Melampus and the Golem

An Apology

By MELAMPUS

The autobiographical account of one man's struggle to refute the claim of strong AI that a computer can be brought to life.

How a solution to the Halting Problem limits the significance of the Theory of Evolution by Natural Selection and alters the balance of belief in the history of ideas in favour of religion.

Account of the proposed solution to the Halting Problem by mathematical induction.

Transcriptions of selected letters and correspondence sent by Melampus relating to all the above.

www.melampus.name

Black's Academy Limited
Kington, England

© Black's Academy Limited, May 2019

This publication is intended for free electronic distribution. The typesetting, arrangement and formatting of the document may not be altered without prior permission of the publisher.

Melampus and the Golem

An Apology

CONTENTS

Preface	2
The announcement	4
The problem of Silence	6
The Turing Test and the Turing Prediction	8
The Press Release	10
Does silence imply consent?	11
About mathematical induction and why the proof is “almost certainly valid”	12
The immediate background to the announcement	15
A red-hot issue - the dialectic and Darwin’s theory of evolution	19
First and second order – why mechanisation of the mind does not survive if the solution to the Halting Problem is valid	21
The history of an illusion	23
What if my solution to the Halting Problem is shown to fail?	25
Thirty-eight years ago – and you reap as you sow	26
Conclusion and appeal	29
Postscript – about privacy	29
Appendix – recipients of the letter to “the academics”	30

Preface

What you witness here is a live experiment in the history of ideas. All that is quoted here are real letters that have been sent to real people.

To illustrate immediately what I mean by this, the following is a letter that was sent to almost every Anglican bishop of England.

> 30 March 2019
>
> Letter sent to the Bishops of the Anglican Church of England
>
> My Lords Spiritual,
>
> I am writing to inform you of an initiative I have undertaken, and to ask for your blessing and practical support.
>
> I am publishing a press-release announcing a solution to the Halting Problem. The significance of this work in mathematical science to religion is clearly explained in the press release, and in the letter that I have sent to selected academics in the field of mathematics and computer science. I attach copies of both.
>
> You have in your community a philosopher and theologian who has his whole life fought for the same cause that I am fighting for now – Mr. John Lucas. Although he is now advanced in years, he can explain the importance of this result should my own comments be insufficient. I believe that his position will be vindicated by my work on the Halting Problem as by other papers published on my website, where there is also a full technical vindication of his interpretation of Gödel's theorem.
>
> I mention in my letter to the academics that I do not belong to a Church. This is correct. However, based on the Scriptural testimony of Mark and Matthew I do accept that Jesus was resurrected from the cross. I believe that our Lord Jesus was crucified, died and resurrected to restore to mankind its belief in the transcendence of human nature and in the possibility of Eternal Life.
>
> But what have we now, but atheism and materialism by stealth? Triple maths and physics, or triple maths and computer science does not equip the person thereby to make comments on spiritual matters, and yet that is in effect what every day the mechanistic philosophy of computing does. This matter is too important to be left to the computer scientists alone. They have set up a closed community and produce work that the "lay" person is not equipped to evaluate. I believe that academia is besotted with the idea that human nature is mechanical, and it acts collectively without reflection. Hence, it was inevitable that a criticism of their approach would come from source entirely outside their provenance. I should say that there is in the academic community a good deal of discomfort with the narrowness of the philosophes, but there is also a fear to speak out.
>
> You can help with your prayers. But you can also help practically by asking those mathematicians and computer scientists you know to properly evaluate the ideas I have presented. Even in the unlikely event that my solution to the Halting Problem being refuted, I do not doubt that the raising of the issue into proper debate, one in which the philosophers will have to justify their ideas to the community as a whole, and in a plain language that we all can understand, will have the effect of being a bolster to religion. I have no doubt that my work, taken as one whole, does firmly refute the pretentions of computer science as to what it claims to be able to demonstrate about human nature. Doubt will be restored, and with it the possibility of faith. Therefore, I ask for your blessing.
>
> Yours truly,
>
> Peter "Melampus"

Since writing the above, further correspondence and letters have ensued. Since most people would regard the writing of such a letter as surprising and even unwarranted, my purpose here is to explain my reasons for doing so. I am confident that the reader will approve my writing of this letter, or at least understand my point of view, once he has grasped the whole background to it.

Before I commence, some words about the rules regarding correspondence that I shall follow here. Copyright law generally prevents the recipient of a letter from publishing it in full. Some restricted quotation may be allowed. However, in general I respect the privacy of my correspondent. In some cases, summaries of the ideas may be appropriate. Thus, for the main part the reader of this autobiographical account will read what I have written, not what has been written to me.

Five Bishops of the Anglican Church did reply to the above letter. I found their letters kind, sympathetic and supportive.

But firstly, what is it that I am protesting?

A Golem is a clay figure brought to life by magic. Many people, academics mainly, believe that a human being is just such a Golem – a material object, a machine, a mechanism, animated by "nature". It is the goal of many scientists and philosophers to create an automaton or robot that is to all intents and purposes human.

This narrative is an example of a kind of writing known as an *apologia* (an "apology") – "a formal written defence of one's opinions or conduct." The popular meaning of the word "apology" conveys the idea of "a regretful acknowledgement of offence or failure," but the genre in which this writing is conducted does not imply "regret", so I am not apologising for anything in that sense. The most famous example of an *apologia* is the account given by Plato of the speech of Socrates at his trial. As is well known, Socrates was condemned to death by drinking hemlock, and though he could have avoided this punishment, he chose to abide by the verdict of the court rather than attempt escape. Socrates apologies for nothing in the modern and popular sense of the word. He defends his faith that obedience to conscience takes precedence over the desire for life and justifies his role as the scourge of men who pretend to know everything as a pretext to the exercise of tyranny.

An apologia does convey the idea that there is something that one has done, or in my case, is doing, that would *appear in the eyes of the public* to be in some sense a disturbance of the peace, an extraordinary way of going about things that ought not to be undertaken. This is very much the case here, for I am at this time going about announcing ideas in ways that may appear unusual, writing to people that I do not know personally, and, by the tone of my letters, not merely asking them to deal with something, but one might say, even *commanding* them to respond. One may ask, *by what authority do I do this?*

At it happens, conduct of this kind is not unusual, and has been a feature of every age. In contemporary times, outrage at the destruction of the planet and its species by global warming, pollution and wanton human exploitation of nature does rightly evoke protest. My apology is a justification of my protest, and especially of the manner I am going about it, which, I hasten to remark, is entirely legal and, contrary to appearances, according to custom.

As well as making an apology, then, in this classical sense, I am also making an appeal. I am appealing directly to the public to support my protest. The protest is not an easy one to make. It involves ideas that would at first glance appear to be difficult and recondite. What on earth does such an intellectual matter have to do with me? – is a natural response. It is my purpose to explain that this is not a minor issue, but in truth *the only issue on which the entire survival of the human race depends*, and that is of course my justification for making such a fuss about it. I believe must educate my reader about matters spiritual – explain to my reader the underlying and real foundation to life.

The great philosopher Kant wrote that the unavoidable problems of pure reason are God, freedom, and immortality. Apparently, contemporary academic philosophy has nothing to do with any of these. But this disregard for the fundamental questions is illusory. Contemporary philosophy appears to have nothing do with it precisely because it has *everything* to do with it. It is the background assumption of the dominant academic philosophy that (a) there is

no God (atheism), (b) that all events in nature are determined (no moral freedom), and (c) that the person (mind, soul, spirit) is annihilated at death (denial of immortality).

I say "dominant" academic philosophy, because from my experience, throughout the epoch in which this dominant academic philosophy has arisen there has always been a minority opinion. Not all academics are convinced, but the voice of this minority has been suppressed. The evidence indicates that the protest of the minority is confused, and at the same time, where the protest comes, it is met with strong opposition from the dominant party, whose conduct may be said to be bullying. The minority voice is being bullied out of expression.

Mine was just such a voice. I could describe myself as an ostracised academic. And ostracism has not been of necessity my fate, but rather a condition that I have accepted, even willingly, as a consequence of non-conformity, my refusal to go along with the dominant academic consensus, for where I have incidentally contributed to their academic program, my offerings have been gratefully received and acknowledged. But the time for meek and silent submission to the will of a majority that would if unchecked lead humanity into the ultimate blind ally of a society without mores, wholly given to the fear of death and the egoism that accompanies it, is over. No more silence in this matter. The protest has been made, and will continue to be made, until it is heard, and a proper response is elicited.

I am not the first to make this protest – the protest itself has a history. It seems that the last clear minded thinker to make it was Dostoyevsky, who told us that if God does not exist then everything is permitted.

> "... he [Ivan Karamazov] solemnly declared during an argument that there was absolutely nothing in the whole world to make men love their fellow-men, that there was no law in nature that man should love mankind, and that if love did exist on earth, it was not because of any natural law but solely because men believed in immortality. He added in parenthesis that all natural law consisted of that belief, and that if you were to destroy the belief in immortality in mankind, not only love but every living force on which the continuation of all life in the world depended, would dry up at once. Moreover, there would be nothing immoral then, everything would be permitted, even cannibalism. But that is not all: he wound up with the assertion that for every individual, like myself, for instance, who does not believe in God or in his own immortality, the moral laws of nature must at once be changed into the exact opposite of the former religious laws, and that self-interest, even if it were to lead to crime, must not only be permitted but even recognized as the necessary, the most rational, and practically the most honourable motive for a man in his position." [The character Miusov is speaking. Fyodor Dostoyevsky, The Brothers Karamazov, trans. David Magarshack, Penguin. 1958, p. 77.]

All human morality depends upon the desire for a *better state*, a *purer state of being*. Please note that I do make any assertions about *sin* or anything like that; I talk merely of the desire for greater purity. There is a very good understanding of this problem in the work of William Blake, his notions of the states of Innocence and Experience, and the desire to return to or achieve a higher state of purity in which one transcends the bitter state of Experience to a state of recovered Innocence. I cannot digress further here to quote from or analyse his work, but the reader who familiarises him or herself with it will be inspired.

But we soon realise that this purity, while it may be partly expressed in life, only has meaning in the context of a purer state of being that transcends this life, and we call that state *immortality*. Unless we can attain in some measure *purity* and *immortality* there is no meaning whatsoever to the desire for purity; purity is then for materialists, who must be cynical in this matter, just a muddy word for confused self-love and self-delusion. Without the belief in purity, and its attendant concept of immortality, then as Shakespeare says,

> Humanity must perforce prey on itself,
> Like monsters of the deep.

