Are the Gödel incompleteness theorems limitative results for the neurosciences?

Jeff Buechner

Received: 2 December 2008 / Accepted: 22 April 2009 /
Published online: 12 June 2009
© Springer Science + Business Media B.V. 2009

Abstract There are many kinds of limitative results in the sciences, some of which are philosophical. I am interested in examining one kind of limitative result in the neurosciences that is mathematical—a result secured by the Gödel incompleteness theorems. I will view the incompleteness theorems as independence results, develop a connection with independence results in set theory, and then argue that work in the neurosciences (as well as in molecular, systems and synthetic biology) may well avoid these mathematical limitative results. In showing this, I argue that demonstrating that one cannot avoid them is a computational task that is beyond the computational capacities of finitary minds. Along the way, I reformulate three philosophical claims about the nature of consciousness in terms of the Gödel incompleteness theorems and argue that these precise reformulations of the claims can be disarmed.

Keywords Gödel incompleteness theorems · Watts–Strogatz random networks · Barabasi–Albert scale-free networks

1 Introduction: limitative results in the sciences

There are various senses in which there can be limitative results in the sciences, as well as various kinds of limitative results. Here, I will focus on the neurosciences, but my main conclusion carries over to molecular biology, to systems biology, and to synthetic biology. There are limitative results in various categories: philosophical, mathematical, physical,

J. Buechner (✉)
Department of Philosophy, Rutgers University-Newark, Newark, NJ, USA
e-mail: buechner@rci.rutgers.edu, jbuechner@gc.cuny.edu

J. Buechner
The Saul Kripke Center, CUNY, The Graduate Center, New York, NY, USA

economic, social, political, ethical (and more). For instance, an economic limitative result is that there is not enough funding to carry out experiments necessary to test certain theoretical claims in a science. An example from physics is the super-hadron collider in Europe. An attempt to build a comparable apparatus in the United States was rejected by a vote in Congress. In the neurosciences, what experiments are deemed worth carrying out can be a political decision. Thus, certain ideas will not be tested because authorities in the field decide that either there is little merit to them and that it would be an unnecessary expense to test them or else that, although there may be merit to them, there are other ideas with more merit that should be tested first. We might also use ideas in game theory to explain why some ideas will be tested while others will not, using descriptive terms such as payoffs measured in terms of monetary utility. There might be ethical reasons not to conduct such experiments. The money used to conduct the experiment might be taken from a fund for providing food for malnourished children and such an expropriation would be unethical. These limitative results would cease to hold if circumstances had been different. Are there limitative results that persist under (most) changes in circumstance?

1.1 Data-model maps and limitative results

There are concerns about whether mathematical models used in the neurosciences accurately describe the phenomena. Do mathematical models map onto the phenomena? For a case in point, consider whether coupled oscillatory theory maps onto real brain cells. J. L. Perez Velazquez, in an important paper on limitative results in the neurosciences, argues that we rarely find intrinsic oscillators that are weakly and globally coupled in real brain cells [1]. The failure to find a mapping of the right sort between a mathematical model and real brain cells is a limitative result only when it is an instance of one (or more) of the qualitatively distinct limitative results described below. If not, it is an empirical question about what is the appropriate mathematical model for describing the data. Unless there is a reason to think that there can be no mathematical model to describe the data—that *would* be a limitative result—the failure to find the best model is simply an empirical matter.

1.2 Underdetermination limitative results

Philosophical limitative results are another class of limitative results in the sciences. These limitative results are interesting because they reveal limitations that have to do with (1) the essential features of scientific theories and (2) with observation. One general kind of philosophical limitative result is that we will never find the correct theoretical description of the phenomena which we are trying to understand. There are various ways in which this claim can be argued. One way is note that there are infinitely many distinct scientific theories, of a given subject matter, each of which accommodates any finite set of empirical data equally well. Conjoining to any finite data set additional finitary sets of data points will not help matters. Although the new data set may well eliminate from contention many scientific theories, which were compatible with the old data set, there are infinitely many new scientific theories, each of which are compatible with the new data set. It follows from this that the only way to break the deadlock—to show that the theory in question and no other is confirmed by the data—is to construct a data set consisting of infinitely many data points. Unfortunately, human beings are finitary and cannot either collect or survey infinite data sets. It is our finitary limitations that preclude circumventing this form of philosophical limitative result, known as the underdetermination of theory by data problem [2].

1.3 Data interpretation limitative results

Another closely related philosophical limitative result is that we have to interpret the data we observe by means of a theory and that this is a purely subjective procedure that will depend on the interpreter. Thus, what data we have has a twofold dependency. It is dependent upon the particular theory through which we interpret the data and upon the particular human being—that is, the interpreter of the data—who looks at the data through the particular theory she holds. (For instance, it has been argued that assigning priors in Bayesian decision theory is an ineliminably subjective procedure.) I will not examine these kinds of limitative results here since they are philosophical claims that can and have been argued both pro and con over the years. Moreover, these kinds of claims can be directed against almost any subject matter at all. They are essentially skeptical claims. Skepticism has importance within philosophy and outside as well (think of the advances in statistical methods in theory testing and in discerning causal structure in physical processes, which are partially motivated by an attempt to respond to skepticism) [3, 4].

