds are, or might be, brought into being, but how they operate.

To move on to the second line of defence; it has been vigorously argued that the insights gained by what appear to be the short-cuts of mathematical 'insight' are but the products of a horrendously complex, buried algorithm residing in the brain -having been bequeathed to us by the germplasm. Such an algorithm would by logical necessity have to be beyond the reach of direct inspection and explication:

> "......Humans do not ascertain mathematical truth by means of any knowable algorithm....The algorithm (or formal system) X, or at least the soundness: of X, would have to be something unknowable to human mathematicians. Otherwise, G(x) would be seen to be a mathematical truth although it is inaccessible to X: a contradiction." Penrose (1995)

To qualify for serious consideration, such a proposal would need to clear two formidable hurdles, at least. First, it is hardly conceivable that such an algorithm could arise through the neoDarwin process of filtered noise; to those however, who have no qualms about accepting the unredeemed promises of Darwinism, everything works out very nicely: it could hardly be otherwise.

> "......It is time to turn the burden of proof around, the way Darwin did when he challenged his critics to describe some other *way* -other than natural selection- in which all the wonders of nature could have arisen. Those who think the human mind is non-algorithmic should consider the hubris presupposed by that conviction. If Darwin's dangerous idea is right, an algorithmic process is powerful enough to design a nightingale and a tree. Should it be that much harder for an algorithmic process to write an ode to a nightingale or write a poem as lovely as a tree.? Surely Orgel's second rule is correct: Evolution is cleverer than you are" Dennett 1995.

(I find it to be profoundly ironic that Penrose should make an equally extravagant appeal to Darwinian resources, though on different grounds:

> "......It was not an *algorithm* X that was favoured in man, by natural selection, but this wonderful ability to understand" Penrose 1990)

The second objection is grounded upon the nature of the process of mathematical discovery as it is directly experienced by the questing mind; it bears a simplicity strongly at variance with its alleged horrendous workings; furthermore it seems to deliver altogether more than needed to secure biological survival -supposedly the reason for its emergence:

> ".....This seems to be totally at variance with what mathematicians seem *actually* to be doing when they express their arguments in terms that can (at least in principle) be broken down into assertions that are 'obvious', and agreed by all. I would regard it as far-fetched in the extreme to believe that it is *really* the horrendous and unknowable X, rather than these simple and obvious *ingredients* (emphasis added) that lies lurking behind all our mathematical understanding.....". Penrose 1990

> "......Any approximate algorithm for generating all the mathematical truths we know would have to be incredibly complicated (and in no way related to the kind of things creatures need to survive), whereas mathematical insights are, at root exceedingly simple." Penrose (1990)

As if it weren't already burdened to the breaking point, Dennett and others have heaped the final demand upon it -that it has contrived to make the very complicated process *seem* simpl*e* -in the interests of streamlining its operation.

It is not *logically* impossible for such a device to exist; this much was granted by the discover of the theorem which bears his name:

> ".....it remains possible that there may exist (and even be empirically discovered) a theorem-proving machine which in fact is equivalent to mathematical intuition, but cannot be *proved* to be so, nor proved to yield only *correct* theorems of finitary number theory.". Penrose 1995 p.128

So also Lucas:

> "......It is essential for the mechanist thesis that the mechanical model of the mind shall operate according to 'mechanical principles', that is, that we can understand the operation of the whole in terms of the operations of its parts, and the operation of each part either shall be determined by its initial state and the construction of the machine, or shall be a random choice between.determinate operations. If the mechanist produces a machine which is so complicated that this ceases to hold good for it, then it is no longer a machine for the purposes of our discussion." Lucas

This offers the protesting reductionist the one refuge from which he cannot be dislodged.

> "......Such a complicated algorithm [the horrendous X] would approximate the competence of the perfect understander, and be invisible to its beneficiary. Whenever we say we solved some problem 'by intuition', all that really means is *we don't know how* we solved it. The simplest way of modeling 'intuition' in a computer is simply denying the computer any access to its own inner workings." Dennett 1995

Penrose (1995) concludes with this comment on the matter:

> "......The situation is quite different once we allow 'understanding' to be a non-algorithmic quality. Then, it need not be something so complicated that it is unknowable or incomprehensible. Indeed, it could be much closer to 'what mathematicians think they are doing'. Understanding has the appearance of being a simple and commonsense quality. It is something difficult to define in any clear-cut way, but nevertheless is so familiar to us that it seems hard for us to accept that it might be a quality that cannot be properly simulated, even in principle by a computational procedure. Yet this is what I am contending." Penrose (1995) p.149

> "......AI researchers will ignore the arguments from Gödel's theorem at their peril. Mathematical thinking, although a tiny minority activity, is thought after all, and if that is demonstrably non-algorithmic, then non-algorithmic thought is shown to be possible. That is all we need from the argument" Penrose (1990)

While it is true that just about all mathematicians are Platonists, the acknowledgment of Transcendence thereby implied begins and ends with the truths of mathematics. Penrose's platonism is more robust coming much closer to an affirmation of universals more in keeping with what the author of the concept of universals had in mind:

> "......Language only works at all because of the underlying power of the mind that allows *precise sense* to be extracted from *imprecise language*" Penrose (1990) [my emphasis]

It is deeply ironic that virtually all of the votaries of the 'hidden horrendous algorithm' have sought to stand their ground burdened by a physicalism demanding that the germplasm and brain alike must deliver the goods given only the canon of scientific law as currently recognised and accepted; any question of there being a need for exotic extensions expressly in the service of life and mind have -for reasons beyond my ken- been put beyond the pale. As this volume demonstrates, one may conceive of exotic extensions to the lex naturalis which might provide a substratum complex enough and appropriately organised in a way able to account for the *conservative* aspects of what the mind/brain ensemble is able to deliver, and may well be 'unknowably sound'. However, one would also have to credit the germplasm with an equivalent competence which must somehow be supported by a substratum of a mere 6.7×10^{-12} grams of anhydrous DNA. However, even were this to be granted, we would then be driven to ask how the DNA derived these extraordinary powers. In summary, I cannot see how quite liberal extensions to scientific law can circumvent the manifest phenomena of insight, progressive self-organization, true learning and leaps of creativity of every kind, in the absence a Gödelean appeal.

I now move on to the third proposed way of dodging the implications of Gödel's theorem: if we can extend the scope of a formal system by an open-ended addition of theorems which become axioms, why could we add a front end to a computer program, consisting of a loop which could do exactly the same thing?