And that is precisely what is happening right now. This is why everything is going wrong – cramped society, accumulation of wealth in the hands of a disinterested super-rich, lying rationalisations of such institutions, narrow-

minded self-interested thinking, banking scandals, corruption in dark corners, ineffective and failing political institutions, democracy and decency all in decay. They are in decay because, convinced or being convinced, that there is no state of immortality and no state of meaningful purity worth struggling for, people are increasingly devouring each other, like monsters of the deep. Interpersonal relations are at a low ebb.

Of course, I am aware that not everyone will agree with this frank assessment of the evils of our world. Materialists will affirm that morality can be built upon the ground of subjective affirmations; others will object that only 16 per cent of the world's population explicitly avow atheism. But let my statements stand at present as an opening statement of my perceptions and serve to motivate interest merely in the sequel.

The announcement

30 March 2019

Letter to sixty-five computer scientists, mathematicians and philosophers, including the committee of the Turing Test.

Black's Academy announce possible solution to the Halting Problem

I enclose (a) the press release concerning this solution; (b) the introduction to the paper; (c) a list of recipients of this letter.

In such a matter it is appropriate to write to individuals to ask them to address the issues raised by the paper I have published. Computer science has in our contemporary age become immensely powerful – it constitutes a movement all its own. Its power to transform human reality is immeasurable. It is not my purpose to comment on all such transformations, but one in particular – the way computer science is transforming our ideas of human identity and our attitude to spirituality.

That I write from a broadly religious perspective should not be a cause for concern. My point-of-view is generally religious, but I am not a member of any church, nor do I belong to any faith. Sooner or later the implications of the mechanistic model of human nature that is so evidently promoted by computer science was bound to reach down to the public at large and trigger a crisis of faith. As professors and Doctors of Science you do have a responsibility to address such concerns.

It is evident that if a machine is deemed to have passed the Turing Test, then to all intents and purposes human nature will have been shown to be mechanical. That implies that human identity is annihilated at death, and likewise suggests that all faiths that are founded in Scriptures are false as to that foundation. Hence, computer science and academic philosophy cannot evade the moral responsibility for such changes in belief and must reach down to the public at large and offer clear guidance.

I find the claims made for computing to be wholly exaggerated and unfounded. I find the hagiography that surrounds the founders of the science, such as Alan Turing, to be unjustified. I do not see in Turing's prediction made in his 1950 paper anything other than bald assertion, akin to an article of faith, and I observe that 70 years later no machine has yet "passed" the Turing Test.

In the paper I have written I have presented a clear argument that the Halting Problem is solvable by mathematical induction. I invite you and your colleagues to evaluate it. I offer to withdraw my claim publicly should a clear demonstration of an error in my reasoning be provided. That there is a case to answer is clear; why was it that mathematical induction was never considered as a tool to the solution of such a problem? It is not madness to assert that the problem can be solved by induction.

I trust that should you conclude that my argument is sound, then you will not hesitate to make such an opinion known, so that the old conflict within the mathematical community can be resolved, and humanity may once again find faith in its spiritual transcendence over "dark", that is non-conscious, matter.

Finally, I am using a nom-de-plume here for many reasons, one of which is that the only thing that matters is the truth.

Yours sincerely,

Peter "Melampus"

The problem of Silence

The recipients of this letter derive mainly from the mathematics, logic and computing faculties of Oxford and Imperial College; some researchers chosen at random from Google, the committee for the Turing Test (the Loebner Prize), and other selected leading advocates or researchers in the field of artificial intelligence. For the list, please see the appendix. None of the recipients are known to me personally to be ideologically committed to the cause of the Golem. That the members of the committee for the Turing Test (the Loebner Prize) have some commitment to the cause of strong AI might be inferred but is not certain. However, there are many, many philosophers, scientists and thinkers who are committed in print to the belief in strong AI.

A second letter has been sent to the same sixty-five people. Until this time, only one has replied. To my knowledge the two papers (the "solution" to the Halting Problem and a second paper with solutions to 98 "hold-out" problems, which shall be explained below) have been downloaded four times and five times each respectively. Therefore, it is possible that some evaluation of the claims I am making is in progress.

I think it is self-evident that so large a claim should be evaluated, even if the provenance is so unusual – for instance, the papers have not been submitted to any academic journal, which has become the customary method by which new ideas are introduced to the scientific community. I am also presuming upon the time of academics who may deem the intrusion unwarranted. However, the claim as to the "bigness" of the result, if true, does warrant the intrusion. It is also correct to call anyone to account to clarify their philosophical thinking. The people to whom I have written are professors or leaders of industry and do have a moral responsibility to the public at large to justify their ideas; they are people of influence and must use that influence wisely. Therefore, I do not believe they should ignore the claim.

Besides, if the claims I am making are in their eyes the work of a madman, or a religious "enthusiast", then what simpler course of action is there for them to merely point me in the direction of the truth? They could write privately or publicly that the claim on page such-and-such, line such-and-such is false; that there is such-and-such a result that proves the method is invalid. Thus, in this case the silence of the academics is strange.

It has been suggested to me that six weeks is not enough time to expect an answer from academics who are busy with their own affairs. Or, perhaps they are ring-fenced by secretaries and have not even seen the letter?

It seems to me that the reader may take this matter in several ways. Apart from complete indifference, which also must be explained, one reader may see this at the level of a human drama – without seeking to understand the details of the claim, he or she will understand that a person professing to be a *knight of faith* has come into the field, blown his trumpet and issued a general challenge to the academic community to bring forth a champion to do combat with him. Yes, and if this analogy surprises you, then indeed I regard this as a conflict situation where the metaphor of combat is appropriate. That the conflict is utterly not of my making is part of the apology I present here. The need for the challenge is entirely the fault of those who have willy-nilly pressed ahead with a *bad idea*, and hoisted it upon

the public, contaminating the ideation of the public through-and-through, and a champion must come forward to challenge it. And if all this be false, then what is simpler for them then to produce their champion or champions and reign down their mortal blows upon my arguments? If on the contrary the claims of strong AI are vacuous, then let impartial judges own up to it immediately.

Thus, one reader may see this as a kind of sport – where indeed conflict also occurs according to the rules of sport – and grow curious as to the outcome. This is something I would very much welcome – it is part of the appeal I am making. Such a reader may say, if Melampus is wrong, why doesn't an academic at least come forward and say so? If he is right, then why does not some academic admit as much? I welcome the curious. Curiosity will save us.

Another reader may wish to take the matter further – to understand in detail the arguments I am making and evaluate the claim for him or herself, at the "technical level". For such a person, who is also most welcome, the option to download the papers from the website is freely available. But I shall also endeavour as we progress here to explain in so far as I can and in layman's terms what the arguments are.

Henceforth, I am dealing with two kinds of reader – the curious and non-technical, and the curious and technical. The indifferent readers have by now already given up. As the non-technical reader may find the technical parts distracting, I shall mark those technical glosses with an asterisk, so that the non-technical reader can skip them in the first instance.

* The Turing Test and the Turing Prediction

Before presenting the press release, let me first elaborate on a claim made in the letter to the academics – namely, the one about Alan Turing's 1950 prediction. This will profoundly illustrate the muddy thinking that surrounds the problem of the computer Golem. Strong AI is the claim that digital computers can "think" as we humans do, are "intelligent" and "conscious" just as we are. The idea of a machine imitating a person has become known as the Turing Test.

Turing Test: If a machine can imitate human behaviour to such a degree that human beings cannot tell the difference between the responses of the machine and responses of a human being, then the machine will be deemed to be conscious and to think. In this test the physical appearance of the machine shall not count, and therefore may be disguised to prevent identification on that basis alone. A machine that fulfils this criterion shall be said to have passed the Turing Test.

Turing made his prediction now more than fifty years ago: -

> I believe that in about fifty years' time it will be possible to program computers, with a storage capacity of about 10^9, to make them play the imitation game so well that an average interrogator will not have more than 70 per cent chance of making the right identification after five minutes of questioning. The original question, "Can machines think?" I believe to be too meaningless to deserve discussion. Nevertheless I believe that at the end of the century the use of words and general educated opinion will have altered so much that one will be able to speak of machines thinking without expecting to be contradicted. I believe that no useful purpose is served by concealing these beliefs. (Turing [1950])

One thing to note about this prophecy is that it is more cautious than these latter-day prophecies – Turing does not say that the machine will be alive and conscious; he says vaguely that people will ascribe "thinking" to machines, using the muddy word "thinking". Nonetheless, this prophecy has passed into academic and thence into contemporary consciousness as a confident expectation, and by "thinking" is meant life, self-awareness and consciousness.

An argument of a kind is offered in the above passage of Turing, but this argument may quickly be demonstrated to be circular, in that it assumes more than what it sets out to prove.

> Premise: in the future computers will have a large storage capacity of about 10^9.
> Conclusion: Machines will be deemed to be able to think.

The premise is an extrapolation into the future concerning a physical characteristic of machines. The conclusion is that human reason (a mental characteristic) is a species of physical computation (computation with 1s and 0s in a binary, digital machine). Hence, the premise can only be relevant if it is assumed in advance that human reason is a species of computation, that mental characteristics are derived from physical characteristics, but since that is precisely what the "argument" sets out to prove it cannot be assumed as a premise without committing the vicious circle fallacy. What a muddy thinker Alan Turing is! A genius in his own field of computing science, he had no philosophical training whatsoever, but this has not prevented the above prediction from muddying the waters of academic thinking.

It turns out that the only way to prove that human reason is a species of digital computation is to produce the machine that succeeds *all the time* (not just 70% of it for five minutes) in the "imitation game". No such machine has been invented. There is no premise that makes the fulfilment of this prediction inevitable. The appearance that *it is just a matter of time before the Turing Test is fulfilled* is wholly an illusion.

Hence, this reveals the Turing Prediction to be exactly what it really is – an expression of faith in materialism – not a rational belief at all – in the sense of being grounded upon incontrovertible premises. The proponents of strong AI have not demonstrated in advance that human reason is a species of computation, but they believe this because they are already ideologically committed to materialism. It is an article of faith with them that the machine can be built.

What is really happening is that there is an ideological conflict engendered by a new religious faith arising amongst us – the faith in materialism. What is meant by this is faith that all events are brought about by a non-conscious matter operating according to uniform laws of nature that just exist. This also entails atheism – the materialists have replaced god in explanation by nature, saying, "Thou Nature, art my goddess!"

But materialists are in bad faith, because they do not present, even to themselves, their ideology as faith, but delude themselves into thinking that their faith is grounded in reason – that science tells them that there is only dark (non-conscious) matter in the universe. Science, of course, tells them no such thing. Science is silent as to metaphysics and philosophy and is compatible with all faiths. Science and materialism do not go hand in hand, and it is an illusion of academic materialism to suppose that science entails materialism.

We will examine this problem further when we later look at the relationship between materialism and the Darwinian theory of evolution.

