God and Gödel:

Gödelian Incompleteness in Mathematics

and the Confirmation of Theism

James Baird

In 1931, Kurt Gödel published his monumental findings on undecidable formulas in formal systems of mathematics.1 His incompleteness theorems demonstrated the inability of any strictly formal system of calculation to prove every true mathematical formula. Gödel himself took them to mean that no formal system could capture the full range of human mathematical insight.2 This result changed the discipline of mathematical logic forever.

In what follows, I want to show that Gödel's results also change the way that humans must look at themselves and the world in which they live. Gödel's incompleteness theorems make the philosophical position of naturalism untenable because they imply that human rationality is forever out of reach of complete scientific explanation. Because of this result, Gödel's theorems aid Richard Swinburne's rigorous Bayesian confirmation of theism. They remove an important objection to Swinburne's approach, and they also make available the existence of human rationality itself as another piece of evidence for the hypothesis that God exists and is the creator and sustainer of the world.

The Significance of Gödel's Achievement

Gödel's proof was aimed at answering the question of whether or not it would be possible to prove all the truths of mathematics by completely formal means, that is, simply by following rules about how to manipulate marks on paper, without the need for understanding, insight or intelligence. At the time, confidence in logic was nearing its zenith, and many took it for granted that mathematics could be completely formalized. Gödel's solution was stunning, because it used the simplest part of mathematics, arithmetic, to show that mathematics could not be formalized in this way. In effect, Gödel showed that if a formal system for logic included the ability to do addition and multiplication, and if it was constructed so that it was consistent, i. e. did not prove contradictions3, then a true sentence, call it the Gödel sentence, could be constructed following the rules of that system which could not be proved by that system.4 His proof was completely general, so that even if the system were modified so that it could prove the Gödel sentences which baffled it before, unless the modifications made the system incapable of arithmetic or prone to contradiction, the modifications themselves could be taken into account to generate new Gödel sentences which would be unprovable by the new system. In this way, Gödel dashed all hope for a completely formal mathematics.

This finding was dramatic in its effect on the mathematical community, but it seems to be of limited interest otherwise. But almost immediately, further research showed that Gödel's results have radical implications in other areas as well. In 1936 and 1937, a British mathematician named Alan Turing, described abstract machines which lie at the foundations of all modern computers.5 These machines, now called Turing machines, carry out calculations in extremely simple ways, but given sufficient time and resources, they are apparently capable of carrying out any procedure which can be carried out by strictly formal means. Indeed, Turing showed that any formal system can be translated into a Turing machine with equivalent output, and any Turing machine can be translated into a formal system. An implication of this is that anything which can be done by any modern computer can also be done by a Turing machine, (though for any interesting calculation, the Turing machine is likely to take trillions of years).

Because of the intertranslatability of formal systems and Turing machines, it was to be expected that Gödel's incompleteness would arise for Turing machines. Turing showed that and more in his original works.6 The consequence is that modern computers, and all imaginable extensions of modern computers, face the same incompleteness which Gödel demonstrated in 1931. No computer can generate all mathematical truths. It is sometimes thought that computers which used massively parallel architecture, or heuristic programming, or which can learn from past failures (even past failures to prove Gödel sentences) might be able to escape the bounds of Gödelian incompleteness. But as long as such improvements are mechanical (as long as computer wizards aren't actual wizards), and as long as they do not cause the computer to start proving contradictions or degrade the computer's ability to do arithmetic, then the improved computer will be describable as a consistent, arithmetically competent formal system which is subject to Gödel.7 A computer is subject to Gödel's results even if it incorporates random elements, whether the actual randomness of quantum effects or the pseudo-randomness of some chaotic system, as long the randomness does not lead to inconsistency or arithmetical breakdown.8 Thus, no imaginable advance in computer science will give computers the power to overcome Gödelian incompleteness. There will always be mathematical truths beyond the reach of any particular computer.

Naturalism's problem with Gödel

Given the common assumption that any lawlike natural process can be computationally simulated to any degree of accuracy, if only we have enough time and resources, it has occurred to several thinkers that Gödelian incompleteness may pose a serious threat to naturalism by putting human thinking beyond the power of any scientific theory to explain. The Oxford logician, J. R. Lucas published an article to this effect in 1961, claiming that Gödelian incompleteness implied the incompleteness of any mechanistic model of human thinking.9 He was met with a storm of (often mutually destructive) refutations in succeeding years, but answered them adequately enough to republish his argument as the key element of his 1970 book, The Freedom of the Will.10 Lucas's work helped stimulate Douglas Hofstadter, a dedicated believer in the computational model of the mind, to publish the 1979 best seller and Pulitzer Prize winner, Gödel, Escher, Bach.11 In this book, Hofstadter developed a particular objection to Lucas's argument which will be reviewed later. In 1985, the Harvard philosopher, Hilary Putnam, used Gödel's theorems to show that human rationality could not be prescriptively formalized.12 More recently, the Oxford physicist, Roger Penrose, has published The Emperor's New Mind13 in 1989, and Shadows of the Mind14 in 1994, expounding his argument that Gödel's theorems show that human thinking cannot be entirely the result of mechanical, computable processes. Indeed, Penrose goes to great lengths to show that human thinking cannot be the result of any known physical processes. Penrose has been subjected to his own storm of objections15 mostly the same ones which had been aimed at (and refuted by) Lucas earlier. Whatever else may be concluded about them, the reactions to Lucas and Penrose indicate that they are striking at something that is dear to many.