> "......The assumption that we could suppose ourselves to have given the machine an adequate idea of mathematical truth when we give it the axioms and rules of inference is not true. This would be to suppose the formalists were right, and they were shown by Godel to be wrong. The Godel theorem is no more an obstacle to the computer than to ourselves.we recognize the truth of the unprovable formula by comparing what it says by what we know to be the case, so a computer can do the same." Michael Scriven

For a start, we might recursively add the system to itself as an extra axiom. However, as Lucas demonstrates, this can’t be carried through.

> "......The procedure whereby the Gödelean formula is constructed is a standard procedure--only so could we be sure that a Gödelian formula can be constructed for every formal system. But if it is a standard procedure, then a machine should be able to be programmed to carry it out too .This would correspond to having a system with an additional rule of inference which allowed one to add, as a theorem, the Gödelian formula of the rest of the formal system, and then the Godelian formula of this new, strengthened formal system, and so on. It would be tantamount to adding to the original formal system an infinite sequence of axioms, each the Gödelian formula of the system hitherto obtained .We might expect a mind, faced with a machine that possessed a Gödelizing operator, to take this into account, and out-Gödel the new machine, Gödelizing operator and all. This has in fact, proved to be the case. Even if we adjoin to a formal system the infinite set of axioms consisting of the successive Gödelian formulae, the resulting system is still incomplete and contains a formula which cannot be proved-in-the-system, although a rational being can, standing outside the system, see that it is true. We had expected this, for even if an infinite set of axioms were added, they would have to be specified by some finite rule or specification, and this further rule or specification could then be taken into account by a mind considering the enlarged formal system. In a sense, just because the mind has the last word it can always pick a hole in any formal system presented to it as a model of its own workings. The mechanical model must be, in some sense, finite and definite: and then the mind can always go one better."
>
> J.R.Lucas (1961)

Lucas sought to link this odd Gödelizing affinity of the mind to a psyche seen as an entity which is metaphysically simple and not compound, having a generality which stands above and before *all* differential definition or particularization.

> ".......The paradoxes of consciousness arise because a conscious being can be aware of itself, as well as of other things, and yet cannot really be construed as being divisible into parts. It means that a conscious being can deal with Gödelian questions in a way in which a machine cannot, because a conscious being can both consider itself and its performance and yet not be other than that which did the performance. A machine can be made in a manner of speaking to 'consider' its performance, but it cannot take this 'into account' without thereby becoming a different machine, namely the old machine with a new part added. But it is inherent in our idea of a conscious mind that it can reflect upon itself and criticise its own performances and no extra part is required to do this: it is already complete, and has no Achilles' heel." J.R.Lucas (/)

It is at this juncture, I suspect, that Penrose and Lucas parted company. As a Christian, Lucas saw the psyche as a soul, standing, presumably in cartesian relationship to the brain, whereas Penrose, by all accounts, is a determinist.

In other words, paradoxical though it sounds, even an infinite number of cycles through the loop would be insufficient. Penrose adds what is perhaps a more telling rebuttal when he notes that theorems and axioms cannot just be conceived and added blindly; an inescapable element of judgment is involved in managing the expansion if the process is to be water-tight. He cites, particularly the problem of addressing infinite ordinals in an unambiguous way which could lead to a water-tight Gödelizing loop. Turing himself had conceded earlier that *if* such a program could be created, it would indeed solve the 'halting' problem; such an outcome would be inconsistent with Gödel's theorem, hence, at the least, is very highly suspect.

In somewhat the same vein, the votaries of AI have claimed that 'learning' programs of a general-purpose sort, not jerry-rigged to address a particular class of problems, are essentially delivering a Gödelean performance in the absence of any appeal to any Platonic pipeline. Here is how Penrose (1990) counters this sally:

> "......Neural networks are supposed to 'learn' and modify their structures in ways that are not 'preprogrammed' for the solution of specific problems, that gradually improve their performance. The rules governing this learning and improvement, however, are just themselves algorithmic (otherwise it would not be possible to simulate them on a general-purpose computer).....Similar remarks apply to 'heuristics'....as soon as they have been programmed on a general-purpose computer, their algorithmic nature has been secured"'

Turning now to the 'monkeys and typewriters' argument. Of course, it is trivially true that such a procedure could recreate all of the discoveries of mathematics to date, and even more than this, anticipate mathematical discoveries yet to be made. The whole trouble is, as Penrose has pointed out, that these truths would be totally inundated, first by a huge population of fallaceous near-misses, and beyond this, a near-infinite mass of white noise. Any such procedure would lack the means of guiding itself towards meaningful goals, and more than this, simply wouldn't know when it had arrived at something interesting, and put it on one side for preservation.

I believe it is most important not to lose sight of the fact that should a proof-perfect copy of Gödel's celebrated theorem turn up –as indeed it must- it would not *be* Gödel's theorem until recognised and accepted as such by a qualified reader. What the monkeys would have produced would be but a configuration of alphanumeric characters which are not even symbolic of the celebrated theorem until mapped into the symbolic language resident within receptive minds.

Exactly the same sort of rebuttal may be urged against those who claim that creativity involves nothing other than a novel rearrangement of the resources which are already at hand. Here's Russian-American philosopher Ayn Rand, on the subject ['Philosophy: Who Need it']:

> ....".'Creation' does not (and metaphysically cannot) mean the power to bring something into existence out of nothing. 'Creation' means the power to bring into existence an arrangement (or combination or integration) of natural elements that had not existed before. (This is true of any human product, scientific or aesthetic). man's imagination is nothing more than the ability to rearrange the things he has observed in reality."

This may serve well enough as a *phenomenal statement;* it does not thereby transmute into an *explanatory principle*. The question of course is over the nature of the process driving the rearrangement. It is trivially true that *all* that Sandro Botticelli did in creating his 'Birth of Venus' was to rearrange the molecules of pigment on his palette, translating them onto the canvas in the process. Were Botticelli available to satisfy our curiosity on the matter, I doubt if he could tell us how he achieved the 'rearrangement' other than in terms of arm-waving generalities.

I final assault against the Gödelean embarrassment has been mounted by those who point to the shakiness of the foundations of mathematics. For example, Cohen's theorem tells us that the continuum hypothesis is 'undecidable' while there is no general agreement about the firmness of the Zermelo-Fraenkel system in providing the needed ground floor for the mathematical edifice as a whole. Can we therefore trust our judgment that the very recondite Gödel insight is 'obviously' true? While conceding the open-endedness about our ideas on the foundations of mathematics, very much the same arguments can be voiced about the operation of the proof-process in general.