1.4 Intrinsic/extrinsic and metaphysical/epistemic limitative results

The concept of a limitative result in the sciences is quite complex. There are several dimensions along which it can vary. I make two distinctions that isolate four possibilities: intrinsic/extrinsic and metaphysical/epistemic. A limitative result in the sciences is extrinsic when it applies to our theory of a particular domain of inquiry in that science. A limitative result is intrinsic when it applies to the structures in the science, such as the computation of neurobiological parameters by neural structures [5]. Extrinsic limitative results are concerned with theories conceived of by human beings, and intrinsic limitative results are concerned with actual physical objects such as neurobiological structures.

Prima facie, one would think that any correct theory of, say, a neurobiological process, would necessarily coincide with the computations such neurobiological processes engage in. After all, any scientific theory is an attempt to describe and explain what happens in nature. It would be odd if one's scientific theory, though correct, failed to describe and explain what nature does. However, correct scientific theories may make certain idealizations that are not found in nature. For instance, in statistical thermodynamics, one idealizes to an infinite number of particles. A description of what actually happens at the molecular level in a thermodynamical process would be quite different from what the theory says happens.

A limitative result is metaphysical when it reveals that a limitation is an essential feature of some physical process. The Kochen–Specker theorem revealed that a limitation (indeterminacy of physical parameters) is an essential feature of quantum mechanics [6]. A limitative result is epistemic when it shows we cannot know how some physical process works, even though there is a fact of the matter as to how it works.

1.5 Resource limitative results

Another kind of mathematical limitative result in the sciences is that the proofs of mathematical theorems expressing claims in a particular science are too long to be humanly feasible [7–9]. Here there are mathematical results about the lengths of proofs of theorems, in, for instance, set theory. This kind of mathematical limitative result takes hold because human beings do not have the computational resources necessary to carry out the calculations,

even though the resources necessary to do that are finitary. It is *not* the case that it is mathematically impossible to carry out the calculations. Rather, what is mathematically possible for us to do is prohibitively resource-expensive. For instance, to compute a truth table for a sentence in propositional logic containing 300 atomic sentences—assuming that a single line of the truth table could be inscribed on an elementary particle—would require more matter than is available in the known universe. Our interest here is in mathematical limitative results that state what it is mathematically *impossible* for humans to do in the sciences.

1.6 Mathematical limitative results

My interest is whether there are limitative results in the neurosciences, which are mathematical and which are not wholly philosophical in character. In mathematics, there are well-known impossibility results such as the impossibility of solving arbitrary polynomial equations of degree 5 or greater by radicals. These results are important—they show the mathematician what it is impossible to do. There are no further questions that can be raised. Mathematical limitative results persist under changes in circumstances. Are there impossibility results in mathematics that apply to the neurosciences? That is, are there mathematical theorems which say, for instance, of a claim in the neurosciences, that it cannot in principle be established because to do so would contradict an established result in mathematics? I am not interested in providing a general characterization of any kind of mathematical limitative result whatsoever, principally because to establish such a result one would have to have a characterization of the sciences and of mathematics which is beyond what we now have. Further, if such a characterization were entirely formal, then the very description of the limitative results might itself be limited by the well-known Gödel incompleteness theorems. The mathematical impossibility result I will examine, the Gödel incompleteness theorems, have been recruited in psychology in arguments that it is mathematically impossible to characterize human cognition by a computational description. However, those arguments make questionable philosophical and empirical assumptions about human cognition.

I argue below that it can be conclusively shown that the Gödel incompleteness theorems cannot be used as a mathematical limitative result in psychology (unless it can be proved that the human mind can construct infinite proof trees) and, for similar reasons, that they cannot be used as a mathematical limitative result in the neurosciences (and in molecular biology, systems biology, and in synthetic biology). Before giving these arguments, the claim that the Gödel theorems can be used as mathematical limitative results in the neurosciences will be given “a run for its money”. That is, I will show how someone might argue that the Gödel theorems are mathematical limitative results in the neurosciences. After doing that, I argue that the Gödel theorems cannot be used as mathematical limitative results in the neurosciences.