The Press Release

Let us proceed to view the Press Release that I made also on 30 March 2019.

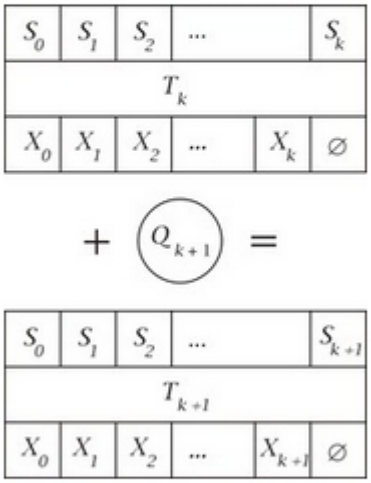
Press Release

Possible solution to the Halting Problem

The Halting Problem is solved by complete induction on the number of states in any Turing Machine. If true, a computer cannot solve the Halting Problem, but human reason can.

Under the imprint of "Melampus" Black's Academy, a publishing company, announce a solution to the Halting Problem. They invite the mathematical and computer science community to evaluate it. Melampus has also published solutions to all 98 unsolved five-state Turing machine hold outs using the methods provided in the main text. The papers are available for download at:

www.melampus.name



Since it is proven that a digital computer cannot solve the Halting Problem, this result would demonstrate that the claim of strong AI is false: a computer cannot emulate human reason.

The Turing Test is the idea that if a computer can "fool" a human into thinking it is human, then to all intents and purposes it is human. Any solution to the Halting Problem would rule out a computer being able to "pass" the Turing Test. A computer cannot solve the Halting Problem, but human reason can.

The abstract of the paper is simple: "The Halting Problem is solved by complete induction on the number of states in any Turing Machine. It is inferred that mathematical induction is a species of synthetic reasoning."

Complete induction is the main tool of that branch of mathematics known as number theory. Hence, if true, the solution demonstrates that number theory is not a species of computation.

Melampus agrees with the opening statement made by the late Professor Edward Nelson (Princeton University) when he wrote in *Warning Signs of a Possible Collapse in Contemporary Mathematics*: "I rejoice that we live in a world of boundless, infinite possibilities, one in which with Blake we can see a world in a grain of sand and a heaven in a wildflower, hold infinity in the palm of your hand and eternity in an hour". Conflict within the mathematical community as to whether all mathematics is a form of computation is at least one hundred years old.

The Introduction to the paper may be evaluated by any competent mathematician in less than 15 minutes. If true, the result demonstrates that mathematics is not a species of computation and would do so using the simplest of methods. A proof by mathematical induction does not seem to have been attempted before.

Melampus also comments: "The impossibility proof is a proof that a computer cannot solve the Halting Problem and no more. It may not be assumed without circularity that human reason is only capable of solving at most all those problems that a computer can solve. In solving the Halting Problem by mathematical induction, we not only demonstrate that the human mind can solve a problem that a computer cannot solve, we demonstrate that mathematical induction is not a computable algorithm."

To comment or ask for further information, please contact melampus@blacksacademy.net.

Does silence imply consent?

It is not modest to claim more than is strictly true. Thus, I wish to avoid hubris while admonishing the hubris of others – for it is hubris to claim that man is a species of machine, for it elevates human pride in its technology, thus elevating Man to the status reserved for God.

Thus, strictly, this is the state-of-affairs: Firstly, I claim to have solved the Halting Problem by a method of proof known as mathematical induction. Secondly, I am confident that this solution is valid, even as to technical details. Thirdly, in the course of developing the proof, I have also made observations not known to me to have appeared in the literature extending our understanding of computers (Turing machines). Fourthly, the claim has not to this date even been challenged, let alone refuted. Fifthly, the silence of the academics is strange and requires explanation, and I feel even emboldened to observe the dictum that applies in law – Qui tacet consentit" – "Silence gives consent".

But, my dear reader, wait, for at this very moment of writing, I have finally received my first holding letter on behalf of Dr. Demis Hassabis of DeepMind. I thank *DeepMind* for this letter. I note also the many wonderful things that

companies are attempting to achieve with artificial intelligence. The future of computing does not necessarily lie in interfering with religious beliefs or setting up a counter-religion of atheism. It is not useful to humanity to have too definite beliefs about the finality of human existence. Doubt and faith are the ground of healthy human existence and mores.

Does silence imply consent? I think it might imply wilful indifference or even arrogance. It might imply a wilful refusal to consider an idea that is disturbing. Your guess is as good as mine.

* About mathematical induction and why the proof is "almost certainly valid".

The paper proposing the solution assumes that the reader is familiar with (a) Turing Machines and (b) proof by mathematical induction. There are notes in the paper on Turing Machines, and the introduction to the paper may be read by someone with only a modest knowledge of mathematics.

However, here I will attempt to explain to the "layperson", the non-specialist, the essence of the proof, and thereby why it cannot be ignored, even if it is only to be refuted. A refutation might be embarrassing to me, but the whole process would so advance our understanding of science, mathematics and religion that I could hardly be said to have done anything other than been a benefit to mankind. However, my grounds for general confidence are deep, as I shall explain.

Let us begin with mathematical induction. This relies on the concept of the potential infinite. The term "potential" is added here because it is an essential detail that first appeared in the work of Aristotle in his *Physics*, so I will continue to use it, though the general reader may simply read this as "infinite". Let us start counting, 1, 2, 3, 4, … . For how long does this list of counting numbers last? We realise that the process of *adding one more* can be repeated *indefinitely*, and thereby reach the concept of the potentially infinite. Potential, because it is based on the idea of being *able* (having the potential) to always add one more to the list of counting numbers.

It is a curious aside to remark that my researches into the history of ideas indicate that until the 6th century B.C. mankind did not have the concept of infinity. For example, the Egyptians and Babylonians could think of enormously large numbers, but they had no symbol or concept for the infinite. The concept of infinity was introduced into human ideation in the C6th BC in Greek Ionia. I call this the *Ionian Revolution in Consciousness* and have written separately about it.

Mathematical induction is one of the fundamental methods of proof in mathematics and works by attaching a result to the process of adding. A result is attached to a starting number, for example, the number 1, and then it is shown that if the result applies to that number, then it applies to the next number up, and so on *ad infinitum*.

Let me give a very simple example of a geometric proof that uses induction. In graph theory a *tree* is a diagram in which there are lines (edges) connected points (vertices). A *tree* is specifically a connected graph in which any two vertices are joined by exactly one path (edge). It is a theorem that in any tree the number of vertices (V) is equal to one more than number of edges (E):

$$V = E + 1$$

This is proven by mathematical induction. Firstly, consider a tree with just one edge (figure 1). It is obvious that this has exactly two vertices.

Figure 1. A tree with one edge ($n = 1$) *and two vertices.*

Now consider that we have a tree with n edges and that we wish to add one more. We suppose that for the tree with n edges we have $V = E + 1$. But when we add one more edge, we add exactly one edge and one vertex (figure 2). Hence, for $n + 1$, we also have $V = E + 1$.



Figure 2. The inductive step in a proof for all trees that $V = E + 1$.

This proof illustrates how many of the simplest things we know in mathematics are really demonstrated by mathematical induction which relies on our concept of the potential infinite. Mathematical induction is *everywhere* in mathematics. It is the foundation of number theory, so summations for infinite series are proven by mathematical induction. It is used in the proof of the binomial theorem. It is the foundation of calculus, being used in the proofs for example of the formulas for differentiation of polynomials, products and chains of functions. Etc. Etc. Etc.

Now let us turn our attention to Turing Machines. These were devised by Alan Turing and demonstrate that when he was not interfering in philosophy, he was a very great genius. These "machines" are simple descriptions of computers that involve single states which operate on a "tape" (a representation of the memory of the computer) and perform simple instructions. The point of the Turing Machine is that every digital computer whatsoever can be described as a Turing Machine – so whatever is true of a Turing Machine is also true of any computer.

Turing Machines have "states", and these can be numbered. The simplest Turing Machines have just 1 state. Then we can add to these 1 more state, and another, and so on *ad infinitum*. But though this can *in theory* be continued *forever*, in practice any real computer is equivalent to a Turing Machine with a finite number of states. Most desktop computers would be equivalent to a Turing Machine of, say, trillions of states. Turing Machines are used to analyse very simple computers and establish theoretical results about all computers, however large.

We all know from experience a computer crash. The computer just freezes while it has entered a loop. It appears that it will never exit this loop, so we press the "break" button to reset the computer and start again, probably cursing the computer, God, luck and the operating system for losing valuable work. This is akin to the Halting Problem for a computer. Will the computer halt or not? Has it really frozen, or does it just appear to have frozen, and will stop what it is doing sometime in the future and unfreeze?

The importance of the Halting Problem to theoretical computing is that there is a definite proof that no single machine could be built that could solve the Halting Problem for all computers whatsoever, including its own. This proof is an instance of a contradiction argument, which is another fundamental method of proof in mathematics and logic, and one that is in my opinion also not in general a computation. Laying that observation to the side for the present, the contradiction proof establishes a theoretical limitation on what a computer can do. If human reason can solve the Halting Problem, then human reason is not an instance of a computer. Human reason is not computational in that sense, and the claim of strong AI is false. A computer-golem cannot be built, and the Turing Test will never be passed by any computer.

Now let us revert to the press release. The reader will see that I have included in it a figure (figure 3). *This figure is an intuitive proof the Halting Problem can be solved by human reason using mathematical induction.* Furthermore, any competent mathematician will immediately recognise it as such.

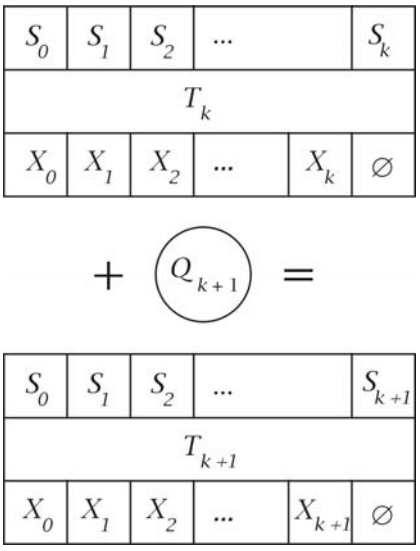


Figure 3. Diagram illustrating the inductive step in the proof of the Complete Criterion Theorem.

Note, the figure refers to something I have called the "Complete Criterion Theorem" – this is in fact an even larger problem than the Halting Problem, and the solution to the Halting Problem is an easy consequence (we say "corollary") of this theorem. It turns out that the solution to the Halting Problem is made *more difficult* by considering it in isolation from this larger problem. For the details of the Complete Criterion, I refer the reader to the text.