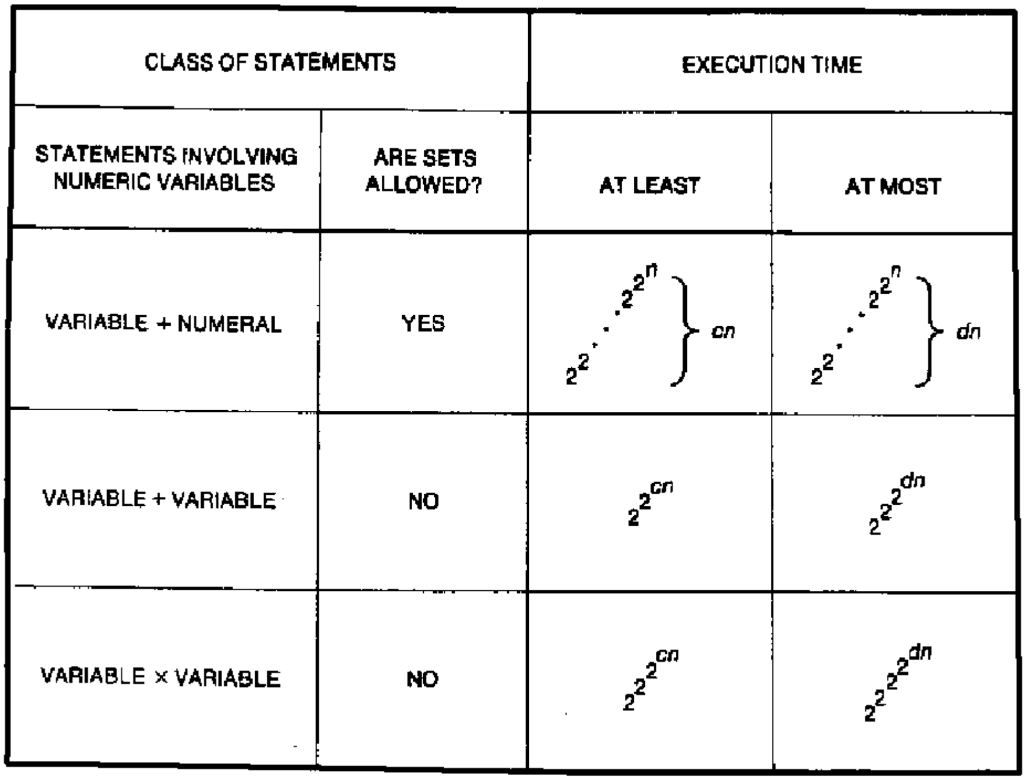
John Lucas was the first to explore the mental implications of Gödel's theorem (Lucas 1961, 1970) Predicably, his claims were subject to a great deal of criticism by the Intellectual Establishment, to the point where -as Penrose remarks [ibid. p 49]- the general impression was given that the whole matter had been satisfactorily disposed of. Penrose's treatment of possible loopholes is extraordinarily thorough (he deals with a round score altogether) and anybody to whom the idea of an intuition coming suddenly into possession of truly novel insights -out of thin air, as it were- is unacceptably repugnant, will have to take on Penrose. And judging by all accounts he or she would be well advised to get a good night's rest in advance of the encounter.

Proof-processes in practice

For first-order systems, automated verification of conjectured theorems can always be devised and executed -at least in principle. Yet the exercise of human gumption can often find a shorter road to home.

Thus S.C.Chou (1988) has successfully automated a general proof process for euclidean geometry, able to demonstrate the truth or falsity of proffered theorems [-and also, perhaps,to make discoveries. A not too complicated geometric figure which is the subject of Thèbault's conjectured theorem, was proved to be genuine, though the process took 44 hours of computer time [(penrose, ibid. p.200)]. But it was discovered that the proof process can be greatly shortened if one first solves the more general problem of which Thèbault's conjecture is a special case - avenue of approach which would not have 'occurred' to Chou's software-hardware ensemble.

As discussed further above, arithmetic can be reduced to a first-order system if appropriate constraints, of one kind or another, are imposed upon the rules of the game. Three such 'cut-price' arithmetics are shown below (Stockmeier et all S/A 1979 p.126:



CLASS OF STATEMENTS		EXECUTION TIME	
STATEMENTS INVOLVING NUMERIC VARIABLES	ARE SETS ALLOWED?	AT LEAST	AT MOST
VARIABLE + NUMERAL	YES	$\left.2^{2^{\cdot^{\cdot^{\cdot^{2^{n}}}}}}\right\}cn$	$\left.2^{2^{\cdot^{\cdot^{\cdot^{2^{n}}}}}}\right\}dn$
VARIABLE + VARIABLE	NO	$2^{2^{cn}}$	$2^{2^{2^{dn}}}$
VARIABLE × VARIABLE	NO	$2^{2^{2^{cn}}}$	$2^{2^{2^{2^{dn}}}}$

Figure 1

c, and d are constants greater than zero

A comparative examination of these discloses two interesting facts. First, that broadening the envelope of operations allowed within the system imposes a dramatic scaling up of the dependence of computing load upon problem size. Yet such leverage is not to be compared with that consequent upon the inclusion of sets - notwithstanding the severe limitations upon allowable operations. As is shown in figure 1 computational demands are *index-level* proportional to problem size.

Penrose gives some delightful examples of how looking at a problem askance, as it were, can cause the solution to jump into focus; i.e. something which might have initially seen as an odd conjecture suddenly becomes *conspicuously* true -that is, its veracity is seen *whole*, in a glance, rather than piecewise through a catenation of deductive steps, no matter how firm such links in the chain may be.

The following is a delightfully simple example, the first of two given by Penrose [ibid., p.55; 69-71]. Does (a x b) = (b x a)? for example let a = 3 and b = 5. Then we may represent (a x b) and (b x a) respectively, as:

(# # # # #)(# # # # #)(# # # # #)

and

(# # #)(# # #)(# # #)(# # #)(# # #)

Now, orthogonalise the two ways of representing the multiplication, and the truth of the identity leaps to the eye:

# # # # #
# # # # #
# # # # #

We can see that this way of presenting it makes it clear that what we actually have is a single object which can be regarded in two different ways.

Of course, we already knew that we get 15 in both case, the point is that this 'proof' process makes it absolutely clear is its generality, and that in consequence we can confidently assert that:

79797000222 x 50000123555 = 50000123555 x 79797000222

-without any need of proving our point through a brute-force demonstration at the computer.

A second, trickier but thoroughly delightful example is equally revealing. It concerns 'hexagonal numbers' whose definition is made obvious by figure 2 below which presents. As shown, we pass from one member of the series to the next by adding a perimeter of dots. In so doing, successive multiples of 6 are added each time: 6, 12, 18 ,24, 30 and so on.