2 The Gödel incompleteness theorems

The Gödel incompleteness theorems are well-known mathematical impossibility results, but their exact mathematical relevance in the mathematics community is debated, even today. Before 1977, it was widely held by mathematicians and scientists that the bizarre self-referential sentences that Gödel devised (that say of themselves that they are unprovable)

had no applications within either pure mathematics or the sciences. In 1977, that changed when Jeff Paris and Leo Harrington proved that an extended version of the finite form of the well-known Ramsey theorem on homogeneous partitions of infinite sets could not be proved nor could its negation in Peano arithmetic [10]. (Peano arithmetic is the set of truths of arithmetic that can be proved from the five axioms that Peano—a nineteenth century mathematician—employed to characterize arithmetic. The most distinctive of the five axioms is the axiom of induction, which is used in proof by mathematical induction.) This was the first instance of an actual mathematical sentence that could not be proved nor disproved within Peano arithmetic. It is worth remarking that the technique employed by Paris and Harrington (and generalized by Saul Kripke in his method of fulfillability [11]) was to construct a model of Peano arithmetic in which the mathematical sentence is true and one in which it is false (that is, one in which its negation is true). This is a model-theoretic method of obtaining an incompleteness theorem and is different from the original syntactical method employed by Gödel in 1931 [12]. Historically, the beauty of the technique is that once one mathematical sentence had been shown to be independent of the axioms of Peano arithmetic, many others followed. Just as the invention of forcing in the 1960s revolutionized set theory, in that it spawned many demonstrations of sentences in set theory that were independent of the axioms of ZFC (Zermelo–Fraenkel set theory with the axiom of choice), so too did the Paris–Harrington construction and its generalization by Kripke revolutionize how working logicians dealt with Peano arithmetic. Independence results for Peano arithmetic, in addition to finite versions of Ramsey theorems, include (but are not exhausted by) generalized Goodstein sequences, ordinal diagrams, dilators, problems concerning graph minors, and variants of Kruskal’s theorem.

The point of speaking of the Gödel sentences as independence results is that there is now a large body of results (many of them formulated by Harvey Friedman [13]) on independence results in areas of mathematics that are the province of working mathematicians. These independence results can be viewed as mathematical limitative results on a par with the Gödel incompleteness theorems since the latter can be naturally viewed as independence results. Moreover, should the criticism be made that the Gödel incompleteness theorems have only a marginal connection with the areas of mathematics that are the province of working mathematicians, it can easily be deflected by pointing to the independence results that do infect these areas of mathematics. In what follows, however, the discussion of mathematical limitative results and how they can be effectively disarmed will be carried out by employing the Gödel incompleteness theorems.

At this point, a brief description of the Gödel incompleteness theorems is necessary. Formal systems are formal languages augmented with a set of axioms and a set of rules of inference (where a formal language consists of a set of symbols and a set of formation rules over the symbols for the construction of well-formed formulas, with which the formal language is identified). Suppose that one is given a formal system of first-order logic that has expressive power rich enough to express sentences of minimal arithmetic (which is weaker than Peano arithmetic) and has the syntactical resources to code (by some system of coding, such as Gödel numbering) relations and properties of various kinds, such as the proof predicate. (Here, I do not distinguish between various strengths of arithmetic, namely, Peano, Robinson, and minimal arithmetic. Minimal arithmetic does not include all of number theory. The point of using minimal arithmetic is to show that the Gödel incompleteness theorems take root in surprisingly weak parts of arithmetic—and thus take root in all extensions of the weak parts as well.) Then, there are truths of arithmetic that can be expressed within the formal system, but which cannot be proved within it, provided

that the formal system is consistent. This is the first Gödel incompleteness theorem. This shows the formal system is not complete, where completeness of a formal system is relative to the class of truths which the formal system, when given the appropriate semantical interpretation, formalizes. (The formal system of Peano arithmetic formalizes the truths of Peano arithmetic—where formalizing a truth of Peano arithmetic is to prove in the formal system for Peano arithmetic the formal sentence which, under the appropriate interpretation, expresses that truth. An interpretation is a map from syntactical objects—symbols—to objects such as mathematical objects.)

Completeness can then be defined as a formal system is complete with respect to some class of truths (such as the truths of Peano arithmetic) if and only if the language of the formal system is adequate for expressing any truth of Peano arithmetic, and every truth of Peano arithmetic is a theorem of the formal system. It is also required that the formal system is consistent, since, if it is not consistent, then anything which can be expressed within the formal system is provable within it. (Consistency can be defined in various ways. Simple consistency is defined as: For no formula F of a formal system, it and its negation ($\text{not-}F$) are provable. This definition is syntactical. The point of a syntactical definition is to produce a definition that is applicable to any formal system independently of any interpretation that might be imposed on any of them. The motivation of the concept of consistency is to capture the idea that nothing which is absurd can be proven in a formal system.)

It follows from the first Gödel incompleteness theorem that, if the formal system is consistent, then that fact about it is unprovable in it. This latter statement is Gödel's second incompleteness theorem. There are other ways in which the first incompleteness theorem can be stated. For instance, there is no complete, consistent, and axiomatizable extension of minimal arithmetic. One last point: within the formal system in which first-order proofs appear, neither the Gödel sentence nor its negation can be proved. Suppose that the formal system is sound and consistent, so that it never proves untruths. Then, the upshot is that if we only use finitary proofs to establish arithmetical truth, we cannot know whether the Gödel sentence or its negation is true.