To see the nature of the problem Gödel poses for naturalism, let's engage in a little science fantasy. Imagine the grandest possible scientific research program, the Human Genome Project. The aim of this project is nothing less than the mapping of human intellect in its entirety. Imagine that after years or centuries of work, the project is completed, and humans beings are furnished for the first time with the completed Human Genome Map. Using the Human Genome Map, psychologists can trace the origin of any thought, at least in principle. Every belief, hope and perception of which humans are capable finds it explanation in the laws and principles of the Human Genome Map. Mental illness can be treated with pinpoint precision. Mental energy can be channeled at maximum efficiency. The dark sources of superstition, war and hatred can be brought into the light and eliminated. The future of mankind seems forever bright.

But there is a catch for those who remember the work of the 20th century logician, Kurt Gödel. The Human Genome Map can be translated into a formal system which should be able to prove every mathematical truth which the human intellect is capable of recognizing. Since human beings are capable of addition and multiplication, we would expect this formal system to be able to prove the truths of addition and multiplication. So the system has one of the characteristics, arithmetical competence, which is needed to make it subject to Gödel.

Would the system prove contradictions? Humans make errors and contradict themselves all the time, so we might assume that the system would be inconsistent in the same way. Of course, the system might well incorporate various guessing mechanisms and heuristic elements which may lead it into error sometimes, but what about when it is functioning at its best? Imagine that it is functioning under ideal conditions in which there are no constraints of time or resources. If it generates contradictions even under such ideal conditions, then the entire project will quickly collapse, because a system which proves a contradiction proves everything.16 Of course, if every sentence can be proved, then no sentence is any better than any other. If everything is proved, then nothing is. Any system which generates such a result becomes absolutely useless.17 To say that the Human Genome Map is such a system and that it truly reflects our best rational capabilities would be to say that all of our beliefs are useless, including our belief in the Human Genome Map. So we had better say that the system derived from the Human Genome Map would be a consistent system when operating at its best under ideal conditions. Therefore, the formal system derived from the Human Genome Map would be a consistent, arithmetically capable system, and so would be subject to Gödel's incompleteness results.

But this leads to a proof that the Human Genome Map does not completely describe our thinking. For using Gödel's techniques, we can derive a Gödel sentence for the Human Genome Map system. This sentence will be true, of course, and we will see it to be so, but it will not be provable by the Human Genome Map system. In recognizing the truth of the Gödel sentence, we will have gone beyond what any reasoning completely described by the Human Genome Map should have been able to do. Of course, it would be possible to work on the Human Genome Map system until it could prove the old Gödel sentence. But the improvements would allow us to generate a new Gödel sentence, which could not be proved in the system of the new Human Genome Map. This process could continue indefinitely, but Gödel's proof guarantees that the human recognition of mathematical truths will always be slightly outside of any Human Genome Map which can be created.

Obviously, this little thought-experiment hands metaphysical naturalism a stunning defeat. Metaphysical naturalism is the view that everything can in principle be explained as the result of purely natural processes guided only by natural laws. But if the Human Genome Project cannot succeed, then human thinking cannot ever be completely explained in this way. There is something to human thought which always escapes naturalism's net.

Naturalism's Defenders

Naturalism is far too robust a position just to give up in the face of Gödel. Naturalists have an entire list of objections against any deployment of Gödel in this way. Most of these, though still popular, have been answered adequately both by Lucas and by Penrose. But there are some objections to the Gödelian attack which show possible escape routes for naturalism. Each however creates interesting problems of its own.

Penrose: The Quantum Brain Defense

The first of these would be from Penrose himself. Penrose goes to great length to show that human thinking cannot be the result of any known physical processes because he is building a case for a new physical process. Penrose believes that a union of quantum analysis and Einstein's relativity theory would come down on the side of quantum analysis, with the result being a theory of Quantum Gravity. Penrose hopes that Quantum Gravity might, in certain special circumstances, lead to processes which are truly non-computable. By using these processes, Penrose believes that the human brain might also be non-computable and so transcend the limits of Gödelian incompleteness. In the Emperor's New Mind, Penrose argues that such processes might be made use of in systems of neurons,18 but in Shadow's of the Mind, he suspects that such processes might actually be used by the brain at the far smaller level of the microtubules of neurons.19 In this way, by suitably expanding physics, the brain might be an entirely physical object and still transcend the limits of Gödel. This could be viewed as a salvation for a suitably expanded naturalism.