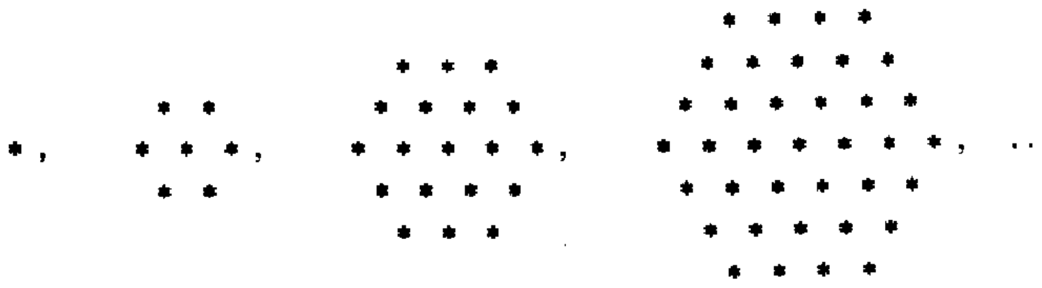


Figure 2

Now here's something which strikes the mathematically untutored eye (such as my own) as odd and unexpected. When hexagonal numbers starting from the first are summed, the resulting totals are 1, 8, 27, 64, 125. what immediately strikes one as odd is that these numbers are the first five perfect cubes: 1^3, 2^3, 3^3, 4^3, 5^3 Is this a coincidence?. let's push on to 6. this gives us 216, which, sure enough is 6^3. We begin to suspect that this may be a general truth, and furthermore, its ordinal character suggests that the matter might be subject proof through mathematical induction, and this turns out to be the case; the conjectured theorem is, indeed true. But there is a much shorter road to home. Figure 3 below organizes the summing process in a way in which the truth of the theorem becomes instantly *perspicuous*:

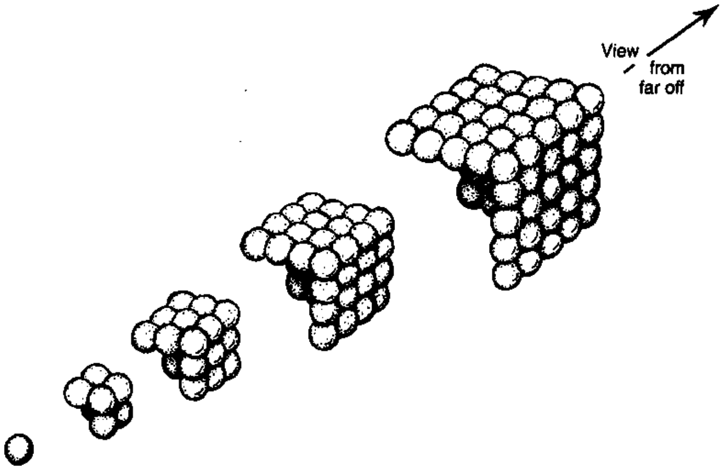


Figure 3

The succession of hexagonal numbers is organized in such a way as is *seen* to produce a succession of cubes. Figure 3 shows how what needs to be added to produce the next cube in order is, at one and the same, time a hexagonal number, and also the surface of three sides of the new cube which is produced. Suddenly, those hexagonal figures 'look' quite different; a latent 3D presence springs into life!!

Proof Processes outside of mathematical induction need not be complex. e.g. the reductio ad absurdum proof that there's no highest prime:

1.Assume that *x* is the greatest prime.

2.Form a product of all primes up to (and including) *x* and add one to the product. This gives us the number y:

y = (2 X 3 X 5 X 7 X 11 X.........X *x*) + 1

3.If y is prime then *x* cannot be the greatest prime since clearly y > *x*.

4.If y is composite, then once again *x* cannot be the highest prime. For if it were, it must contain a prime divisor z differing from any of the primes up to x, hence z must be greater than x.

5.Hence, no matter whether y turns out to be prime or composite, the presence of a prime higher than *x* is established.

An equivalent reductio ad adsurdum was utilized in proving that the diagonal of a square is incommensurate with its sides, and like the one above is compact enough to be written on the back of an envelope. G.H.Hardy (1967) remarks that the "....reductio ad adsurdum, which Euclid loved so much, is one of the mathematician's finest weapons". (p.94)

Following is a truly fascinating example of a problem in which the immediate appearance deceives. One might think at first glance that the magnitude of the number so defined will continue to grow spectacularly until the bitter end, and the ordinal nature of the presentation suggests that perhaps the truth of the conjecture might be demonstrable through mathematical induction:

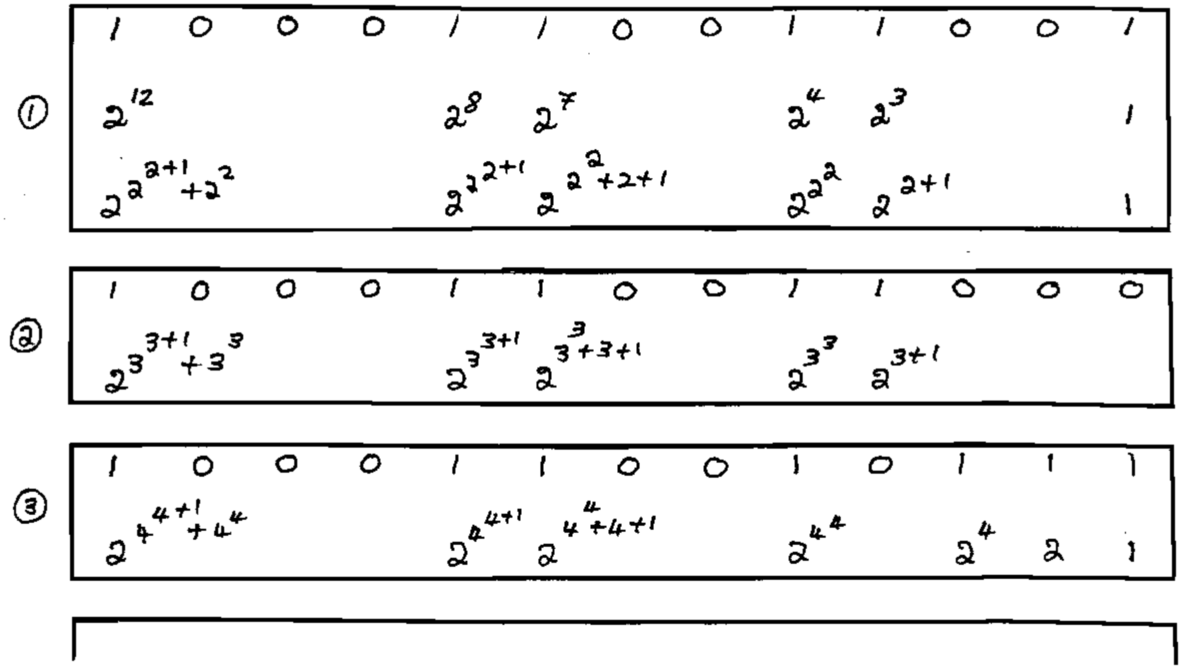
1.Pick any decimal number at random and expand it into its binary form.