The relation of the incompleteness theorems to independence proofs in mathematics is transparent. An independence proof shows that a particular statement cannot be proved from a given set of assumptions. (For instance, the parallels postulate cannot be proved from the axioms of Euclidean geometry.) This means that there are models of the mathematical theory in which the claim is true and models in which it is false. If the only way we can access mathematical truth is through proof, then we will not know whether a statement proved to be independent is true or false unless we can postulate axioms from which it can be proved. But, in that case, the question about the truth or falsity of the independent statement is transferred to the question of the truth or falsity of the axioms employed to prove it.

Gregory Chaitin has proved incompleteness of formal systems strong enough for arithmetic in a different framework—that of algorithmic information theory [14]. Chaitin's basic idea is that information expressed in the axioms of Peano arithmetic is not sufficient to determine all of the truths of arithmetic. He uses notions of information, computability, program size, and randomness. A central point I make below, about the role of mathematical certainty in incompleteness phenomena, carries over to Chaitin's work. For that reason, I use in the exposition below the framework of Gödelian incompleteness as an independence result in mathematics, which I prefer to Chaitin's framework of randomness, information, and noncomputability.

2.1 Harvey Friedman's work on undecidability in linear order systems

Do the Gödel incompleteness theorems surface in the neurosciences, in molecular biology, in systems biology, or in synthetic biology? The point of the following section is to provide reasons why they do. In a recent paper—"Limitations On Our Understanding of the Behavior of Simplified Physical Systems" [15]—Harvey Friedman defines the abstract notion of a Linear Order System (which is an ingenious way of simulating a Turing machine in a physical system). Although we only need to cite some of Friedman's results here, we should say what a Linear Order System consists in. Identify spatial points with integers and identify points in time with positive integers. In any given physical system that is a realization of a Linear Order System, there are n bodies, each of whose positions and times are given by: $B_i[t]$, $1 \leq i \leq n$. " B_i " identifies the i th body, " t " identifies the time, and " $B_i[t]$ " identifies the position of the i th body. An initial configuration of the n bodies can easily be given. Assume that movement of the bodies is wholly deterministic and quite restricted. At any given moment of time, any body can move either to its left one unit of space, to its right one unit of space, or remain at its current position. (Notice that motion is simply a matter of the relative order of all of the bodies. It is computed by noting the positions of a body at two different instants of time.)

Friedman imagines that one can think of each body continuously transmitting some form of signal such as radiation. In that way, one body can be aware of another body either to its left or its right (or even at the same position) by detecting the signal (that is, the radiation) that it transmits. It can distinguish between a body to its immediate left and one to its left—but not its immediate left—by the strength of the radiation signal that it receives. In that way, the body can determine its relative distance from all of the other bodies. It follows that any given body has the capacity to infer the relative order of all n bodies.

Friedman provides limitative theorems for Linear Order Systems of two different kinds. One is an algorithmic undecidability theorem, which generates an infinite class of statements, none of which can be decided by an algorithm. The other kind of undecidability is one which generates a single statement, and it is this kind of undecidability that we think of as at the core of the Gödel incompleteness theorems. Given any single statement, there is always an algorithm that can decide its truth value. If the statement is true, then the algorithm which says it is true is certainly correct. The question is, though, not whether an algorithm is correct, but whether the procedure it provides for obtaining the correct answer is informative and not simply trivial. However, making the distinction is not always possible and, where possible, not always feasible. For that reason, talk of decidability by an algorithm should be replaced by talk of provability within some formal system, such as either ZFC or PM (The *Principia Mathematica* of Bertrand Russell and A. N. Whitehead [16]). One of Friedman's limitative theorems for Linear Order Systems is that it is impossible to prove that a physical Linear Order System is bounded after it evolves over time from some initial configuration [15].

What relevance do Friedman's results have for the neurosciences (or for molecular biology, or systems biology, or synthetic biology)? Given a crude model of neural processing, where each neuron sends and receives signals from all other neurons, the neurons are stationary, and the signals are propagated through space. What if we think of the signals as bodies that change position? This is a way to model multiple interacting signals in a complex brain circuit. Friedman's theorem tells us that, given certain initial conditions, the statement that the signals are bounded (to a certain area of the brain) is neither provable

nor refutable (in a system of reasoning that is strong enough to encompass ZFC). This is significant since it would show that we are mathematically limited in predicting global connections among brain circuits.

There are other applications of the theorem. Certainly, one can think of its usefulness in systems biology and synthetic biology. We may not be able to prove or refute claims about the boundaries of complex networks of interacting molecules, such as long distance gene regulation by promoter and regulatory sequences that, though tens of thousands of base pairs distant from the gene, are nonetheless spatially close to it owing to the double helical three-dimensional structure of DNA. Or, in modeling protein functionality, there are several distinct domains on the protein, each of which has its own functionality and interactions with molecular complexes in the cell. The composite of these interactions is a complex system of molecules. Friedman's theorem tells us that we might not be able to prove or refute claims about the boundaries of the subsystems or the boundaries of the composite evolving system.