In the figure the topmost part represents a Turing Machine of k states as a "black box", whose internal workings are not immediately relevant. This black box is a machine that for given inputs churns out certain outputs. We are entitled to assume that the Halting Problem can be solved for this machine. (It has a Complete Criterion.) Then, in the middle of the figure we *add one more state* to the machine. Now such an extra state can only result in a certain number of modifications to the pre-existing machine of k states. Thus, in the proof, we simply exhaust all these modifications, and show in each case that if we know the Complete Criterion for the machine with k states, then we can find the Complete Criterion for the machine of $k + 1$ states. Hence, we solve the Halting Problem.

The proof could only run into difficulties if in the process of adding one more state the machine suddenly developed an utterly unpredictable pattern of behaviour. That this is impossible derives from the observation that what happens in a Turing Machine falls into cycles that have a finite period. However, it turns out that there is one modification to the machine which does seem at first to introduce unpredictable behaviour – I call it a *permutator* – but on analysis this is also shown to be as predictable and determined as the other modifications. Thus, in the process of developing the proof I make several additions to our current understanding of Turing Machines: (a) that they have a Complete Criterion, that is (b) based on cycles that have periods, (c) that there are methods of exits and inputs (see the text) that enable one to determine the Complete Criterion, and (d) how to understand the effect of the permutator.

It is true that there are many philosophers and scientists who have committed their reputations to the claim of strong AI. If they see the diagram, and have any competence in mathematics at all, as I presume them to have, then it immediately follows that they will experience a nervous shudder. The figure is all that is needed to indicate that *there is a case to answer*. Only if the details of the proof are wrong, if the understanding of the permutator for instance is false, only then can the conclusion be controverted. Human reason can potentially solve the Halting Problem for any computer whatsoever.

I say "potentially" solve, because there is a practical limitation. It has been demonstrated that there are some approximately 8 million Turing Machines of five states, and 7.5 billion Turing Machines of six states. The method

of inputs to solve the Halting Problem involves exponentiation and hence leads potentially to solutions that are very long indeed, and for some Turing Machines, solving the problem would in some sense exhaust all the physical resources of the universe. However, that does not demonstrate that the Halting Problem cannot be solved, because the proof is absolutely without reference to any physical constraint. It establishes a method that works for *all machines whatsoever,* whatever their size, and demonstrates that we *know* that we can solve the problem. Furthermore, this situation is the same as the one that pertains in all mathematical inductions, for it is never possible to get to the end of all the finite numbers – to reach infinity – simply by counting up. Thus, this constraint is the same one that applies to all mathematical theorems established by induction and does not controvert the proof.

Now a new question arises – the proposed solution to the Halting Problem uses mathematical induction and a background knowledge of Turing machines that would take the average student at most five hours to master. It is not "rocket science", it is not erudite or difficult or recondite or anything like that. How is it possible that in the whole course of the history of ideas, no one else has ever thought of using it? The mathematics involved is scarcely beyond "A" level in the British system, or High School in the American. Furthermore, the person writing this is sixty years old, and began studying the problem thirty-eight years ago. Why did it take so long?

The immediate background to the announcement

This apology has two themes – firstly, that of my reasons for believing in the truth and relevance of my claim to have solved the Halting problem; secondly, my justification of the methods I am using to announce that claim. Here we revert to the second theme.

I have explained that I am an "ostracised academic". By this I mean that I do not have any academic affiliation – I do not belong to any university or research institute – I am an "independent" working in isolation and am now cut-off from all academic contacts. But this has not come about because I have wished it so. This is the story of the failure of my other attempts to achieve an evaluation of my ideas. Some of this is very personal, as friendships have been broken over the matter. I don't think it will be too easy to explain it, and I will not attempt to explain it all.

I took up the study of the Halting Problem while an undergraduate student of philosophy thirty-eight years ago and wrote a dissertation on it. What happened then constitutes the deep background to my personal drama and that part of the story shall be told below. At the time I first took it up, my mathematics background was insufficient, so subsequently I studied mathematics obtaining a top first as an undergraduate and then an MSc. I would have studied for a PhD in mathematics, but economic and other difficulties got in my way, and I had to drop the study of higher mathematics for fifteen years in order to work incessantly. Thereafter, working on my own I recovered all my maths background, and then pushed into the study of the mind/machine problem.

By this time, I had dropped the direct assault on the Halting Problem and Turing Machines in favour of another approach. But about six years ago, while working on the problem of the continuum, the idea suddenly flooded into my head, *why not simply use mathematical induction?* My first paper took just seven days to thrash out, but such a direct assault on the problem, while devastating in its impact, is so hot in the political sense of the word, that I preferred at the time to shelve the work and pursue other approaches, hoping to establish results in other areas first before returning to it.

Two things then happened to make me realise that the time had come to do something decisive. Both are connected with my last "friend" in the academic world, whom I will not name. For years we had been meeting in my apartment for discussion. I would cook sumptuous meals (I am a good cook) and as he is a bit of a wine-buff, I would go out and buy the sorts of wine that I don't normally drink because of the expense, though of course, wines I very much enjoy when I do drink them. Now though we had been getting together for several years, from an early stage I had

realised that this was not going to lead to the kind of collaboration that I had always dreamed of. Firstly, I could never understand his papers, save perhaps the very first – he is a theoretical physicist. Secondly, he ignored such suggestions of mine as – why do we not meet up and take it in turns to teach the other everything we know? Thirdly, though he had by some instinct a sympathy with my general point of view, which he did grasp, his contributions to my project tended to be niggling criticisms rather than constructive remarks, though they did stimulate me to an extent.

But there was another benefit of these meetings – from time to time my friend would bring to my attention a paper of another academic that would stimulate my own ideas. Fairly recently, he drew my attention to a paper by the now late Professor Edward Nelson, protesting the formalism of the day, and arguing a point sympathetic to mine – that mathematics is not a species of computing. When I looked at the book, I saw above it the end of an article about research into Turing machines. I wrote to the eminent author (he is a Fields Medal winner) and received an immediate reply – he told me that his information was taken from Wikipedia. This in turn directed me to the paper by Own Kellett on the five-state Turing Machines, supervised by Soren Bringsjord of the Reneslaar Polytechnic Institute. Their work had reduced the 8 million Turing Machines of five states to just 98 unsolved "holdouts". Right, I said to myself, if my ideas are correct, I should be able to solve all 98 of those problems. And I believe I have. In the process I rewrote my paper, and that is the work you may now download from my website. There is an additional paper with the solutions to the 98 "holdouts" as Kellett called them.

Now it is not likely that anyone in such a situation will want his work to sit on the shelf for another six years, but then my "friend" produced another stimulus to action. In addition to claiming to have solved the Halting Problem, I also claim to have solved the problem known as the Continuum Hypothesis, an altogether more difficult problem. To solve this problem, I claim that it is an error to construct the continuum out of points alone; we need two primitives, points and extensions. Let us view this purely has a human drama: Melampus had an idea and he told his friend about it.

Melampus then wrote a paper specifically for his friend, and his friend read it and said he did not understand more than a third of it, but he did cite it in an essay of his own on the continuum and at a local seminar. Then the friend wrote another essay that he has submitted to one of the journals. Melampus and his friend quarrelled after two or three bottles of wine over a trivial matter, and their friendship ended. In the sequel the friend wrote to Melampus stating that he has removed all the work of Melampus from his computers, and also stating that he felt no need to reference Melampus in "the introductory section of the manuscript I gave to you yesterday"; he prefers to attribute "the problem with the extension in the current formalization of the continuum" to Weyl.

I decided that the time has come. The reader will learn below that ideas of mine that were suppressed before did subsequently appear independently in the writings of other prominent academics. If indeed the continuum cannot be understood without the concept of the indestructibility of extension, then the first person to say that explicitly is not Weyl but myself.

Weyl was a very great man, and very important to the development of my own ideas. I have read his book, and it was a stimulus to the development of my idea about the continuum and the indestructibility of extension. My friend initially derived this idea from me, and I cannot permit the omission of some acknowledgement of that to pass without comment.

Therefore, I decided that I would announce my result to the Halting Problem. Being by this time bereft of all academic connections, I decided that it would be a good idea to write privately to someone who might be sympathetic. There were just three possibilities: John Lucas, a philosopher and theologian of Merton College, Oxford; Roger Penrose, head of the Mathematics Research Institute of Oxford, and Soren Bringsjord, professor at RPI and supervisor of the Kellett paper from which I derived my result for the 98 holdouts. John Lucas is someone I know, I have been in touch with him before, he attempted to share academic connections with me on a previous occasion, those connections either did not reply to my letters or ceased replying, he is very elderly and is not a mathematician. I wrote to Roger Penrose on a previous occasion without reply. He is arguably the foremost theoretical physicist in the world, but is surely ring-fenced by secretaries; therefore, not contactable. So, I wrote to Soren Bringsjord twice. Neither letter was

acknowledged, so I have no way of knowing whether he received the letters or not – they were emails. The two letters announced my results and asked him to advise and/or collaborate. They are of course very polite letters, but I will not cite them in full here, as in many ways they repeat information that has already been provided.

When a person produces a result, he prefers to have a friend with whom he can consult. If he is wrong, then his friend will tell him so in private. There is no loss of face, no humiliation, and perhaps the result can be patched up. Failing that, he reaches out to some man or woman in a prominent place, a "sage" or "great man or woman" or "lord or lady" to whom he can make his appeal.

Six or so years ago I was working on a result known as Gödel's theorem, and it led me to write a whole book that I called *Poincaré's Thesis* after the eminent early C20th mathematical physicist. (That book has now been broken up into several separate papers.) Poincaré advanced several theses, but the one that I am referring to here arose in the context of his objections to the work of Russell and Peano. It was Poincaré who first explicitly came up with the idea that mathematical induction is not a form of computation – that it has a synthetic character – so I attributed the idea to Poincaré, as I should. We'll have a look at this too below.

Then I went to see my ex tutor. Some years' before I had donated to the College £1,000 and I think he may have thought that I was interested in making another donation. At the time, my financial situation was a little straightened, and I would not have easily found that kind of money for such a purpose, but I was told all about the boat house. No, I'd come to ask him for a conversation and a contact. I said I'd like to chat with someone about ideas to do with Gödel's theorem. It was not my intention to mention the Halting Problem, for I knew that to be a very sensitive topic, too hot even for a talk with my friend. But I inadvertently said that the Halting Problem could be solved – I said nothing else. I did not get the conversation, and we have never been in contact since – neither he nor I has communicated. Did I dream it or not? I am under the impression that as we parted I was being marched to the doors of the college and being cast out. He pointed me to the bookshop and advised me to buy a book on Gödel's theorem by one of his colleges. Thirty plus years before I got a first on the Mathematical Logic paper while doing my undergraduate degree, and he ought to have remembered that. He certainly knew I had two degrees in mathematics by that time as he himself added them to the college record. It is sad, sad indeed, when ideology gets in the way of friendship and natural affinity.