Our brief response to Penrose would have to be: too early to tell, but unlikely. To be non-computable is to be non-lawlike. We hate the idea of physical lawlessness, and resist it at every opportunity. Quantum analysis is the only example of scientific acceptance of fundamentally lawless behavior on the part of nature, and nobody is too thrilled about it. Einstein was notoriously dissatisfied with the lawlessness of quantum analysis until he died, and very few physicists even today view the lawlessness as real. The standard interpretation of quantum analysis, the so-called Copenhagen interpretation, tends to see quantum lawlessness as a limitation on our knowledge rather than a real fact about the world. For Penrose's proposal to succeed, a realist interpretation of quantum events is not only essential, but must be expanded to include gravity.

Current quantum analysis is tolerated, in spite of its lawlessness, because there is no rival theory which can match its enormous predictive ability. But here Penrose runs into trouble. The only grounds for believing in the yet-to-be-discovered quantum gravity is that it would allow human rationality to be physical and still escape Gödelian limits. If this is the only area in which quantum gravity is relevant, then it is unlikely to enjoy the kind of predictive success which coaxes physicists to a grudging acceptance of more traditional quantum analysis.

For these reasons, naturalism does not receive much comfort from Penrose. He correctly diagnoses the dangers which Gödel implies for traditional naturalism, but as a defense of naturalism he offers only the possibility of a yet-to-be-discovered theory of quantum gravity, leading to yet-to-be-discovered non-computable processes, which can be used by yet-to-be-discovered mechanisms in the microtubules of the brain. There is no particular evidence for any of these, so they seem justified only as an effort to save naturalism in the face of Gödel's theorems. Naturalism would need far stronger support than it has to warrant such a leap. Small wonder that naturalists have been Penrose's most aggressive critics.

Dennett: The Human Fallibility Defense

The most outspoken of these critics has been the philosopher Daniel Dennett.20 Dennett proposes that naturalism, in its computational guise of strong Artificial Intelligence, can defend itself against Gödelian limits by appealing to the fallibility of human reasoning. He admits that Gödel's Theorem tells us that no formal system or equivalent algorithm (computer) can prove all the mathematical truths humans can recognize. But, he says:

Gödel's Theorem in particular has nothing at all to tell us about whether there might be algorithms that could do an impressive job of "producing as true" or "detecting as true or false" candidate sentences of arithmetic. If human mathematicians can do an impressive job of just seeing with mathematical intuition that certain propositions are true, perhaps a computer can imitate this talent, the same way it can imitate chess-playing or conversation holding: imperfectly, but impressively. That is exactly what people in AI believe: that there are risky, heuristic algorithms for human intelligence in general, _.21

Dennett's argument is that humans guess and take short cuts in order to survive, and that their rationality thus falls short of the kind of rock solid proof found in the systems Gödel deals with.

This is an interesting suggestion, but it falls to pieces when we begin to ask a few questions. First of all, how extensive is the rational fallibility Dennett has in mind? Dennett clearly suspects, though cautiously, that we are up to the task of pushing back the boundaries of knowledge until we explain all aspects of our world in purely scientific terms.22 But if our intellects are too fallible, we simply may not have what it takes to reach his dream. Perhaps Dennett means that we are fallible, but also self-correcting, so that we can keep working by trial and error until we get things right. But in this case, Gödel becomes relevant once more, for as Lucas pointed out 25 years ago, "A fallible but self-correcting system would still be subject to Gödel's results."23 So Dennett must be hoping that we have a very special kind of fallibility. It must be severe enough that we will never under any circumstances overcome it completely, but it must not be so severe that it finally prevents us from completing our scientific explanation of the world. This is a possible hope, I suppose, but it cannot be said that it has a lot going for it other than a faith in naturalism itself. Just to cite one problem, it is hard to imagine what sort of evolutionary pressures could produce just this almost unnoticeable sort of fallibility. How could such a retiring problem have any reproductive significance?

Things are even worse for Dennett if we imagine his hoped for completion of science. Suppose for argument's sake that we have the very well-behaved kind of fallibility which Dennett's position calls for. Once again imagine the completion of the Human Genome Map. Certainly by this time, if not long before, the sources of human fallibility will have all come to light. Will we, or will we not, be able to eliminate our fallibility at this point, and gain the ability in principle (thought perhaps not the time or resources in practice) to reason correctly? Either option is catastrophic for naturalism. If we can eliminate our errors, even in principle, then the formal system derived from the Human Genome Map will be a consistent, arithmetically competent system and so subject to Gödel. In this case, we will discover that we are something other than the Human Genome Map or any improvement thereof, and naturalism fails.

On the other hand, if even with the help of Human Genome Map we are unable to eliminate our errors, then we must be unavoidably inconsistent reasoners. In this case, as we have seen earlier, we will have an unavoidable proof of every sentence we can say or think. Everything will be true for us, and so nothing will be, including naturalism.