There are network models of cortical circuits that are used to model the dynamical processes such circuits undergo which are far more realistic than the earlier family of lattice models. For instance, Watts–Strogatz random networks [17] or Barabasi–Albert scale-free networks [18] are prominent examples of new classes of network models that are capable of treating dynamical processes in cortical circuits, such as synchronization and plasticity properties among large populations of neurons. The patterns of activity represented by activated nodes in the networks change over time. Where there are mathematical claims about the distribution of activity in the network, Friedman's theorem points to a restriction on what we can prove (or disprove) about the boundary conditions of the activated nodes in the network.

2.2 Do the Gödel incompleteness theorems arise in areas of mathematics other than finite combinatorics and Peano arithmetic?

There are many different mathematical approaches which are employed in the neurosciences. Should those who use such approaches worry about the Gödel incompleteness theorems? After all, one objection might be that the kinds of true but unprovable statements that are the province of the Gödel incompleteness theorems are fairly limited, mainly to finite combinatorics and arithmetic. The neuroscientist, pretty much no matter which mathematics she uses, will not have to worry about the Gödel incompleteness theorems.

This objection is entirely bankrupt, however, since there is a good deal of mathematical evidence that the phenomenon of incompleteness is pervasive throughout mathematics (contrary to the impression those outside logic might have of it). Since Gödelian incompleteness is an independence phenomenon, the results on independence in set theory are relevant to the objection. In particular, the spectacular development of Boolean Relation Theory within the last 20 years by Harvey Friedman is significant here, since it shows that there are many instances of finite Boolean relations that are part and parcel of many different branches of mathematics and that are independent of reasoning in the system of set theory ZFC. Here is Friedman on the mathematics of the near future: "Every interesting substantial mathematical theorem can be recast as one among a natural finite set of statements, all of which can be decided using well studied extensions of ZFC, but not all of which can be decided within ZFC itself." [19]

2.3 Intrinsic and extrinsic limitative results and the Gödel theorems

Suppose that on our theory of biological processes, there is a Gödelian limitative result that applies to the values of a certain parameter specified by the theory. Now, suppose that at the biological level, nature somehow contrives to avoid that Gödelian limitative result by engaging in a biological process that is not subject to those limitations and which nicely approximates to those values for that parameter (governing the biological processes) which the biological system would satisfy if it did not suffer from Gödelian limitations.

The point is that nature is artful, and one way in which it can be so is in designing mechanisms that somehow overcome limitative results. If the design approximates very closely to the actual values that are constrained by the Gödelian limitations, we—as humans investigating this process—will not be able to detect it. We will believe both that our theoretical claims about the biological process cannot be proved (since those claims are constrained by the Gödelian limitative results) and that the biological process is constrained in a similar way. In this case, we would miss what nature actually does because we are misled by the limitative results. That is, in believing that nature is similarly constrained by the limitative results, we are under the illusion that our theoretical claims about the natural process, constrained by the Gödelian limitative results, accurately describe that biological process. From our correct belief that our theoretical claims are constrained by the Gödelian limitative results, we incorrectly infer that the biological processes about which we theorize are similarly constrained. True, the biological processes are similarly constrained if they make computations that are Gödel constrained in the same way that our theoretical description of those computations is Gödel constrained. Because we have this true belief, we fail to see that nature sidesteps these limitations. Nature does not engage in Gödel constrained computations, but our theory says it does.

3 The use of the Gödel incompleteness theorems as limitative results in psychology

The first use of the Gödel incompleteness theorems as limitative results in the sciences is in psychology [20]. They were recruited to show that human cognition has no computational description. This kind of limitation, if true, would put a severe crimp in the program of cognitive science, which takes human cognition to have a computational description. The limitation is established by showing that machine computational states are limited in their computational powers by the Gödel incompleteness theorems, while human cognitive states are not similarly limited. Thus, human cognitive states are not identical with machine computational states. As limitative results in psychology, the Gödel theorems show that human cognition surpasses the computational powers of any machine designed to simulate it. As limitative results in the neurosciences, the Gödel theorems show the human nervous system surpasses the power of any theory (expressed in logic) designed to explain and understand it.

I will now examine how the Gödel incompleteness theorems are used as mathematical limitative results in psychology, argue that they fail to achieve that purpose, and then argue that the Gödel incompleteness theorems cannot be used as limitative results in the neurosciences (and molecular biology, systems biology, and synthetic biology as well). The point to keep in mind in the following is that the argument for why Gödel incompleteness theorems fail as limitative results in psychology directly carries over to the neurosciences.