Personal appeals don't work, and there have been others. Meanwhile, what of the idea of submitting the paper to a journal? Here we encounter the immediate problem that a journal generally takes up to a year or so to review a paper, and there is no guarantee that it will publish. In the meantime, one cannot submit elsewhere. I also do not think that a solution to the Halting Problem should become the property of a journal – the result is too significant for that. But there is another reason why I did not take this customary option.

In the ideal world politics and scientific (or philosophical or mathematical) truth do not interfere with one another. The truth is the truth is the truth, and respecting the truth is to scientists as the Hippocratic oath is doctors. (Whoops! I almost wrote "hypocritic" – Freudian slip.) But that ideal is all hogwash in the real world, as in our even lesser cynical moments we must acknowledge. The whole thing has been admirably analysed by Kuhn in his justly famous book, *The Structure of Scientific Revolutions*. Now for my brief summary of this work. Most science is not ground-breaking revolutionary stuff, but what is called "normal science" by Kuhn – in other words, just churning out results according to agreed procedures – Kuhn calls these "paradigms" – and this is what most scientists are doing. And when they do so, they look to their colleagues and are basically copying what everyone else is doing. *There is a great deal of fear in the scientific community* – fear of stepping out, fear of rocking the boat, fear of being not-normal, fear of upsetting this or that eminent person, or funding body. Scientists are under a great deal of pressure to publish, for otherwise they become a kind of nothing. Thus, when there is a new idea, and one that challenges everything, and strikes at the foundations of beliefs, and holds up attitudes to wider scrutiny, the scientists cannot be relied upon to get it right, because they are afraid of getting it wrong. The truth is subordinate to political considerations, and when the context is religion, and a book is placed on the Index, we call that *censorship*.

Some few years ago I decided upon a philosophical approach. The philosophical position that all mathematics is a form of mere computation is known as *formalism* – it might also be called *the empirical theory of mathematics.* I decided to write an essay in which I would go through the types of mathematical reasoning that are found in mathematics and argue the case for the alternative, and I called the paper *The Rationalist Philosophy of Mathematics*. I submitted it to a journal, and some four or so months later got short reply stating that they would not publish the paper, saying that the essay was discussed "orally, and at some length" and refusing to make constructive comments.

I would like to make some observations about this letter. Firstly, the paper I submitted is said to have been discussed at length. Does that imply that there was a minority view in favour of publication that was overruled by a majority? Secondly, the author declined to offer any constructive comments. Given that the paper was discussed "at some length" that is surely odd, and contrary to custom. If the paper generated discussion, then it surely had some merits and could have been improved. What were the reasons for not publishing the paper? The reader will observe that *no reason was offered* for not publishing the paper.

The paper can be downloaded from my website. This takes us into a discussion of the dialectic.

Before I go into the topic of the dialectic, this is the right moment to make some explanation as to why I am not using my birth name, but a non-de-plume – Melampus. This has something to do with an oracular dream (the name was given to me in the dream), but I know that readers find talk of such things disturbing and tend to look askance. But here I will explain some of the prudential reasons for the use of a nom-de-plume. The reader will appreciate that my personal relations with "friends" and "acquaintances" have been adversely affected by my idealism – the break-up of two "friendships" has been described above. This problem extends even into my regular friendships, where there is no conflict, where even my friends share my general outlook. So, I've decided to separate my personal and my "missionary role". By and by, I have a great sense of humour, very ironic, and am quite capable of laughing at myself. Missionary indeed! What a laugh!

Another point arises in connection with the submission of the paper described above. All the journals describe the impartiality of their procedures, how they read the papers "double-blind". Not many people search for me on Google, but a certain academic website looking for money (what else?) tends to report every time it happens. Just before my paper was rejected, someone searched for me from the country where the journal is situated. It aroused my suspicions. You see, every now and then some famous mathematician on his deathbed wishes to protest the crushing atmosphere in academic culture of materialism and anti-faith. The essay by Edward Nelson is just such a case in point. My paper *The Rationalist Philosophy of Mathematics* is not some work of a junior high-school country bumkin; it is sophisticated and *might just have been perceived to be the work of some important mathematician.* I am not entirely convinced that science and politics are wholly separate. Of course, I make no animadversion against the journal to which I submitted, which I am convinced acted throughout in good faith. It is not named here in any event.

Finally, on this topic of nom-de-plume, I mention that it is a time-worn custom for authors to adopt such a device. Charlotte Bronte wrote as Currer Bell, Mary Anne Evans as George Eliot, William Gosset as Student, Soren Kierkegaard as Johannes de Silentio, Desiderius Roterodamus as Erasmus, and Martin Luther as Eleutherius. The choice of Greek name is also relevant; I thereby associate myself with a humanist tradition.

The opening sentiments expressed by the late Edward Nelson, a professor of mathematics at Princeton, that I partly quoted in the Press Release, are worth quoting in full. He wrote:

> I rejoice that we live in a world of boundless, infinite possibilities, one in which, with Blake, we can see a world in a grain of sand and a heaven in a wildflower, hold infinity in the palm of our hand and eternity in an hour. I rejoice that the sacred scriptures of our faith portray a God who listens to prayer, who loves us and longs to lead us. I rejoice that my chosen line of work, mathematics, has enabled me to bring into being new things that did not exist before and to greet with wonder and awe many amazing inventions of my fellow workers. I rejoice that daily we live immersed in infinity, that we have the freedom not only to make choices but at times to be the agent, by will or by grace, to sing to the Lord a new song. [Edward Nelson, Warning signs of a possible collapse of

contemporary mathematics from *Infinity, New Research Frontiers,* ed. Michael Heller and W. Hugh Woodin, CUP.]

Nelson then goes on to defend number theory in mathematics against the claim that it is a species of computation – just as I am doing, though on different grounds. We do not normally associate mathematics with great passion, but that is another thorough-going illusion. Nelson felt the need to make his protest clearly because he felt that mathematics is being crushed by atheism.

A red-hot issue - the dialectic and Darwin's Theory of Evolution

Plato introduced the dialectic. At its most basic level the dialectic takes the form of debate – it is the idea that philosophical understanding progresses through the opposition of opposed ideologies, which fight it out over the centuries and are modified and updated to their mutual advantage by the conflict. Hegel expressed this notion through the idea of the opposition of thesis and antithesis leading to a synthesis, and thence the synthesis becoming a new thesis. The history of ideas shows that in philosophy there have always been opposed schools of thought on any subject, and of these, arguably the most important is the opposition between empiricism and rationalism. Empiricism is the theory that all our knowledge comes from sense-experience, and rationalism is the opposed view that some of our knowledge is derived from ideas. Empiricism readily allies itself with materialism and since rationalists tend to account for ideas as originating with God, then rationalism allies itself with theism. Thus, we have:

Rationalism	Empiricism
Theism	Materialism

Plato in his dialogue the *Theatetus* described this as the war between the Gods and the Giants. Here I am writing somewhat loosely, and specialists might crow that I am making mistakes or false oppositions. My detailed exposition of all these ideas may be found in my work *Immortality: the Groundwork to the History of Western Consciousness* in three volumes. Here I am merely trying to give the non-specialist reader an idea about the background conflict.

In the Western tradition, rationalism has been associated closely with Christian faith and a religious attitude, and empiricism and materialism with the opposite. Indeed, it can be shown that the ideas of the first rationalist philosopher, Plato, were instrumental in the formation of the Christian theology of the early Church. But here I remind the reader that I do not identify myself precisely with Christianity, and that my work is motivated by a superordinate concern for religion, not for a particular Church. Also, I am very aware of the negative aspects of religion, both Christian and non-Christian, and I've written all about this in my longer work.



Figure 4. The dialectic c. 1800.

Now something very odd happened during the twentieth century concerning this conflict between the two halves of the dialectic. Up until about 1900 rationalism was in the ascendancy and most academic philosophers would have

accounted themselves Christians; by now, I think it is possible to say that *there is not a single rationalist philosopher working in mainstream academia.*

That may be an exaggeration, but if there is a rationalist school of thought, it exists *as a suppressed minority view*. The dominant ideology is empiricist through-and-through; rationalist sympathisers who have managed to eek out an existence in the cut-throat world of academia, do so on the margin, working not in logic or philosophy of science, but in other areas, perhaps the history of ideas, perhaps as existentialists. One exception to this rule might be Alvin Plantinger, who still defends an argument for the existence of God known as the ontological argument, but his is a relatively isolated case. I don't think either he or, say, John Lucas, could be regarded as mainstream academic philosophers commanding widespread following within the academic community.

The academic community has become monolithic and monolithic empiricist at that.



Figure. 5. The dialectic c. 2000.
A monolithic culture of empiricism has arisen in the wake of the Theory of Evolution.
Rationalism has been crushed to a mere marginal belief held outside academia.

The term "materialist" is actually rather muddy and complex, and some academics that I might with a broad-brush call "materialist" would cry out at such a nomenclature. Hence, "empiricist" is best to describe the flavour of the monolithic ideology. In place of "materialist" the term "mechanist" might be more acceptable. It will be hard to find a mainstream academic philosopher who is not at heart advocating a form of mechanism. Hence, the popularity of strong AI, which stands as the pinnacle expression of such an idea, for it makes man into a machine, a computer.

This monolithic academic culture came about through a natural shift in the dialectic that roller-coasted through society from the early nineteenth century onwards, and the appearance of Darwin's theory of evolution and natural selection is clearly a landmark of this process. It is simply that a way of seeing the world arose in academic circles and it became increasingly difficult for alternative views to gain acceptance. All the "big thinkers" allied themselves with empiricism and every advance in natural science was assumed to lend support to empiricism, and soon it became risky to defend the contrary.

I partly describe the history of this process in the third volume of my longer book, in the chapter called, *The Wilful Overthrow of Kantian Consciousness*. I have put that entire chapter onto the Melampus website so that it can be freely downloaded.

By now the reader will understand my predicament. As an advocate of rationalism, I am clearly swimming against the tide, rocking the boat, upsetting the applecart. Furthermore, if true, the solution to the Halting Problem will dramatically reverse this situation. That is because, it must constitute a defeat of a large part of the claim being made implicitly on behalf of the Theory of Evolution. This also accounts for my disappointments in my personal relations – academics who are naturally inclined to like, respect and in the wider sense of the word, love me, who are even under personal obligations to me, suddenly turn volt-face, a "quarrel" ensues and the friendship comes to an end. The reason for it is natural – conflict of loyalties leads them into an impossible situation. On the one hand they wish to remain my friend, and are either attracted to my arguments, or at least recognise that they are not obviously stupid; on the other hand loyalty to the group, commitment to the mainstream ideology, here described for want of more detailed analysis as "academic, monolithic empiricism" forces them to break with me.