Faced with these looming catastrophes, Dennett seems forced into the claim that our own understanding of the true Human Genome Map will never be clear enough to carry out the Gödelian refutation of it. This indeed seems to be the escape he was exploring at one time.24 It is certainly the position of his friend, Douglas Hofstadter.

Hofstadter: The Necessary Ignorance Defense

Hofstadter has given a new twist to a long-established defense against Gödelian refutations of naturalism, put forward most effectively by Paul Benacerraf in his dispute with J. R. Lucas. Benacerraf made several attacks on Lucas's version of the Gödelian argument, and he concluded with the claim that at best, Gödel's theorems proves if I am a Turing-machine "I cannot ascertain which one."25. This argument hinges on the fact that we can only see the truth of Gödelian formulas of formal systems we can understand. This means that we can continue to believe that our minds operate entirely according to some algorithm as long as we can come up with some plausible reason why we can never discover enough about the algorithm to be able to derive and understand its equivalent formal system and see the truth of the relevant Gödelian formula. Benacerraf suggested that the complexity of our own algorithm might be sufficient, and Hofstadter has elaborated on this by suggesting that algorithms can be of any finite amount of complexity, so that sooner or later any particular human will be so overwhelmed by complexity as to be unable to apply Gödel's procedure.26

The trouble with this defense in either version is a failure to take the power of Gödel's proof seriously. Gödel proved, not just that it takes too long or costs too much or is practically or physically impossible for a plausible formal system to prove its own Gödel sentence. He proved that it is logically impossible. Therefore, if our minds are the embodiments of sound formal systems, it is logically impossible for us to see the truth of our own Gödelian formula. The complexity of the system or the length of the proof is entirely irrelevant to this result. Even if finding some Gödel incompleteness result would take billions upon billions of years, and consume more paper and ink than could be supplied by a universe filled with nothing but ball-point pens and legal pads, it would still be logically possible to carry it out. It is therefore useless for Benacerraf or Hofstadter to appeal to mere complexity or length to avoid the refutation Gödel hands naturalism. The only way to escape the implications of Gödel is if it is logically impossible for us to know our mechanism well enough to perform the Gödel operation on it. Our ignorance must be logically necessary.27

I suppose it is possible for naturalists to take this line if they are prepared to admit that our minds are the results of real but necessarily mysterious processes. But this last ditch maneuver generates problems of its own, problems which have a very familiar sound. One of the great appeals of naturalism is its promise of a non-mysterious universe. It is ironic that in the face of the argument from Gödel, naturalism cannot escape refutation except by hoping that there is a logically necessary mystery right in the middle of the human head.

Worse still, versions of almost all the standard objections to dualism now arise within naturalism. How do necessarily mysterious processes interact with non-mysterious processes? Is the interaction itself necessarily mysterious or not? If not, then at what point does the mystery set in? But if so, what prevents our necessary ignorance of mental processes from spreading to all causally related processes, and from those on out, until we are necessarily ignorant of everything? Can we have a non-mysterious theory of the process by which the capacity to carry out necessarily mysterious processes develops in the human fetus? Will the theory of the evolution of the brain be mysterious or non-mysterious? What mathematically definable mutations and selection pressures could be imagined to bring necessarily mysterious processes into being?

In this light we can see that, at best, naturalism has almost all the disadvantages of dualism, as well as its own notorious difficulties with morality, meaning, freedom, values and so forth, there seems to be nothing left to attract us to it. Because of Gödel, naturalism is in ruins.

Gödel and the Bayesian Confirmation of Theism

The defeat which Gödel's theorem hands naturalism can be used to aid theism in several ways. In what follows, I will show that the argument from Gödel can be of use in the rigorous program of confirmation of theism articulated by Richard Swinburne.

Swinburne's Bayesean argument

Swinburne uses Bayes theorem and other elements of the calculus of probability to build a cumulative case for theism.28 His argument is that theism can be supported by using a variety of inductive arguments to build up a cumulative case for it, in much the same way that we would for our more important large-scale scientific hypotheses. Using Bayesian techniques, each bit of evidence can be judged to determine the approximate amount of confirmation or disconfirmation it lends to the theistic hypothesis. In his book, The Existence of God, Swinburne builds a case using versions of the cosmological and design arguments, as well as arguments from consciousness and the apparent providence of God, supplemented with a rather different kind of argument based on the testimony of experience of God by many witnesses. This case, he claims, makes the theistic hypothesis more probable than not.

The argument from Gödel aids Swinburne's case in two ways. It helps establish his claim that personal explanation is a separate, irreducible category of explanation. This provides Swinburne a way to respond to an important criticism leveled by John Mackie. In addition, the argument from Gödel provides another piece of evidence which counts in favor of theism, namely the scientifically inexplicable existence of human rationality.