3.1 Arguing that human cognition has no computational description

Although Gödel had the idea in 1951 [21], J. R. Lucas published a paper in 1961 [22], where he proposed a magic bullet to destroy the view that human cognition has a computational description. He proposed a way to mathematically distinguish human cognition from computing machines. The idea is simple. A computing machine capable of activities of a certain complexity—in this case, being able to prove theorems of arithmetic—is subject to the Gödel incompleteness theorems. Human cognition, if it does not have a computational description, is not subject to them. Gödel's theorems provide a precise way of expressing the difference between human cognition and computing machines. If human cognition is not subject to the Gödel theorems, it can prove the truths in arithmetic that the machine (conjectured to simulate it) cannot prove and that everything it proves in arithmetic is true. Given a machine subject to the Gödel theorems, it is easy to formulate truths in arithmetic it cannot prove. Lucas formulated a clever strategy to show human cognition is not captured by a computational description. Hypothesize human cognition is a machine of type M . Construct M 's Gödel sentence. Human cognition can prove it, contradicting the hypothesis that it is a machine of type M . This strategy holds for arbitrary M . Thus for any M , there are no machines of type M that capture human cognition. Computational functionalism—the philosophical view that human cognition is captured in a computational description—is refuted.

Gödel's theorems have fascinated anti-functionalist philosophers. Their hope is that the theorems will provide a mathematical proof that mechanism is false. They also fascinate those who wish to show that the neurosciences are fundamentally limited and, beyond that, that the aims of molecular biology, systems biology, and synthetic biology cannot be attained. If it turns out that brain circuits can be modeled in a logic strong enough to express arithmetic, then there will be truths about those circuits that cannot be proved in any theory in the neurosciences. If that is the case, there is an ineliminable obstacle in the program of reducing human cognitive states to neurophysiological states since there is no complete description of the neurophysiological states. The Gödel incompleteness theorems, in that case, would have revealed limitations in what we can know about human cognitive processes.

Similarly, there may be truths in systems biology that cannot be proved. This would be a blow to systems biologists whose goal is to artificially create cellular structures. These limitative claims are extraordinarily general. They articulate constraints on any possible scientific approach to either the neurosciences or systems biology since they are claims about the mathematical possibility of success in those sciences. If the Gödel theorems apply to the neurosciences or to systems biology, then there are justifiable doubts that these programs can succeed. Moreover, if a biological system requires information that can only be supplied if some computation is performed that is subject to the Gödel theorems, then the limitation is not just a limitation on what human beings can scientifically demonstrate about either neurophysiology or systems biology. It is also a limitation on what cells can do: how they function. It is a limitation on cells that constrains their essential nature. Further, it is a limitation on the kinds of synthetic cellular structures we can build. Thus, the project of synthetic biology might face two kinds of limitations. One is a limitation on what we can scientifically demonstrate; the other is a limitation on what we can successfully design.

3.2 Hilary Putnam's problem for using the Gödel incompleteness theorems as limitative results in psychology

If human cognition has noncomputable powers, it can prove truths of arithmetic that a computational device intended to simulate it cannot prove. Conjecture a formal system for human arithmetical ability. That formal system—in which some truths of arithmetic cannot be proved—must be consistent. If it is not consistent, then the Gödel sentence for any computational device physically realizing that formal system is provable in it since in an inconsistent formal system anything expressible in it can be proved in it. If human cognition can prove the Gödel sentence of a computational device conjectured to simulate it, then it must also be able to prove that the formal system physically realized in that computational device is consistent. If it is not consistent, then the computational device can also prove its Gödel sentence, in which case there is no way to distinguish human cognition from the computational device intended to simulate it. In that case, the Gödel incompleteness theorems could not be used as limitative results in psychology.

Suppose that the computer program for human arithmetical ability is so long that no human being can survey it. We can imagine a program that consists of 10^{50} lines of code (in a high-level programming language for which it would be futile—because the resulting program would be totally opaque to the user—to create MACROS to decrease program length). If so, no human being could prove that it is consistent. It follows that there would be no way of distinguishing human arithmetical ability from a computational device that is intended to simulate that ability. This is the idea Hilary Putnam employed [23] to show that the Gödel incompleteness theorems are not limitative results in psychology. His target was Roger Penrose and, in particular, Penrose's defense of his argument that the mind has no computational description given in *Shadows of the Mind* [24]. A difficulty with Putnam's argument is that the length of the computer program describing the computational structure of human arithmetical ability is not known. Since it is not one among a number of possibilities we can examine by search through a search space, it does not refute the claim that the Gödel incompleteness theorems are limitative results in psychology.

In 1984, Putnam proposed an ingenious argument, which he claimed avoided Penrose's error and which restored the Gödel incompleteness theorems as limitative results in psychology. That his argument is invalid is argued in detail in my book *Gödel, Putnam and Functionalism* [20]. As we shall see below, even if human beings *could* prove the consistency of any formal system strong enough to express the truths of arithmetic, the Gödel incompleteness theorems could not be used as limitative results in psychology. The reason is straightforward, but it has eluded most thinkers who have weighed in on the role of the Gödel theorems as limitative results in psychology.