So, let me explain why the solution the Halting Problem is such a red-hot issue in the dialectic.

The theory of evolution is primarily a theory of the development of physical characteristics – for example, it is supposed that whales evolved from dog-sized members of the extinct raoellidae ungulate family, for their ears, teeth and bones resemble those of whales. Since a large part of the survival of an animal species depends on its behaviour, it is assumed that behaviour evolves as an adaptation to environment, in much the same way as a physical characteristic – as a mechanical response to changing stimulus. Karl Vogt wrote in *Physiological Epistles* (1847) that 'the brain secretes thought, just as the liver secretes bile'.



Figure 6. Possible configuration of the dialectic
if the solution to the Halting Problem is accepted as valid.

But if the solution the Halting Problem goes through, then it will not be possible to advance the view that "the brain secretes thought". In fact, it will no longer in the ordinary way be possible to maintain that human reason (the mind, soul, personality) are mechanical in any absolute sense. (It may be possible to assert that aspects of the human mind are modelled by mechanical systems, an altogether different proposition that I could support.) The absolute point-of-view will be taken out of the dialectic. My next task is to explain why. My claim is that the solution to the Halting Problem, if accepted as valid, will ease the pressure on religion; there will be a substantial return to religion. That will be a good thing, for whatever the problems of religion, and that they are very many I will freely allow, man cannot live in a state of despair. Despair arising from the thought of his inevitable annihilation as a conscious being in death, rots joy in life and rots morality. Dostoyevsky is right, without the belief in immortality, everything is permitted.

* First and second order – why mechanisation of the mind does not survive if the solution to the Halting Problem is valid

What if the solution to the Halting Problem is valid? We will have shown that human reason (the mind, soul) cannot be replicated by a digital computer, but what follows of that? May it just not be some other kind of machine?

To answer this "objection" let us consider again Vogt's claim that 'the brain secretes thought, just as the liver secretes bile'. In this claim Vogt assumes the pre-existence of non-conscious matter. This would be material substance that has existed from the beginning and continues to exist; it is indestructible. The Ancient Greeks called it *dark matter*, meaning, matter devoid of consciousness. Then, in accordance with Vogt's materialist claim, certain configurations arise among the atoms of this dark matter, and, hey presto (as if by magic) we have consciousness – we have life, intellect and self-awareness. This is exactly what the proponents of strong AI are advocating – they claim that when they make an arrangement of non-conscious computer components (made primarily of silicon) and so long as it behaves as men behave, then, hey presto, you have a conscious *thinking* (vague word) intelligence. Just how this "happens" they do not deign to tell us – it just happens.

To understand the utter collapse of this explanation, we need some further facts about computers. I shall compare computing to a game of chess. Among other things, chess involves three components: (a) a chess board, (b) chess pieces and (c) rules telling us what moves can be made with the pieces. There are also the notions of two players, black and white pieces, and the idea of a competition. These additional aspects, that might be relevant as an analogy to the dialectic, shall not concern us here.

A computer may be likened to chess. In the place of the chess board we have an abstract structure known as a Boolean Algebra; this is also an instance of an abstract structure known as a lattice. In the place of the chess pieces we shall have arguments – deductions from premises to conclusions. The analogy is a bit loose here but let us think of arguments as moves of the pieces around the abstract surface (the "chess board") from one position to another. Finally, in place of the rules, we shall have "first-order logic".

In short, computation involves arguments made as moves on a Boolean algebra according to the rules of first-order logic. Chess involves games made by moving pieces on a chess board according to the rules of chess.

Thus, the arguments are the moves that can be made on the board (an abstract "surface") called a Boolean Algebra by moves allowed in first-order logic. And a Turing Machine is precisely that kind of machine that can make those moves. Furthermore, we can automate those moves. Any move that is a kind of argument made on the chess board known as a Boolean Algebra in accordance with the rules of first-order logic is mechanical. If all of mathematics can be played out on this board, then mathematics is a form of computation. If every argument or expression of human reason can be played out on this board under the rules of first-order logic, then human reason is a form of computer.

The solution to the Halting Problem demonstrates the falsity of this view. There is an argument, mathematical induction, that cannot be played out by a computer. Using mathematical induction, we can solve the Halting Problem, which no computer can solve. Therefore, human reason is not confined to first order logic; it is not constrained to moves on the lattice of a Boolean Algebra.

Then, we have a need for another kind of logic.

The description of that logic shall be "second-order or higher". (What this means shall be explained in a moment.) But being second order the idea that we derive consciousness from non-conscious dark matter collapses. Henceforth, all our explanations, including those of the deepest theoretical physics, must involve the mind *from the first*. We don't secret consciousness from dark matter, rather matter is intelligible matter – another term from the Ancient Greek epoch. What we understand of the universe is an amalgam of mental considerations (concepts) combining with a vague notion of something non-mental, the atoms perhaps. An atom cannot be conceived apart from the mind that conceives it. It is an end of that kind of mechanism forever. It is the end of the idea of dark matter. All matter must be intelligible matter – that is, *material substance presented to self-awareness (the conscious mind, soul or spirit)*; henceforth, we cannot have a physics that is also not a psychology; all physics is really a species of psychology.

For the technicians, the reason for this conclusion has to do with quantification. Now the idea of dark matter (matter devoid of any reference to the conscious beings that perceive it) has never seemed to me to be really a coherent idea – to me, it makes little sense. But I am willing to grant my dialectical opponent as much as I can, so I will acknowledge that if he can construct a language in which explanations are offered that make no references to the intelligible observer, then the idea is coherent, even if I don't personally like it. To construct such a language no reference or *quantification* must be made over concepts – the only quantification allowed shall be over "things", or as the technicians say, "individuals". This language is first-order logic.

Hence, if first-order logic fails to account for human reason, then in the ensuing second-order logic we quantify (refer to) properties and concepts; and we must have an explanation of the world that is thoroughly mentalistic as well as material. Hence, the solution to the Halting Problem not only collapses the idea that a digital computer can be a living human-being, but it also collapses the idea of a dark matter that can be described in a language that makes no reference to ideas.

As it happens, the idea of dark matter has always been suspect. The matter that we see and touch has always been *intelligible matter*. The proof of this is as easy as squidging your eyeball. When you move your eyeball the thing you are looking at *also moves*. You think that is dark matter you are looking at? No, it is merely a representation of an object that is "in" or "before" your mind. But this is not the place to elaborate upon the philosophy of perception. Furthermore, if science can describe objects in a first-order language and attain to reality through an objective method,

then perhaps we can talk of dark matter after all. Hence, it is the solution to the Halting Problem that really negates the idea of dark matter.

Thereby, the impact of the Theory of Evolution by Natural Selection on religious belief is also negated. There is no evolution of purely physical characteristics that accounts for mental (spiritual) characteristics. If there is an evolution (and I think there is) it is of *both the physical and the spiritual hand-in-hand*. Breathe a sigh of relief, oh religion, for thou are not negated by evolution. Spirit evolves.

* The history of an illusion

The illusion that the Halting Problem could not possibly be solved is due to historical and cultural forces that made it appear that human reason in general and mathematical induction in particular are species of computation. The history of this illusion is a subject in itself.

It is not assumed in the proof of the solution to the Halting Problem that mathematical induction is not an algorithm. However, there is no reason prima facie to assume that it is. It is self-evident that mathematical induction is a synthetic principle of reasoning since the conclusion it entails contains a concept of a potentially infinite collection that is not expressed in meaning in either or both premises. This point was made repeatedly by Poincaré in the early C20th in response to the rise of mathematical logic and was ignored. Poincaré wrote:

> ... syllogistic reasoning remains incapable of adding anything to the data given in it; these data reduce themselves to a few given axioms, and we should find nothing else in the conclusions. ... mathematical reasoning has of itself a sort of creative virtue and consequently differs from the syllogism.
>
> The difference must even be profound. We shall not, for example, find the key to the mystery in the frequent use of that rule according to which one and the same uniform operation applied to two equal numbers will give identical results.
>
> All these modes of reasoning, whether or not they be reducible to the syllogism properly so called, retain the analytic character, and just because of that are powerless. [See the text for the references.]

Since Poincaré was the greatest mathematician of his day, the claim that mathematical induction is not algorithmic and is synthetic is a respectable and defensible one. It is not insanity to defend it.

[* This paragraph is especially technical, and is intended for specialists – but, the general reader will also gain by reading it. What would prove that mathematical induction was a species of computation? Only this, an algorithm that churned out systematically every result possible by means of mathematical induction. But this is impossible. Peano arithmetic is categorical for number theory, but the principle of mathematical induction is expressed by it in second order logic and generates a totality that is non-denumerable; hence, no algorithm that generated every result provable by mathematical induction could be created, since any such algorithm would have to be first-order and generate a denumerable collection of results. Therefore, a solution by mathematical induction to the Halting Problem is not ruled out by any known result of mathematics. As detailed in the introduction and elsewhere in the main paper, the so-called impossibility proof for the Halting Problem does not refer to or delimit (a) mathematical induction, or (b) human reason. In fact, it is not even an impossibility proof that legislates against the deployment of algorithms in the solution to the Halting Problem. This is clarified in the text. Mathematical induction employs pattern recognition in the solution to a problem in number theory. It attaches the same pattern to every number in succession. An impossibility proof for the Halting Problem that ruled out the use of mathematical induction would have to demonstrate that the same pattern did not recur in the case of some particular number *n*. Such proofs do exist. For example, the proof that there can be no solution to the quintic equation attaches a group of permutations of roots of

the quintic that could not possibly be like that of the solution to the quartic. Thus, for $n = 5$ an inductive proof that all equations of any degree were solvable must break down. Such a principle has not been established in the case of the Halting Problem.]

In the early C20th philosophers, logicians and mathematicians set about systematically dismantling the logic of the late C19th, represented, for instance, in the work of F.H. Bradley. Late C19th logic was a logic of judgements, in other words, conceptual from the outset. We call it a *logic of intensions* – intension spelt with an 's' is a synonym of 'meaning' or 'concept'. The replacement logic is a *logic of extensions*, a logic that deals only with things not meanings, and replaces mental judgements by computable sentences. This has led to the elevation of "first-order logic", described by a "friend" of mine, who is a devotee of it, as "the voice of God" – which is curious, because I think he is a card-carrying atheist. First-order logicians have elevated an early developer of this language, Gottlob Frege, into a kind of saint of their movement, but it is clear that Frege himself had no such intentions, (spelt with a 't', meaning "purpose"), as he called his logic *concept-writing*. However, Bertrand Russell seized upon the work of Frege, and gave impetus to the movement. Around the same time, the academic community became obsessed with the idea of a "thinking-machine", a computer that would turn out all the results of mathematics. This was proven impossible by Gödel's theorem, but in the post-war period the obsession transferred to the desire to make a Golem – the machine that could imitate human verbal behaviour, and we are in the wake of that movement, coming to the end of its cycle, as the age-old dream of making such a machine is exposed as an illusion.