The Gödelian defense of Personal Explanation

One of the crucial pieces in Swinburne's case for the probability of theism is the claim that personal explanation cannot be reduced to scientific explanation. Swinburne argues this on conceptual grounds, showing that every attempt to reduce personal explanation to scientific explanation involves diminishing or altering the concept. Then he points out that we know perfectly well how to use the concept, and so it can be accorded an independent status. This is important for Swinburne because theism relies on a special kind of personal explanation, namely the action of God. For his Bayesian analysis to succeed, Swinburne has to argue that the theistic form of personal explanation is not terribly unlikely. In this way, the irreducibility of personal explanation is crucial for the success of Swinburne's case.

The late John Mackie attacks Swinburne on this point in The Miracle of Theism. Mackie claims that the prior probability of the theistic hypothesis is fatally lowered by the theistic appeal to directly fulfilled intentions.29 Since all of our experience is of persons whose intentions are fulfilled only indirectly through physical bodies operating according to natural laws, the hypothesis that there exists a being whose intentions are fulfilled independently of such means must have an enormously low prior probability.

Swinburne has responded directly to Mackie on this point. His primary complaint is that Mackie has failed to give due attention to his intention of judging the prior probability of the theistic hypothesis on the basis of tautological background knowledge alone.30 Swinburne claims the right to proceed in this way because the distinction between background knowledge and evidence is largely a matter of choice and he wants to use the existence of the universe as one bit of evidence in his cumulative case. This leaves nothing but logical truth as background knowledge. But, Swinburne urges, when we try, on the basis of tautological background knowledge alone, to judge the prior probability of an all-inclusive hypothesis like theism, the primary factor is its simplicity. The simplicity of a hypothesis will have some inverse relation to the number of entities and kinds of entities the hypothesis postulates, and the complexity of the properties the hypothesis attributes to the entities it postulates.31 The crucial point of this in answer to Mackie is that the simplicity of a hypothesis can be evaluated apart from its familiarity or fit with our ordinary experiences and expectations. Therefore, it does not matter so much whether we are familiar with the direct fulfillment of intentions as long as the concept is a simple one. Swinburne's claim is that it is simple, and that this is made obvious by the fact that we learn to use this concept in our own case long before we become aware of the complexities of the processes which actually connect our intentions with their fulfillments.32

I certainly think Swinburne's response to Mackie is defensible, but the argument from Gödel means that even if we, incorrectly, accept Mackie's claim that the prior probability of the theistic hypothesis must be assessed on the basis of our ordinary expectations, we can still resist his criticism. For Mackie's doubt about direct fulfillment of intentions is clearly a corollary of his confidence that there can be, in principle at least, a complete scientific explanation of human thought processes, and the argument I have developed shows that this confidence is misplaced. Mackie says "any personal explanations that we can actually give, as applied to ordinary actions, constitute, when properly spelled out, a sub-class of causal explanations, not a rival mode of explanations to the causal one."33 Later in the same paragraph, he adds:

Teleological description may be distinct from anything involving causation; but teleological explanation of anything's coming about is, in all ordinary cases, only a special kind of explanation in terms of efficient causes. For example, to explain an action as purposive is to indicate that it is causally brought about by the agent's desires, beliefs and decisions. If we say that a plant or an animal has such and such organs, or behaves in a certain way, because this serves some function or tends to produce some result, this is shorthand for a causal account of how these features have been developed by natural selection.34

By 'causal explanation,' Mackie means an explanation in terms of laws and initial conditions.35 And Mackie has argued elsewhere what he here assumes, that teleological causation, of which causation by human action is one type, can be seen as a sub-type of this single type of causation by laws and initial conditions.36 Explanations by laws and initial conditions are what I have been calling scientific explanations. Therefore, it is clear that Mackie's objection depends upon the assumption that we can, in principle at least, complete a scientific explanation of human action. But the argument from Gödel shows that this assumption is false. In view of apparently unavoidable interdependence of human rationality and human action, it is implausible to suggest that we could find a complete scientific explanation of the second when we are prevented from finding a complete scientific explanation of the first. Therefore, Mackie is simply wrong when he claims that our ordinary personal explanations are, when properly spelled out, really just special cases of causal explanations. We are permanently unable to complete the 'spelling out' he has in mind. Consequently, the personal explanations which we give in relation to ordinary actions cannot be abbreviated or promissory causal explanations. Indeed, we must now regard most of the events for which we give personal explanations every day to be scientifically inexplicable. It is plausible to suppose that we will continue to believe that most of these personal explanations are correct. And it is plausible to suppose that we are unjustified in asserting the presence of natural laws where we know that no coherent theory can be constructed, even in principle. On these suppositions, we must conclude that in every correct personal explanation, the causal story of the connection of the intention with its fulfillment will stop short at some point. That is, after science has said everything it can about the events which take place in muscles, nerves, neurons and so forth, it will not have shown the link between the intention and its fulfillment. The connection will only be completed if the intention of the agent has some direct result which begins the scientifically explicable process leading to the fulfillment of the intention. We must posit that such direct results of intentions take place in every instance of correct personal explanation. But this conclusion is devastating to Mackie's objection. For it means that we have daily experience of the kind of direct connection between intentions and physical results, which is what Mackie actually finds so improbable in the theistic hypothesis.