3.3 The role of mathematical certainty in Gödelian limitative results

What eluded Hilary Putnam, philosophers, mathematicians, cognitive scientists, and neuroscientists is that the Gödel theorems show that no one—whether the Gödel theorems apply to them or not—can finitistically prove the consistency of Peano arithmetic *with mathematical certainty*. They do *not* show that one cannot prove the consistency of Peano arithmetic with less than mathematical certainty. The proof relation of a formal system

confers mathematical certainty upon everything that is proved in it. This importantly qualifies any claim about what one can and cannot prove in a formal system. The *only* way finitary beings can achieve mathematical certainty in what they prove is to prove it in a finitary formal system. There are few results in mathematics that are proved with mathematical certainty since few mathematicians prove their results in a finitary formal system (such as first-order logic). No being—not even God—could prove a Gödel sentence with mathematical certainty in a finitary formal system. The only way to prove a Gödel sentence with mathematical certainty is to either use a stronger finitary formal system—in which case there will be a new Gödel sentence that cannot be proved in it—or to employ an infinitary system in which one constructs infinitary proofs. The latter is within the powers of God, but it is not within the powers of finitary human beings. We cannot construct infinitary proof trees.

The upshot is that no finitary human being can use the Gödel incompleteness theorems to show there are proof-theoretic powers human cognition has that no computational device intended to simulate it can capture. In general, in any formal system to which the Gödel theorems apply, finitary beings cannot prove a mathematical truth in the epistemic modality of the proof relation defined in it. The first-order formal system for Peano arithmetic provides a proof relation whose epistemic modality is mathematical certainty. Where mathematical theorem-proving is not done in a formal system, the epistemic modality of the proofs is less than mathematical certainty.

If we demand that the standard of proof in the sciences is mathematical certainty, then the Gödel theorems show that we are limited in what we can prove in the sciences. The Gödel incompleteness theorems are limitative results wherever we require that we prove our claims with mathematical certainty. What we will now argue is that when we relax those standards of scientific proof, it cannot be shown that the Gödel incompleteness theorems are limitative results in the sciences—unless we are able to compute the noncomputable. If our argument is valid, then the Gödel incompleteness theorems are not mathematical limitative results in the neurosciences, in molecular biology, in systems biology, and in synthetic biology.

4 A fundamental logical problem for all arguments that claim the Gödel incompleteness theorems are limitative results in the sciences

Call any method (whether a computational procedure, formal system of reasoning, scientific method, ...) weak if it does not confer upon its results mathematical certainty. There are two different possibilities. A system of reasoning of any kind can confer upon its results less than mathematical certainty. We take proofs in mathematics that are not proved within a formal system to do just that. Call mathematical certainty the epistemic modality of a proof in a formal system. The second possibility is that there are proof relations in alternative systems of epistemic modality other than that of mathematical certainty. Thus, weak methods could confer upon its proofs either less than mathematical certainty or whatever is the epistemic modality of the proof relation of some alternative to mathematical certainty.

The idea of weak methods is somewhat new and has met with resistance by both philosophers of mathematics and mathematicians. Today, there are important results in mathematics that come from experimental mathematics, from physics, and from proofs within probabilistic systems of reasoning. In the early 1970s, Georg Kreisel [25] noted it does not logically follow from the fact that a formal system subject to the second Gödel incompleteness theorem cannot, with mathematical certainty, prove its own consistency,

that there are absolutely no means available to prove its consistency. It only follows logically that its consistency cannot be mathematically demonstrated with mathematical certainty in the system. It is left open that consistency is proved by other means, viz., mathematically with less than mathematical certainty (perhaps by statistical reasoning) or nonmathematically, in some epistemic modality other than that of mathematical certainty, by abstract philosophical reasoning (that is, by a priori reasoning that is not encodable into a formal system). In either case, one would have circumvented the Gödel incompleteness theorems since they make claims only about what can be proved with mathematical certainty.

At the time Kreisel's paper appeared—1972—the arsenal of methods not susceptible to the Gödel incompleteness theorems that were available to demonstrate the consistency of a formal system strong enough to express truths of arithmetic were virtually unknown. Kreisel mentions statistical methods for mathematically demonstrating the consistency of Peano arithmetic (that is, $\text{CON}(\text{PA})$) with less than mathematical certainty. He was pessimistic about the prospects of statistical methods being up to this task. “Closer inspection shows that we have in fact very little experience of establishing $\text{CON}(\text{PA})$ by inductive methods and thus we have little knowledge of the *statistical principles* proper to evaluating the hypothetical inductive evidence. More specifically, there are certainly no statistical studies to make sure that the evidence supports the whole of PM or ZF rather than only subsystems! At present the neoformalist position is a sham.” [25]