So, the history of the belief in the Golem is the product of a massive, collective mania that was accompanied by groupthink, which Wikipedia explains is "is a psychological phenomenon that occurs within a *group* of people in which the desire for harmony or conformity in the *group* results in an irrational or dysfunctional decision-making outcome." Generally, groupthink is supported by devices that amount to bullying, and exclusion: selection on principles of adherence to the ideology of the group. Criticism of the group ideas was excluded by excluding critics at all stages, particularly the undergraduate. It is to be noted that the chief critics to date of this groupthink are not philosophers or latterly logicians. They were Kurt Gödel, who was ostracised – an early logician saved from obscurity by the chance intervention of John von Neumann – John Lucas, a Merton college fellow, a Christian theologian, but who was never awarded the degree of doctor – and Roger Penrose, a mathematician. So, what happened to philosophy during the twentieth century? Why are there no academic rationalists, saving perhaps Plantinger? I point my reader in the direction of my chapter in my book *Kantian Consciousness*, entitled *The Wilful destruction of Kantian Consciousness* that I have placed onto my website.

To deal with the embarrassment of second-order logic, which quickly manifested properties that were not adaptable to the Golem project they came up with a number of devices, and these really did confuse this author for about thirty years. Of course, my own training in mathematics has followed a bizarre course without doubt, and I suppose that it is just possible that had I been familiar with mathematical induction when I was an undergraduate it might even then have occurred to me to use it, but I think not somehow. The problem is that mathematical induction can only be written fully as a second-order statement, and hence is irredeemably conceptual at outset. But the logicians came up with two or more devices for hoodwinking themselves into thinking that this did not really matter. They replaced the second-order "axiom" with a first-order statement, and said to themselves – this is all that it really amounts to. They hoodwinked themselves into thinking mathematical induction was a species of mechanical reasoning thereby. Indeed, they have even written computer programs that emulate to an extent part of the work of mathematical induction, and now mathematicians use those programs to search for results in number theory. The rot in mathematics is now very deep. Another device concerns a fine structure known as axiomatic set-theory. That set theory is really a form of concept writing has been ignored, but be that as it may, they created an illusion that mathematical induction is subsumed within set theory by an even more powerful device known as transfinite induction. The whole thing is a ghastly history of one illusion leading to another.

Numbers are not sets or anything like that – they are a primary category of understanding, and proof in number theory is by mathematical induction. The fusion and coherence of number theory with theories of the continuum (space and time) is a whole research of its own. I have delved deeply into it, but this is not the place for too much consideration

of that topic, saving one point about indestructibility of extensions that I have made because it relates to the wider aims of this apology – in this case, to justify my viewpoint as truth.

What if my solution to the Halting Problem is shown to fail?

Firstly, I assure my reader that I do think my solution the Halting Problem is valid. I think it is correct both as to principles and to details. I think that if some minor glitch in the proof is found, then it can be quickly patched up, and I also think that it is the duty of the academic world to help me with it in any way shape or form that it can.

But, secondly, everyone who goes into the ring to box for his life, must consider the possibility in advance of defeat. So, as a man, I also consider this. In the letter to the Anglican bishops that I quoted in full at the outset of this apology, I wrote:

> Even in the unlikely event that my solution to the Halting Problem being refuted, I do not doubt that the raising of the issue into proper debate, one in which the philosophers will have to justify their ideas to the community as a whole, and in a plain language that we all can understand, will have the effect of being a bolster to religion. I have no doubt that my work, taken as one whole, does firmly refute the pretentions of computer science as to what it claims to be able to demonstrate about human nature. Doubt will be restored, and with it the possibility of faith. Therefore, I ask for your blessing.

Here I wish to refer to what I call "my work, taken as one whole".

The beauty of any solution to the Halting Problem is that it represents a decisive victory in battle for rationalism against empiricism. If the solution goes through, that is an end to that kind of mania. I imagine that sometime in the future, probably the distant future, some clever person will resurrect the idea of a mechanical Golem, and find some clever way of doing so, but for the near future all those confident expectations of artificial intelligence will cease. Artificial intelligence will become what it ought always to have been – a seriously useful aid to human development – and we hope a salvation of the planet and its other species too.

But here I remark that my reader will find *multiple refutations of the Golem idea* in my work on my website – most of these, I confess, not extensively read to date. (But, it seems I am gaining a reputation in my own name as it happens. I am not sure what it is all about, but I keep being informed by that website that wants a life-subscription of money from me, that I am being referenced by many people. Ironic, since I am now Melampus.)

One important refutation of the Golem arises from my solution to the problem of the continuum. This is the idea of the indestructibility of extension that I referenced in a technical section above, the one that my "friend" wrote to say he derived from the book by Weyl. Here I will not go further into this, because I have already written a lot, but the basic idea is very simple indeed, and is explained on one sheet of paper on my website and can be understood by anyone – mathematical expertise is not required. (I do include some very technical stuff on that one side of paper, but this is partly to demonstrate to the scientific community that I am not some country bumpkin and have taken the trouble to rightly inform myself of the technical details. Also, it shows them instantly the falsity of their approach. They will understand it.)

There are many essays on my website, and some are philosophical, and some are technical. Among the philosophical papers there is that essay on the *Rationalist Philosophy of Mathematics*, which is a very good starting point, and, once again, one does not have to have mathematical expertise to understand much of it. There is also my *Answer to Hamming*, which also does not require mathematical expertise.

Hence, I stand by the statement I made to the Anglican Bishops – whatever happens, the one-sided obsession with the Golem will come to an end. It will become possible to challenge this, and if I am right about the solution to the

Halting Problem, the obsession with it will be killed instantly. It just won't be possible to defend it in public. In the wake of that a whole lot of other nonsense will end – such as the obsession with obtaining physical immortality by manipulation of genes that reverse the natural ageing process, or, most absurd of all, by "downloading consciousness into cyberspace."

But before I close my apology, I should like to tell the reader about how it all began.

Thirty-eight years ago – and you reap as you sow

It was a little more than thirty-eight years ago that this story began. I describe this in the Preface to my book, *Kantian Consciousness*.

> At the age of eighteen I decided to study some formal logic. I acquired a copy of E.J. Lemmon's *Beginning Logic* and worked my way through the exercises. "This is all very well," I said to myself, "and even interesting, but this is the way a mechanical device works. The human mind does not think like this." In such a step, I was unaware that I had gone beyond any current "master" of logic, for formal logic, more specifically first-order formal logic, is the very language of contemporary philosophy. That the mind could think without a mechanical substratum, or that there are inferences not codable in just this language of machines, has not even occurred to them.

Thus, I was already an outcast and a rebel before I'd even set foot in the University. It was one of those "good" ones; one where you feel really proud of yourself to have passed their special exam and "got in"; one of those divided into thirty or more colleges, each of which boasts to be a university within a university. I found pretty well from the start that I could not make head or tail of their non-existent course – there never was a syllabus, and essays were set on recondite topics, many having something to do with an obscure branch of philosophy called "philosophy of language". The lectures were hopeless, but I attended at least some of them out of courtesy. The teachers, teaching being done mainly on an individual basis, were on the whole kind. I never got to write an essay on any topic of my own choice – except *once* – I chose to write an essay *against materialism*. My teacher, who was also my Director of Studies, said almost nothing about the essay. Generally, the methodology was to pick over the student's essay word-by-word, which usually had the effect of making one feel very stupid and like an intellectual pigmy. This time nothing much was said, but maybe sometime after that, he said to me, "You know, you're not very good". That surprised me a little, because I am aware that one *never* says that to a pupil, and furthermore all my other teachers did say quite contrary things, and though I never got the top score in the year, I was always near the top, did very well in exams, and was even awarded prizes for it every year.

In my third year, realising that the hard core of the subject lay in logic, I took the papers in mathematical logic, philosophy of mathematics, metaphysics and philosophical logic. Then, there was a dissertation to write. At that time, I encountered Turing machines and the Halting Problem, and became convinced that as a Turing Machine is absolutely nothing but a deterministic system, a machine whose outputs are fixed by the given inputs, so indeed, the Halting Problem must be soluble. I said as much to my teacher. He actively discouraged me from writing the paper, and I will confess that my first two drafts were pretty poor stuff, and whether he saw the final essay I do not know. What I do know is that in the final stages, I realised that the problem I was tackling was too big for me at the time, and I revised the paper and took another approach, looking rather at the "softer" ways in which human thinking might differ from that of a machine.

This dissertation contained ideas that later appeared in the essay of big wigs, such as Roger Penrose.

> Let us define a "monster" to be an argument which, though intuitively valid, cannot be formalised. … In producing a monster we must discover an argument with both resists formalisation and is

> certainly valid. … [We] must produce something that falls in the domain of science and especially mathematics. … it will not be easy to establish a mathematical monster. [Written by Melampus, 1981.]

The idea of a "monster" is an allusion to the work of Imré Lakatos in *Proofs and Refutations*. The reader will appreciate that this idea of a monster is the basis of the work of both John Lucas and Roger Penrose, both of whom seek to use Gödel's theorem as a monster. At the time of writing the dissertation I was unaware of the work of Lucas, which was a pity. Strange that my Director of Studies did not point me in that direction. All the stranger, because I also met John Lucas at the time. I was asked to manage the philosophical debate club and asked to invite J.R. Lucas to speak at the club, which I did. I still have the card that I designed with the program, complete with a rather crude drawing of mine on the cover of a plump naked woman plucking an apple from a tree. There was a rumour around the faculty that it was a depiction of Gertrude Anscombe, an imputation that I thoroughly reject. Anyways, Lucas was personally known to me at that time, and yet I was unaware of his papers on Gödel's theorem. Perhaps I was dimly aware that he was supposed to have made a monumental fool of himself in some way, but exactly how he had done this, I did not know. But both Penrose and Lucas attempt to use Gödel's theorem as a monster – as I do myself. It was the study of Gödel's theorem and the linking of that to real-number generators by means of Cohen forcing (excuse the momentary technical aside) that lead me to my solution to the problem of the continuum, the reason being that the continuum is a model of Gödel's theorem, and the inductive procedure for generating extensions to first-order logics defines a transcendental number. Gödel's theorem corresponds to a transcendental number.