Furthermore, my argument leaves science unable to complete any account of the origin (either evolutionary or in terms of individual growth and learning) of the capacities humans have to act intentionally as they do. Nevertheless, our daily experience confirms that humans do act intentionally. So against Mackie's doubts, we can say that we are entirely familiar with beings whose intentions bring about physical results, and yet for whose development we can give no complete causal account.

Viewed from a slightly different angle, these same considerations answer Mackie's doubts about the category of personal explanation as something separate from causal or scientific explanation. We have seen that true explanations in terms of the intentional actions of humans are scientifically inexplicable. Therefore it is not just that we are faced with a lot of gaps in what it is possible for science to explain it is that we already know and use an alternate scheme to fill a great many of those gaps, and that scheme is the scheme of personal explanation. Therefore Mackie has no right to complain that Swinburne is doing something irregular when he appeals to the category of personal explanation in theism as something separate from scientific explanation. For this reason and for the ones Swinburne has pressed, we can see that Mackie's criticisms give us no ground to reject Swinburne's Bayesian methodology.

The Existence of Human Rationality as Confirmation of Theism

In this section, I will use Swinburne's Bayesian methodology to assess the extent to which the existence of human rationality confirms theism. I will argue that the existence of human rationality, that is, the human ability, given finite time and resources, to recognize true propositions, raises the probability of the theistic hypothesis.

The first condition which must be met if the existence of human rationality is going to raise the probability of theism is that human rationality must be scientifically inexplicable. As we have seen, the argument from Gödel satisfies this condition. Human rationality is scientifically inexplicable, since every attempt has scientific explanation can be show to be incomplete by the Gödelian method.

The second condition which must be met for a Swinburne style inductive argument is that the existence of human rationality must be epistemically more probable on the hypothesis of theism than on the hypothesis that it exists uncaused, (i.e. with no explanation at all.) Of course, we can see that the probability of human rationality existing uncaused, as a brute fact of the universe, is very, very low. Human rationality is an extremely complex and orderly phenomenon. It unites the bewildering array of markings, noises and conscious episodes with which we are continually faced. The propositions signified by these markings, noises and conscious episodes have intricate relationships, many of which are perceived through our rationality. Such ordered diversity is just not to be expected unless there is some reason for its existence. So human rationality is very, very unlikely to exist uncaused.

But we can also see that there is more reason to expect the phenomenon of human rationality if God exists as described in Swinburne's theistic hypothesis37. First, if God exists, then there exists a being with the power to bring human rationality into existence. Second, as Swinburne has already shown in his argument from consciousness, God has good reason to bring about the existence of agents of limited power and knowledge who have the ability to increase their power and knowledge.38 But beings will only be able to move from limited knowledge to greater knowledge if they are in some sense rational. So God has good reason to bring about the existence of rational beings.

Furthermore, we can see that God might have good reason for choosing to create our particular type of rationality. One of the important features of our rationality is our ability to create and use algorithms and recognize truths about them. We can imagine having, and often wish to have, a more immediate grasp of complex logical points and a greater capacity to remember all the relevant issues of some problem. But our inadequacies in these areas can be overcome to a certain extent by our ability to construct rules for reasoning. With our rules for logic and arithmetic and so forth, we can use the limited memory and reasoning ability we have to solve intricate and lengthy problems. And often, when we find and train ourselves in largely automatic procedures in some domain, we achieve a level of speed and accuracy that closely approximates what we would expect from beings with far more internal processing power.

Now, there may well be good reasons for God to bring about beings with far more unconscious reasoning power and brute memory than we have, who for these reasons would not have to rely as heavily on algorithmic reasoning as we do. But we can see that he also has good reason to bring about beings like us. For constructing algorithms in any domain requires effort, and therefore requires a choice on our part. While algorithmic reasoning can impart great power and knowledge, these benefits do not come automatically. If we do not make the effort, we remain ignorant and weak, but with effort, we are able to create algorithmic structures which tremendously expand our knowledge and power. So by creating beings who must depend on algorithmic procedures, God creates beings with a wide range of choice concerning the amount of knowledge and power they achieve. It seems to be true, as Swinburne argues, that it is good thing that there be at least some beings in a largely do-it-yourself world. Consequently, it is a good thing that there should be beings with our kind of rationality. Therefore, we can see that God has good reason to bring about human rationality.

Thus, if God exists, it is not all that unlikely that he will choose to create human rationality. Since we said that human rationality is very, very unlikely to exist uncaused, the probability that something like human rationality will exist is increased by the hypothesis that God exists. By Bayesian analysis it follows that the existence of human rationality raises the probability of the theistic hypothesis.