2.Convert the indices into powers of 2, i.e. reducing all higher indices to powers of two, where possible.

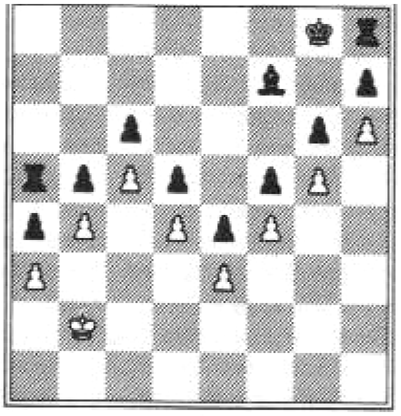
3.Subtract 1, recompute the indices, then replace all of the twos with threes.

4.Reiterate step 3, raising all of the threes to fours. and so on.

Yet a closer inspection makes it suddenly obvious that to the contrary, it implodes into zero at the final step.



Obviously, the number climbs very rapidly and might well pass beyond a brute-force computational evaluation, but it only takes a moment's reflection for it to become obvious that after n iterations, where n is the size of the initial number, the final term must be zero because radix upon which the indices are piled passes out of existence! The grin on the Cheshire cat may be impressive, but Lewis Carroll notwithstanding, cannot persist after the cat has vanished!



Finally, Penrose offers this example. White is to play and draw – easy for humans but the Computer chess program’Deep Thought’ took the rook!...

From all of these examples, it is evident that what is constantly called upon, in any novel situation, is a 'bright idea' in which the true nature of the problem and how its proof must be approached suddenly becomes perspicuous -either in the grand sense of the whole proof process is conceived integrally, that is to say, in a single glance, or failing this, that the matter becomes catenate-perspicuous, in which a perceptively-sound chain arguments suddenly comes into view. Ramanujan was said, by those who knew him, to be intuitive rather than vigorous or 'sound' in his approach:

> ".....Ramanujan was self-taught; he knew nothing of the modern rigour; in a sense, he didn't know what a proof was....In fact, Hardy was obliged to teach him mathematics as though Ramanujan had been a scholarship candidate at Winchester." Hardy (1967) p.36

As one example, he conjectured that $e^{\pi\sqrt{163}}$ was an integer. It was later discovered to be an 18-digit number followed by decimal .9999999... up to 2 million digits. I imagine that most creative mathematicians have access to such perspicuous viewing -of which this is an extreme instance. I believe it was Eric Temple Bell who commented that a mathematician was one to whom the truth of the equation $e^{i\pi} = -1$ was as obvious as Buddha's navel!

However, less spectacularly, one's feeling towards a solution may be guided by semi-formal hunches and 'what-if' arguments falling under the general heading of *heuristics*. Automated computation may also play a part, especially where searches are called upon as one step in the proof process. Perhaps the best known example here was the solution of the 4-colour map conjecture. Unlike the situation with first-order systems, the computer has no way of taking on the whole job, but can only be brought into service after a mathematician has used his wits to *structure* the investigation, isolating those steps in the process which can be computer-offloaded.

There is one class of mathematical performances which is hard to classify, with regard to whether they are or are not strictly and entirely algorithmic, or whether something else is also involved. I am referring to those remarkable individuals customarily alluded to as 'idiots savants', more decently alluded to by Penrose as 'autistic savants':

> "......Even more peculiar perhaps are the cases of 'autistic savants', people who are mentally handicapped, and may have difficulty performing even the most elementary formal arithmetic manipulations, but who nevertheless possess the uncanny ability to produce correct answers to mathematical problems that appear to ordinary people to be impossibly hard. Two American brothers, for instance, can consistently outdo a computer in finding prime numbers even though they are both mentally retarded". Penrose (1995) ibid p.154

The question is- are they pulling these numbers out of thin air, as though through a Platonic pipeline, or are they in possession of an exotic algorithm into which we have not yet come into possession. My own instinct is to back the second option, this being very much coupled to by belief that the substance of mind is exotic in nature, and that it is able to support the kind of structures needed for 'alternative' ways of executing familiar mathematical tasks. In embarking on such speculations, This position is, perhaps easier to maintain in another case, reproted by Penrose, of an illiterate farmer's son who was able to multiply two 8-digit numbers together in less than a minute I'm not for a moment suggesting that any escape from turing machine equivalence and reducibility is involved; the substance and the architecture which it supports may be exotic but not the logic under which it operates.

The great mystery which remains, however, is just why it should be that such outlandish skills should be closely -and almost exclusively- associated with amentia or mental retardation. Is it that the architecture of the mind/brain ensemble which forms its substratum is simply incompatible with more commonplace forms of mental competence? What, one wonders, did the germ plasm 'have in mind'. Does it bear some analogy to sickle cell anaemia which suits the individual concerned to a particular environmental niche -i.e. that where malaria is prevalent- while making him less able to survive than his fellows in more conventional climbs?

GÖDEL'S PROOF

Preliminary

Hilbert's general program was that of patterning the whole of mathematics after the example of his successful reformulation of Euclidean geometry. Implied here was a belief that it could all be shown to be first-order, that is to say, totally tautological (in Wittgensteinian sense) in which all truths would automatically be theorems within the system. All systems could, he hoped, be placed upon an axiomatic grounding which was both consistent and complete so that all conjectures would be either true or false -there would be no question of undecidability. For example, he hoped that Peano's formulation of arithmetic could be extended in such a manner as to make it both complete and consistent. Not clear is how his program -operating under the 'restriction of finitude' could hope to cope the kind of antinomies uncovered by Russell (at the turn of the century), within infinite sets (for example, the infamous 'Barber' paradox and its many alter egos).

Hilbert sought to drain mathematics of all 'meaning' -that is to say purge it of all attachments to those domains 'out there' from which much of mathematics took origin. Mathematics might be said to be 'applied' before it was 'pure', and Hilbert sought to bring such purification to completion. Such a movement was to markedly alter the significance of the meaning of 'truth' in mathematics, making of it an entirely internal matter of verifiability from the axioms of the system. 'Transcendent' or External truth 'outside' of the system no longer had meaning -for the proffered reason that one had only to modify the set of axioms to cause different 'truths' to emerge. This made of mathematics a rarefied intellectual game which pointed to nothing beyond itself.

One can think of Gödel's work as stemming ultimately from the epimenides paradox: "All cretans always lie." -uttered by the cretan epimenides. Likewise with "This sentence is false", and in its many other forms, reaching its apotheosis in Russell's "barber" (and related) paradoxes. All allow both the thesis and its antithesis to be proved true within the system, hence proving the latter to be inconsistent.