For a taste of Penrose's work, consider:

> Mathematical truth is *not* something that we ascertain merely by use of an algorithm. I believe, also, that our *consciousness* is a crucial ingredient in our comprehension of mathematical truth. We must 'see' the truth of a mathematical argument to be convinced of its validity. This 'seeing' is the very essence of consciousness. It must be present *whenever* we directly perceive mathematical truth. When we convince ourselves of the validity of Gödel's theorem we not only 'see' it, but by so doing we reveal the very non-algorithmic nature of the 'seeing' process itself. [Roger Penrose, *The Emperor's New Mind,* chapter 10, published 1990.]

Another detailed suggestion in my dissertation concerned *pattern recognition*.

> Solving the halting problem for a particular machine lies in recognising the pattern into which the machine must enter if it never halts. … there is no reason to suppose that we cannot come to recognise these patterns. [Melampus, *Dissertation*, 1981.]

The use of reasoning from "visual description" is a theme of the research conducted by Owen Kellett under the supervision of Soren Bringsjord at the RPI. For instance, the group paper [*A New Millennium Attack on the Busy Beaver Problem*, RPI, 2003] mentions "a fertile testbed for exploring both "visual" reasoning and possible super-Turing computation." Kellett's paper makes frequent use of the idea, for instance, "It should be abundantly clear from the diagram and this verbal description of the machine's behaviour that the pattern will continue indefinitely…" [Owen Kellett, *A multifaced attack …* , p. 83, July 2005.]

I would like to quote again from my dissertation on this point.

> We look upon the machine as doing something, and may even ascribe to it an intentional activity: "It is *searching* for a 1 and will only halt if it *finds* one," from which awareness of the halting conditions is immediate. (i. Teleological reasoning.) The activity of searching for a 1 goes on for ever if no 1 exists, yet we are not obliged to continually go through the notion of moving one square to the right and checking that in order to conclude that the machine will not halt. (ii. Analog methods.) We can recognise inherently infinite processes, and in so doing understand them and see them as taking values. (iii. The capacity to handle the infinite.) Thus, coming to understand under what circumstances which, if repeated twice, is repeated always. (iv. Recognition of patterns.) And

> underlying all of this is our ability to see signs and symbols as meaningful. (v. The semiotic function of mind.)

The reader will appreciate that in the first sentence I no more ascribe intentional behaviour to a Turing machine than does Roger Penrose, when he writes, "It has been helpful … to use the anthropomorphic term 'know' in reference to an algorithm. However, is it not *we* who are doing the 'knowing', while the algorithm just follows the rules we have told it to follow?" [*Emperor's New Mind, Chapter II.*]

My final papers were taken over two days, Mathematical Logic and Metaphysics on day 1, Philosophy of Mathematics and Philosophical Logic on day 2. I knew the first day went well, but between the two days an insufferable gloom descended upon me (I am not prone to depression generally) and I felt the task before me was hopeless as I took the papers on the second day. The whole experience of studying that course (if it was a course) was immensely depressing. Therefore, I was not too surprised when the results came out to learn that I had a 2.i, and not a 1st. Then I had this note from my Director of Studies. He expressed his sympathy that I had narrowly missed out on a First, despite obtaining "a very high" 2.i, and wishes me all the best for the coming year [dated 23 June, 1981], while praising me for the respect I earned from Faculty members.

But I spoke to my Director of Studies about taking a PhD, and he declined, offering me a place on the M.Litt course instead for one year. (PhD students take the M.Phil.) I knew this was a path leading nowhere, and was sick of the place, and of writing essays using a language that was not natural, and it being at the onset of the Thatcher recession decided that my best course lay in entering the world of work. I had some concerns about my family too. My father was in difficulties, and my older brother's mental health was suspect. (He subsequently died at an early age from a Aids related illness. It was as well that I was in work by that time.) But one week later, I received a letter from my teacher in Philosophy of Mathematics, a Dr. L. He told me all about the examination meeting. How I had been awarded firsts in Metaphysics and Mathematical Logic, high 2.1s in Philosophy of Logic and Philosophy of Mathematics, and a split result on my dissertation – a First from one examiner and a low 2.i from the other. They debated my papers at length, re-read them, and finally the dissertation was read by a third examiner, and the person who awarded me a First on it agreed to downgrade it to a 2.1. So I missed out on a First by that downgrade Dr. L. concludes, "P.S. In all this, I think your performance represents a remarkable personal achievement, and I'm very proud of the work you have done."

Frankly, I would say that in all respects I've had a "hard life" – there was the mental illness of my brother, for instance, the early collapse of my father's final business, all the bullying I came on the receiving end of in my first employment, and so on. But I would not wish to give the reader any impression that I am either bitter or self-pitying or anything like that. It's been a great life – extremely interesting, full of deep experiences, and I think I would have hated the academic environment anyways. After leaving that college, I often rubbed shoulders with academics, but frankly I find their whole lifestyle sterile and I would not wish to belong to it. I'll stay on the margins of that culture now, whatever happens with my thesis.

It is very odd that I should receive a blow-by-blow account of the marking of my papers from someone who was *not* my Director of Studies, and nothing about my actual results in those papers from that self-same Director of Studies.

And I do think it was a shame that my thesis about Turing machines and the nature of human reason was never developed at the time. However, what goes around comes around they say, and they also say, "you reap as you sow". Thus, even the events in that examination room were recorded for posterity. My ideas subsequently came out in the work of Penrose, Kellett and Bringsjord, and here the reader can make up his or her own mind about the value of those ideas. I've explained clearly what the implications for the applecart are should my solution to the Halting Problem go through. They are the same implications that existed thirty-eight years ago. And in all events, this is surely the end to the unopposed march of the Golem idea. The dialectic must surely shift back towards a just balance, and we must surely be able to at least *consider that a machine is not human*.

"You reap as you sow." It seems to me that my letters to "the academics" and the announcement of my claim should not go ignored. Until now, only one recipient or institution has deemed it worthy to send in what is frankly only a

holding letter. The relevance is now acknowledged, and I trust that the evaluation will now not be too long delayed. But I feel that a word of warning is due. A professor receives a letter, and, let us suppose, that the contents of that letter are in some way disagreeable to him. He is advised by the sender that he expects a reply, however short. We call in Eastern philosophy the outcome of our actions their "karmic consequences". There are karmic consequences to all our actions – if we ignore a person or a duty there are consequences to that action, just as much as if we rose to the challenge.

Silence is not a correct way to deal with this problem.

The awfulness of a Golem has entered into popular culture through the myth of Frankenstein's monster. I am beginning to wonder what the real Golem is.

Conclusion and appeal

I am hoping by now that my reader has got interested in my personal drama and will want to know its outcome. The reader will surely be interested to know whether the solution to the Halting Problem is accepted as valid or not.

The website is www.melampus.name, where the reader may follow the story by signing up for the newsletter.

If you wish to be more actively involved, then please use the contact form at the website. A personal email address shall be attached to every instance of the newsletter as well. This is part of what I wrote to Prof. Bringsjord, a letter that was not acknowledged, and I thereby have no idea whether he received it or not. But it was sent to his email address.

> I also belong to the category of men who are possessed of modesty coupled to clarity, but utterly determined in the pursuit of a cause of singular importance to humanity to do whatever it takes to bring a truth to light. I am willing to be refuted in private, but if not, then in public.

I conclude with a quotation from Tennyson (*Ulysses*):

> To strive, to seek, to find, and not to yield.

Postscript – about privacy

My email address is: melampus@blacksacademy.net; you are invited also to subscribe to the newsletter. By copyright law you own the text of anything you write, but I could legally quote selectively from your letter or summarise. I would ask your consent before quoting. (The reader will note a couple of selected quotations from historic letters here, but I am sure will allow that as no infringement of this general privacy policy.) If you are expressly against being quoted or summarised at all, you can always mark the letter, "Private and not for publication". Alternatively, if you would like your views to be included in the newsletter, then write, "For publication". Whether I publish a view thus marked or not is at my discretion, but I am sure to be glad of any constructive comment. Unfortunately, a man who upsets the applecart may expect some hostility as well; hence, in part, the use of screening devices. Please do write to me, whatever your views. Criticism is most welcome, and support even more so. MELAMPUS.

Appendix – recipients of the letter to "the academics"

Peter Molnar, Amazon
Alvin Foo, Reprise Digital
IBM Thomas J Watson Research Centre
Hanson Robotics
Prof. Pedro Dominigos, University of Washington

DeepMind

Mustafa Suleyman
Dr. Dennis Hassabis
Shane Legg

Imperial College

Dr. Francesco Berlardinelli
Dr. Krysia Broda
Dr. Mahdi Cheraghchi
Dr. Antonine Cully
Prof. Abbas Edelat
Dr. Aldo Faisal
Prof. Philippa Gardner
Prof. Ian Hodkinson
Prof. Stephen Muggleton
Prof. Murray Shanahan
Dr. Steffen van Bakel

Oxford

Prof. Christopher Strachey
Prof. Richard Bird
Dr. Stephen Cameron
Prof. Luca Cardelli
Prof. Bernardo Ceunca Grau
Prof. Yarin Gal
Prof. Jeremy Gibbons
Prof. Leslie Ann Goldberg
Sir Tony Hoare
Prof. Ian Horrocks
Prof. Peter Jeavons
Prof. Marina Jirotka
Prof. Elias Koutsoupias
Prof. Thomas Lukasiewicz
Prof. Stephen Pulman
Prof. Michael Woodridge

Oxford Mathematics Institute

Department of Mathematical Logic
Prof. Jochen Koenigsmann
Prof. Jonathan Pila
Dr. Jamshid Derkhshan
Prof. Hilary Priestley
Prof. Boris Zilber
Prof. Samuel Cohen

Google

Martin Abadi
Douglas Aberdeen
Eirikur Agustsson
Jay Adams
Jyrki Alakuijala
Stanislav Angelov
Dave Bacon
Yasaman Bahri
Yunfei Bai
Rami Berends
Stephanie C.Y. Chan
Richard Chiburis
David Dohan
Neil Fraser
Sergio Guadarrama
Jasmin Hsu

Turing Test (Loebner Prize)

Rob Wortham
Dr. Bertie Müller
Dr. Yasemin J Erden
Dr. Andrew Martin
Dr. Ed Keedwell
Dr. Nir Oren
Dr. Floriana Grasso

There are so many people working in AI that it is difficult to know who to select. The main idea was to choose two excellent research centres closest geographically. In addition, representatives of who appear to be leading industries, with Google chosen in particular. At Google there are approximately 1,500 researchers in AI. None of the people listed here are known to me personally.

MELAMPUS