Conclusion

Gödel's theorems have been called the most important logical theorems of this century. I concur with this assessment, and would add that their significance for wider areas of thought has not yet been fully appreciated by the philosophical world. As I showed in my first section, they clearly reduce naturalism to a position which is, at best, too weak to be tenable. Further, as I showed in my second section, they strengthen Swinburne's rigorous Bayesian case for theism in two ways. First, one implication of Gödel's theorems is that personal explanation is not reducible to scientific explanation. This raises the plausibility of theism, which relies on the special case of personal explanation by God's action. Second, since Gödel's theorems show that human rationality is scientifically inexplicable, and since there seem to be clear reasons why God might well create beings with our kind of rationality, human rationality is more to be expected with the theistic hypothesis than without it. In this way, the existence of human rationality becomes further confirmation of the existence of God.

ENDNOTES

1Kurt Gödel, "Über formal unentscheidbare Sätze der Principia Mathematica und verwandter Systeme," Part I, Monatshefte für Mathematik und Physik, vol. 38 (1931) 173-198; "Concerning Formally Undecidable Sentences of Principia Mathematica and Related Systems," trans. Elliot Mendelson, The Undecidable: Basic Papers on Undecidable Propositions, Unsolvable Problems and Computable Functions , ed. Martin Davis (Hewlett, New York: Raven Press, 1965) 5-38.

2See Hao Wang's account of Gödel's views in From Mathematics to Philosophy (New York: Humanities Press, 1974) 324.

3Gödel's paper actually used a special notion of consistency, called omega-consistency, but in 1936 J. B. Rosser extended Gödel's results to simply consistency.

4Here is a simplified description of how Gödel did it. Gödel showed that any sentence allowed in the formal system could be represented by a unique number, and by extension, how any series of legal moves leading to a proof of any sentence could also be represented by a unique number. Thus, the formal relationship of proof to proved sentence could be exactly represented by a particular mathematical relationship between the number of the proof and the number of the theorem. Gödel also showed how to construct sentences which dealt with this proof relationship, and so could talk about whether certain sentences were provable by the system by asking if any possible number stood in the proof relation to the number of the sentence. In addition, Gödel showed how sentences could be made to talk about their own provability by asking if any number stood in the proof relation to their own unique numbers. With these tools available, Gödel showed how to construct a sentence, the Gödel sentence, which said that no number stood in the proof relation to the unique number of the Gödel sentence. Clearly, if any number does in fact stand in the proof relation to the number of Gödel sentence, then the Gödel sentence is proved. But if the Gödel sentence is proved, then it is proved that no number stands in that relation. Thus it is proved both that there is and is not a number that stands in the proof relation to the number of the Gödel sentence. Since the formal system was said to be unable to prove contradictions, this result is impossible. So we must reject the supposition which led to the contradiction, and affirm instead that it is true that no number stands in the proof relation to the number of the Gödel sentence. Since this is what the Gödel sentence claimed, we have a sentence stating a clear mathematical truth, but which cannot be proved by the system. The system thus fails to formalize all mathematical truth.

5The easiest access to Turing's work, as well as the work of researchers like Emil Post and Alonzo Church who helped show the more general significance of Gödel's techniques, is The Undecidable, Martin Davis (ed.), (Raven Press; Hewlett, New York, 1965).

6Roger Penrose has an accessible discussion of Turing machines and Turing's incompleteness results in The Emperor's New Mind, (Oxford: Oxford University Press, 1989), 30-73.

7Roger Penrose has a similar treatment of learning computers in Behavior and Brain Sciences 13:4, (1990), 698.

8J. R. Lucas spells out this argument in a clear way in The Freedom of the Will, (Oxford: Clarendon Press, 1970) 134-136.

9J. R. Lucas, "Minds, Machines and Gödel", Philosophy 36 (1961), 112-127.

10Lucas, Freedom.

11Douglas Hofstadter, Gödel, Escher, Bach: An Eternal Golden Braid, (New York: Basic Books, 1979). He credits Lucas with stimulating his early thought on this subject on p. 472 of the paperback edition (New York: Vintage Books, 1989).

12Hilary Putnam, "Reflexive Reflections," Erkenntnis 22 (1985), 143-153.

13Penrose, Emperor's.

14Roger Penrose, Shadows of the Mind, (Oxford: Oxford University Press, 1994).

15For a massive single collection of responses, see "The Emperor's New Mind: Précis, commentary and author response," Behavioral and Brain Sciences 13:4 (1990) 643-705.

16To see why, suppose that we have some formal system of logic which proves a contradiction such as "Pi is a real number" and "It is not true that pi is a real number." Since "Pi is a real number" is proved, it is easy to prove another sentence "Either pi is a real number or the moon is made of Gouda." We can prove this because it is the nature of either or sentences that if one part is true, the whole sentence is true. But we have also already proved "It is not true that pi is a real number." Using this plus the proved sentence "Either pi is a real number or the moon is made of Gouda" we can prove "The moon is made of Gouda." The problem is that this scheme of proof works no matter what you plug in as a contradictions and no matter what you plug in the place of "The moon is made of Gouda." This scheme will prove that "There is a God" and that "There is no God;" that "IBM stock will split next Thursday" and that "IBM will declare bankruptcy next Thursday." and so on for absolutely every sentence which can be formed. Some technically minded readers might object that there are several logical systems, error tolerant logics and some relevance logics, which reject the quick proof used here. My response is that an error still spreads in these systems. Thus, it requires some work to claim that a contradiction does not lead to a proof of all sentences at closure, and a lot more work to show that a contradiction will not lead to such a massive breakdown of reliability as to rule out such a system as an adequate candidate to be the HUMAN GENOME MAP.