One form is of special relevance to the matter at hand. It is the Richardian paradox which was to surface in 1905. Jules Richard came up with the notion of describing the properties of integers in terms of brief sentences in natural language -which, for the convenience of the present discussion we will take to be English. The set of sentences so obtained were then ordered on length (in terms of the number of characters each contained; any ties in length were to be broken in terms of alphabetic precedence). The sentences were then assigned natural numbers in order of length. Richard then created an epimerides-type self-referencing loop by comparing what the sentence was asserting about numbers with the properties of the number which had been assigned to the sentence. For example, if a sentence defining prime numbers as "Those numbers divisible only by themselves and one" had been assigned the tag 17, then this number clearly reflected the property in question; had the number been 16, then clearly it would not. In the first case, the number is said to be 'Non-Richardian' and in the second 'Richardian'. We have now but to ask "Is n Ricardian" for the trap to be sprung. The epimerides paradox re-emerges, for we immediately see that n is Richardian if, and only if, it isn't.

Clearly, this is highly suggestive of the Gödelian self-reference which was to emerge a quarter of a century later, and it is a matter to which Gödel specifically refers in his classic paper -something, as it were, projecting an uncanny shadow into the future. The difference is that the numbers which Richard assigned to his sentences are *mere* tags which do not mirror in any way the actual structure of the sentence. 'n' was a numeral which stood directly for the corresponding number; 'n' was the numeral for n.

Gödel sentences also have natural numbers which can be calculated -with the all-important difference that the numeric which defines the number is structured in such a way that it maps into the content of the sentence in a perfectly seamless way; it can absorb its own image into itself because the two are perfectly matched, being in the same language. Clearly, the loop establishing the self-referencing closure is tighter than that of Richard -with magical consequences; there is all the difference in the world between "This sentence is untrue" and "This sentence is unprovable". They may have the appearance of brothers but actually are no more than cousins. The Gödel number is, indeed, precisely a number, but one which can be unwrapped to describe what the sentence is asserting. Of cause, many of the Richard numbers will in fact be Gödelian in constitution, but this distinction is lost on Richard to whom all numerical tags are equivalent. Suddenly, we are no longer obliged to infer inconsistency thanks to an alternative choice which appears seemingly by magic; consistency may be affirmed at the acceptable price of incompleteness.

One wonders how Hilbert managed to maintain heart in his enterprise after Gödel's bombshell of 1931; here is the cannonade which opens his thesis:

> ".....The development of mathematics in the direction of greater precision has led to large areas of it being formalized, so that proofs can be carried out according to a few mechanical rules. The most comprehensive formal systems to date are, on the one hand, the Principia Mathematica of Russell and Whitehead and, on the other hand, the Zermelo-Fraenkel system of axiomatic set theory. Both systems are so extensive that all methods of proof used in mathematics today can be formalized within them -i.e. can be reduced to a few axioms and rules of inference. It would seem reasonable, therefore, to surmise that these axioms and rules of inference are sufficient to decide *all* mathematical questions which can be formulated in the system concerned. In what follows, it will be shown, that this is not the case, but rather that, in both cited systems, there exist relatively simple problems of the theory of ordinary whole numbers which cannot be decided on the basis of the axioms." Smullyan (1992) p.1

It is particularly ironic that in putting paid to Hilbert, Gödel drew upon Hilbert's Metamathematics, and the technique of mapping which hilbert had done much to foster. While Hilbert, in his meta-mathematics, cast mathematical arguments into a formal language, Gödel closed the circle by mapping metamathematical statements back into the system of arithmetic itself.

Mapping metamathematics into arithmetic

> ".....Gödel first showed that it was possible to assign a *unique number* to each elementary sign, each formula (or sequence of signs), and each proof (or finite sequence of formulas). This number, which serves as a distinctive tag or label, is called the 'Gödel number of the sign, formula or proof."
>
> Nagel & Newman 1958 p.69

First of all, we need a systematic way of unambiguously assigning Gödel numbers to constants, numerical variables x,y,z,...; predicate variables P,Q,R,....; and such elements of logic as ∀and ∃ Following is such an appropriate set of assignations:

~	1			
V	2	V = Logical-OR;		
⊂	3	(Logical -AND is not primitive; = p.q = ~(~pV~q))		
∃	4	(x) is used instead of ∀		
=	5			
O	6	Other Numerals not primitive; thus 4 = SSSSO		
S	7			
(	8	x,y,z	11, 13, 17..	
)	9	p,q,r	11, 13, 17...	Raised to Power 2.
,	10	P,Q,R	11, 13, 17..	Raised to Power 3

e.g., "For every x there is an immediate successor y":

(	∃	x	)	(	x	=	s	y	)
↓	↓	↓	↓	↓	↓	↓	↓	↓	↓
8	4	11	9	8	11	5	7	13	9

This having been done, we now need a rule for consolidating these into a single number M. This may be achieved by taking the primes in their natural order, and raising them to the power corresponding to the Gödel numbers of the sequence of elements, obtained above, then forming their product.

(∃x)(x=sy) --> 2^8 X 3^4 X 5^{11} X 7^9 X 11^8 X 13^{11} X 17^5 X 19^7......X 23^{13} X 29^9

Suppose this has the Gödel number m.

If we want to make the sentence self-referencing, that is, if we wish to obtain:

(∃x)(x=sm),

-then we will need to replace the Gödel number of the symbol for y with a string of length m of the Gödel number of 'successor of', i.e. the number 7. The Gödel number of this new self-referencing sentence will be given by:

2^8 X 3^4 X 5^{11} X 7^9 X 11^8 X 13^{11} X 17^5 X 19^7 X 23^7 X 29^7 X 31^7 X .... X $p^9{}_{m+10}$.

-where p is the next available prime after the m primes taken up by the number of primes taken up by the m 'successors of'.

Stacks of formulas may now be consolidating by multiplying together the succession of primes each of which is raised to the Gödel number of the formula in question. e.g., if:

(∃x)(x=sy) has Gödel number M, and

(∃x)(x=O) the corresponding number N,

-then their combined Gödel number will be given by:

2^M X 3^N

Because the sentence addresses itself, we need a way of absorbing the corresponding number back into the sentence.

Finally, the logical relationship of 'Demonstrable' or 'provable within the system' is rendered, in the metamathematical sentence simply by as a product of the two corresponding Gödel numbers of the Proof Process and the Conjecture which it affirms. Thus, "The stack of formulas with Gödel number x are a valid demonstration of the conjecture with Gödel number y", or Dem(x,y) may be explicitly renedered as "'...'If you factor x into primes, then the largest prime has exponent y. Furthermore, the exponent of every prime is either the Gödel number of an axiom, or it can be derived from one or more of the previous exponents by means of the transformational rules'. All this, of course, has to be written out in formal language, including the Gödel numbers of all the axioms (written as ssssss...ss0), expanding out the definition of prime factors and exponents, etc. What you get is a very long formal *statement* which includes the variables x and y." (Gerver 1997).