17This is essentially Lucas's position. See Lucas, Freedom,158.

18Penrose, Emperor's, 437-9.

19Penrose, Shadows, 371 ff.

20Daniel Dennett, "Murmurs in the Cathedral," Times Literary Supplement (Sept. 26-Oct. 5, 1989), 1066-68; "Betting your life on an algorithm," Behavioral and Brain Sciences 13:4 (1990), 660&1; Darwin's Dangerous Idea, (New York: Simon & Schuster, 1995), chapter 15.

21Dennett, Darwin's, 438.

22See for instance Dennett, Darwin's, 17-23 or 511-521.

23Lucas, Freedom, p. 158.

24At least one of his refutations of Lucas followed this tack by appealing to the essentially interpretive nature of intentionality. See Dennett, Brainstorms, (The Harvester Press, 1979), 256-266. Dennett has been backing away from the non-realism involved in this position in more recent works, so it hard to tell where he would come down now.

25Paul Benacerraf, "God, the Devil, and Gödel," The Monist 61 (1967): 29.

26Hofstadter, 473-476.

27Lucas, Freedom, 154.

28Richard Swinburne, The Existence of God (Oxford: Clarendon Press, 1979).

29John Mackie, TheMiracle of Theism. (Oxford: Clarendon Press, 1982), 100; 129-132; 149.

30Richard Swinburne, "Mackie, Induction, and God," Religious Studies 19 (1983): 387. This article has been reprinted as an appendix in the 1991 edition of Swinburne's The Existence of God (Oxford: Clarendon Press, 1991) 293-299.

31Swinburne, "Mackie," 386.

32Swinburne, "Mackie," 387.

33Mackie, Miracle 130.

34Mackie, Miracle 130.

35See his comments on 19&20.

36John Mackie, The Cement of the Universe (Oxford: Clarendon Press, 1974), ch. 11.

37Swinburne, God 8.

38Swinburne, God 155-160.

WORKS CITED

Benacerraf, Paul. "God, the Devil, and Gödel," The Monist 61 (1967): 29.

Dennett, Daniel. "Betting your life on an algorithm," Behavioral and Brain Sciences 13:4 (1990): 660&1.

---. Brainstorms. The Harvester Press, 1979.

---. Consciousness Explained. Boston: Little, Brown & Co., 1991.

---. Darwin's Dangerous Idea. New York: Simon & Schuster, 1995.

---. "Murmurs in the Cathedral," Times Literary Supplement (Sept. 26-Oct. 5, 1989): 1066-68.

Gödel, Kurt. "Concerning formally undecidable sentences of Principia Mathematica and related systems." Trans. Elliot Mendelson. The Undecidable: Basic Papers on Undecidable Propositions, Unsolvable Problems and Computable Functions. Ed. Martin Davis. Raven Press; Hewlett, New York, 1965. 5-38.

---. "Über formal unentscheidbare Sätze der Principia Mathematica und verwandter Systeme," Part I, Monatshefte für Mathematik und Physik, Vol. 38 (1931): 173-198.

Hofstadter, Douglas. Gödel, Escher, Bach: An Eternal Golden Braid. New York: Basic Books, 1979. New York: Vintage Books, 1989.

Lucas, J. R. "Minds, Machines and Gödel," Philosophy 36 (1961):. 112-127.

---. The Freedom of the Will. Oxford: Clarendon Press, 1970.

Mackie, John. The Cement of the Universe. Oxford: Clarendon Press, 1974.

---. The Miracle of Theism. Oxford: Clarendon Press, 1982.

Penrose, Roger. Shadows of the Mind. Oxford: Oxford University Press, 1994.

---. The Emperor's New Mind. Oxford: Oxford University Press, 1989.

Penrose, Roger et al. "The Emperor's New Mind: Précis, commentary and author response," Behavioral and Brain Sciences 13:4 (1990): 643-705.

Putnam, Hilary. "Reflexive Reflections," Erkenntnis 22 (1985): 143-153.

Rosser, J. B. "Extensions of Some Theorems of Gödel and Church." Journal of Symbolic Logic 1 (1936): 87-91.

Swinburne, Richard. The Evolution of the Soul. Oxford: Clarendon Press, 1986.

---. "Mackie, Induction, and God," Religious Studies 19 (1983): 385-391.

---. The Existence of God. Oxford: Clarendon Press, 1979.

Wang, Hao. From Mathematics to Philosophy. New York: Humanities Press, 1974.