It follows directly from all of the above rules for assigning Gödel numbers to Metamathematical sentences that every such sentence yields a unique Gödel number. And it follows from the 'Fundamental Theorem of Arithmetic" (that every composite number may be *uniquely* factored into its primes) that the process is reversible, i.e. given any Gödel number, the numerical expression from which it was derived can be unambiguously reconstructed. However, not every number is a Gödel number. For example, 100 cannot be a Gödel number because it factors into 2^2 X 5^2. Here, the prime number 3 has been skipped, thus breaking one of the rules of formation of Gödel sentences. Also, because 100 is greater than 10 it cannot be an elementary logical sign, nor is it a prime greater than ten nor is it the square or cube of such a prime.

An outline of Gödel's proof will now be presented; figure 5 summarises the way in which the steps are sequenced.

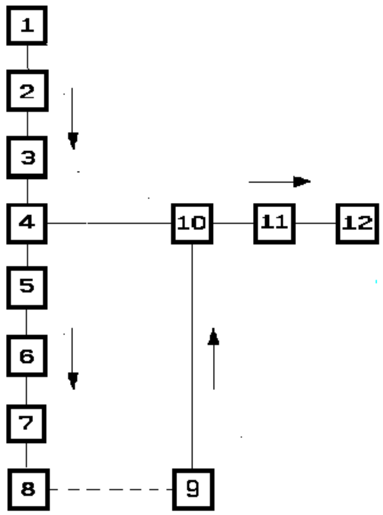


Figure 4

We may start with a Gödel sentence asserting the following:

1."The (particular) stack of formulas with Gödel number x does not constitute a demonstration of the (particular) conjecture with Gödel number z" -or "~Dem(x,z)".

Next, generalize this sentence to apply to *any* set of proofs which might be devised; the first term signifies 'for every x':

2."For every x, the stack of formulas with Gödel number x is not a demonstration of the (particular) conjecture with Gödel number z" -or, "(x)~dem(x,z)".

I'm not sure of the significance of the next step:

3."For every x, the stack of formulas with Gödel number x is not a demonstration of the conjecture with (any) Gödel number y -where y is the Gödel numeral of this sentence" -or, "(x)~(x,sub(y, 13, y))". designate the number of this sentence by n

Now, finally, we substitute the Gödel number for the symbol y with the numeral n of the sentence in (3) above, ending up with the sentence designated, traditionally by G:

4."For every x, the stack of formulas with Gödel number x is not a demonstration of the Gödel number of the sentence G (i.e., this sentence)" -or, "(x)~Dem(x,sub(n, 13, n))"

Hence, sentence (4) proclaims of itself that it is not demonstrable. But if a thesis is non-demonstrable, then so also is the contradictory formula. Hence we end up with:

5."Sentence G is demonstrable if, and only if it is non-demonstrable"

Which confronts us with two alternative inferences:

6.Either Sentence G is undecidable, or the system of arithmetic is inconsistent. If the conjecture be true, it is undecidable; if it be false, then it is provable within the system, that is to say, the system is capable of proving falsehoods, hence is inconsistent.

That the possibility of inconsistency should raise its head should occasion no surprise, since given such a system, anything and everything can be proved to be true.

So:

7.If arithmetic is consistent, Sentence G is undecidable

Gödel was then able to prove that:

8.The consistency of arithmetic cannot be proved within the system.

However, this does not foreclose upon demonstrations of consistency being made by arguments pursued outside of the system. Gentzen was to do just this. His procedure may be summarized as follows. Consider all arithmetic expressions to be arranged in 'order of simplicity'; this sequence is of 'trans-finite ordinal' type. The proof of consistency then follows by the process of 'transfinite induction'. Such a procedure, needless to say, cannot be cast into that of Peanoese mathematical induction. [Nagel 1948 p.97] Although it doesn't fully meet Hilbert's criterion of consistency, virtually no-one doubts the claim; if it has been said that if any inconsistency were ever to turn up in arithmetic, all hell would break loose. Hence, finally:

9.Arithmetic is Consistent

So:

10.Gödel's Sentence is undecidable, hence arithmetic is incomplete.

But this is not quite the end. For we can adjoin Gödel's sentence to itself, and then, by cycling through the proof process, establish that:

11.Arithmetic so augmented is still incomplete.

And since this process can be repeated endlessly, we finally end up with:

12 Arithmetic is incorrigibly incomplete.

<u>The Turing Approach -a Shorter Road to Home</u>?

Gödel's landmark paper of 1931 was to be reformulated a few years later by Turing (1937) in terms of the behaviour of generalized computing machines. The account which follows draws heavily upon Penrose's discussion (1995, p 73-76).

We have to think of two such machines running in tandem; the first machine has the job of determining whether the second will ever achieve its computational aim. This may be most easily visualized by thinking of the second machine as being engaged in an enumerative proof-process, while the first is addressing the same problem in a more analytical way. That is to say, we may think of the second machine as systematically testing some conjecture for each numeral n which is involved. Let us take Goldbach's Conjecture as an example. The machine may be thought of as doggedly making its way through the sequence of even natural numbers, in an attempt to discover one which was not the sum of two primes. Naturally, as the numbers become large, more and more time will be spent for each instance, even though we endow the machine with all of the dodges which have been discovered of shortening the process. However, we are not concerned with this increasing burden, no matter how large it may become. The main point is that the machine can always complete its check within a finite time, for each n, no matter how great this time may be. And we have arranged matters such that our machine will automatically halt, when it first discovers an even number greater than 2 which is not the sum of two primes. Hence, if the machine halts, it means that the conjecture has been refuted.

Now let's look at the first machine, which is assumed to be a general purpose theorem solver embodying all possible proof processes which in principle are available to mathematicians -which is to say a great deal more than that it is a compendium of mathematical knowledge acquired to date. Such a program is assumed to be able to tailor itself to the problem at hand, that is to say, its behaviour will be a function of the particular task being enumeratively addressed by the second Turing machine. Our first machine, like the second, has been constructed so as to stop if it achieves its aim, that is manages to derive a general proof of the conjecture. So, a halting of the first machine implies that the second will never halt -provided only, of course that its construction be sound, that is, does not harbour formal inconsistencies.

In order to proceed with Turing's 'undecidable' theorem, we now need to take a closer look at the second machine of the tandem pair. We are offered an infinity of such machines to choose from, the total population including every possible computational regimen which can be conceived, and not just those which might seem appropriate or promising for the task at hand; one restriction, of course, is that each such machine shall be able to act upon all natural numbers, so that any given process C will be able to execute C(0), C(1), C(2), C(3), and so on, or more generally, C(n). We can order this population of machines in terms of the executions they perform: thus, C1(n), C2(n), C3(n), C4(n),..... or, more generally, Cq(n). I find it convenient to think of the machines as being ordered in some appropriate way upon complexity; this way of thinking of the matter engenders an appropriate mind-set which anticipates that one can hardly expect to find any machines of real utility until q becomes quite large.

The first machine of the tandem pair will embody the universal theorem-prover which we may designate as A, which when presented with a pair of numbers will seek to prove that the second machine, executing $C_q(n)$, will never stop.

The next step in the argument is that of setting q = n, so that the procedure now depends upon a single parameter n. So, we have: if A(n,n) stops then $C_n(n)$ does not.

Now, since the $C_q(n)$ stack includes every possible computational regimen, A(n,n) must be identical one of them: let us designate this by q = k. and since q = k = n, we arrive at the identity A(k,k) = $C_k(k)$. This presents us with the self-referencing proposition: If $C_k(k)$ stops, then $C_k(k)$ does not. From this it immediately follows that $C_k(k)$ never stops. For if it did, then according to its identity A(k,k) it didn't!! But if $C_k(k)$ doesn't stop, then this must also be true of its alter ego A(k,k) The final conclusion is that a proof has been established by a means which is beyond A(k,k)'s ken. Thus, suppose that this pair of Turing machines were addressing the Goldbach conjecture. Evidently, should the theorem turned out -as everyone has long suspected- to be true, this fact would be unknown to A(k,k) so we would remain none the wiser! This is because although $C_K(k)$ and A(k,k) are identities, yet we have to think of A(k,k) as an analytical theorem-solver and $C_k(k)$ as engaged in an enumerative search; it as if the same person were wearing two different hats. Surely such algorithms must be very rare, but since we have an infinite envelope of algorithms open to us -the whole diagonal sequence in the infinite square matrix ordered on n, on one side, and q = n on the other, there must be plenty available; maybe like prime numbers, they are denumerably infinite while being of zero density. And also, of course, to qualify for any kind of consideration, the algorithms mustn't be duds in any of a numbers of ways (Penrose 1995 p. 80). Such duds might indeed come to a standstill, but always in ways which could be shown to stem one form or another of malconstruction.

APPENDIX

These remarks have been added after receipt of Joe Gerver's responses to this document.

(1) Roots of quintics are not transcendental, but are still considered to be algebraic, although they are odder than ordinary algebraic irrationals which involve surds. Such roots cannot be expressed by any finite combination of + - * / and $^n\surd$. Yet this oddness does not break out of denumerable infinity. Gerver notes: "...Algebraic numbers are indeed denumerable. To see this, it suffices to order all polynomials with integer coefficients. What you have to do is to compute, for each polynomial, the sum of its degree and the absolute value of all the coefficients. This sum is a positive integer for every polynomial, and there are only a finite number of polynomials for each value of the sum. Hence all the polynomials can be ordered acccoding to the value of that sum, with someother rule (it doesn't matter what) to break ties. The transcendentals have the same cardinality as the reals. This is greater than aleph null. Whether it is aleph one depends on whether you believe the continuum hypothesis."

(2) Intrinsic curvature of one's space, and whether or not this implies embedment in a space of higher dimensionality.

O First, some general points. In two-dimensional space, the min and max intrinsic curvature are always mutually orthogonal. These two numbers completely characterise the curvature but cannot be unambigously measured from within. What you can measure from within is their product -the Gaussian curvature. But this is ambiguous. Thus a discovered Gaussian curvature of 1 might equally signify a sphere (1 x 1) or a football (1/2 x 2). But on could discover whether one was living in a space which was flat, or positively or negatively curved by inspecting the sum of angles within a sufficiently large triangle. When you get to three-space, you now need three principle curvatures.

O Apparently, no such embedment is demanded.

O In addition, it by no means follows that it is possible.

1. If a two dimensional space of zero gaussian curvature everywhere is such that youo get back towhere you started in all directions, then if you want to speak of embedment you will need a flat torus (in effect a square with wrap-arounds in both x and y directions) -which demands not 3 but 4 dimensons.

2. Similarly, if they found their space had uniform Gaussian curvature of -1 then again no embedment is possible within 3-space. Beltrami's pseudosphere (trumpet) and saddles are only good over a finite region; if you keep going, you encounter a boundary or singularity.

O If, nevertheless, one does wish to insist that ones space is embedded in one of higher dimensionality, then:

1. One had better be sure one has invoked a sufficiency of dimensions.

2. Why, Gerver asks, should one postulate a *Euclidean* space? He claims this is Euclidean chauvinism --one could equally postulate a curved space for embedment.

Chou, S.C. Mechanical Geometry Theorem Proving Ridel 1988

Cohen, P.C. Set Theory and the Continuum Hypothesis. Benjamin Menlo Park, C.A. 1966

De Long, Howard [?] [Tital?] Scientific American March 1971 p. 50-60 [?]

Dennett, Danieal C. Darwin's dangerous idea Touchstone New York NY 1995

Gerver, Joseph. Private Communication 1997

Gödel, K. Über Formal Unentscheidbare Sätze der Principia Mathematica und Verwandter Systeme. Monatshefte für Mathematik und Physik, 38, p.173-98

Hardy, G.H. A Mathematician's Apology Cambridge Cambridge University Press 1967

Lewis, Harry R., and Papadimitriou, Christos H. The Efficiency of Algorithms Scientific American Jan 1978 p. 96-108

Lucas, John Minds, Machines and Gödel Philosophy, 36, p.120-4 1961

Lucas, John The Freedom of the Will Oxford University Press 1970

Nagel, E. and Newman, J.R. Gödel's Proof. New York University Press 1958

Penrose: Emperor's New Mind. Behavioral and Brain Sciences 1990

Penrose (1996). Penrose presented this example during a talk given at the National Institute of Standards and Technology at Gaithersberg, Maryland in the spring of 1996.)

Penrose, R. Shadows of Mind Vintage. Random House London 1995

Perlis, Don The Emerors Old Hat, in Behavioral and Brain Sciences (1990)

Smullyan, Raymond Gödel's Incompleteness Theorems. New York Oxford Oxford university Press 1992

Stockmeier, Larry J. and Chandra, Ashok Intrinsically Difficult Problems Scientific American May 1979 p.140-159

Turing, A. M.(1937) On Computable Numbers, with an Application to the Entscheidungsproblem. Proc Lond. Math. Soc. (ser.2) 42, p.230-65.