Kreisel also entertains the possibility of proving $\text{CON}(\text{PA})$ by an abstract, but non-mathematical interpretation. That is, there is an interpretation mapping into some non-mathematical system in which all of the theorems of PA are true. He gives us an example from intuitionism. “An analogue... is provided by the intuitionistic position which identifies *mathematical* with *intuitionistically acceptable* and regards set-theoretic concepts as *metaphysical*: so it leaves open the possibility of establishing $\text{CON}(\text{PA})$ by means of metaphysical nonmathematical interpretations.” [25]

Suppose one takes the view that there are no weak methods that can circumvent the Gödel incompleteness theorems. That is, there are either no mathematical methods for proving a Gödel sentence with less than mathematical certainty or no methods in the proof relation of an alternative epistemic modality for proving a Gödel sentence in the proof relation bearing that epistemic modality. If there are no such weak methods, then whatever limitations the Gödel theorems impose on what we can know within the neurosciences are absolute. We simply cannot evade them. Moreover, whatever limitations the Gödel theorems impose on what cellular or neural structures can compute are absolute. What we will argue in the following section is that demonstrating this point of view exacts a price that is too high to be paid by any finitary mind. That is, in order to reliably prove that there are no weak methods that can circumvent the Gödel incompleteness theorems, one would have to solve a problem of logical complexity Π_2^1 . That is, one would have to compute the noncomputable. However, no being with a finitary mind has the capacity to solve such a problem. Let us call anyone who claims that there are no weak methods for circumventing the Gödel incompleteness theorems as limitative results in the neurosciences LIM (for: the Gödel incompleteness theorems are limitative results in the neurosciences).

4.1 The logical complexity of the problem

The possibilities to which Kreisel alludes for proving $\text{CON}(\text{PA})$ with less than mathematical certainty or in the proof relation of some other epistemic modality must be taken seriously

by LIM. Failure to take them into account is an error in such arguments. That it is an error presupposes that demonstrations of our claims in these sciences do not require mathematical certainty. If a requirement for success (that is, demonstrative adequacy) in cognitive science, the neurosciences, molecular, systems, and synthetic biology is that all of their claims be demonstrated with mathematical certainty, then, of course, the Gödel theorems will show that success is logically impossible to achieve (provided it can also be shown that neither cellular components nor biological systems built from cellular components nor human minds can construct infinite proof trees). That is, if the standard of adequate demonstration in these sciences is mathematical certainty, then the Gödel incompleteness theorems are genuine limitative results for these sciences.

But why should we opt for such a criterion of adequate demonstration in these sciences? If we opt for a less stringent criterion of adequate demonstration, then there are possibilities that show success can be achieved. That is, under less stringent criteria of adequate demonstration in these sciences, the Gödel incompleteness theorems are *not* limitative results for them.

We will not argue here that the neurosciences do not need standards of adequate demonstration that require mathematical certainty. We will now show that if LIM tries to avoid this error by taking into account the possibility that there are weak methods, she will confront a computationally daunting task that cannot be solved by finitary beings (or finitary computing processes). We call the task “DISJUNCTION”. The structure of DISJUNCTION is discussed below.

LIM must show either:

1. All weak methods for mathematically proving, with less than mathematical certainty, or nonmathematically proving, in some other epistemic modality, the correctness of either the ultimate computer program (call it “*C*”) that completely describes human cognition or the correctness of a theory in the neurosciences or the correctness of a neurobiological computation are, in fact, Gödel susceptible

or

2. If that cannot be done because there are Gödel-insusceptible methods available for proving $\text{CON}(C)$, show that the proofs delivered by those methods are not epistemically warranted.

Thus, a disturbing dilemma is in store for LIM:

- 1 horn: LIM must show, for each possible method capable of demonstrating the correctness of *C* with less than mathematical certainty or in some other epistemic modality, that it is either Gödel susceptible or that, where it is Gödel insusceptible, it is epistemologically inadequate.
- 2 horn: If she does not enumerate all of these possibilities, she commits a logical error in her argument.

We will now argue that DISJUNCTION has logical complexity Π_2^1 . A rough classification of mathematical problems is in terms of computability: a problem is either computable or noncomputable. Computable problems can be solved by agents possessing only finitely many computational resources, but noncomputable problems require an infinite supply of computational resources to solve. Finitary agents cannot, in the absence of an oracle, solve (that is, compute solutions to) noncomputable problems. Among noncomputable problems, some are harder to compute solutions to than others. Significant work

in recursion theory has resulted in articulation of hierarchies of recursive unsolvability in which a noncomputable problem is assigned a degree of unsolvability.

What does it mean to say that a problem has logical complexity Π_2^1 ? First of all, it means that the problem cannot be solved if an agent possesses only finitely many computational resources. Secondly, the problem has the quantificational structure of a universal quantifier over predicates followed by an existential quantifier over objects.

Where LIM argues that there are no weak methods, she must be able to compute an infinitary computational task. If she possesses infinitely many resources, she will be able to complete the task. If not, then not. But if she does not complete the task, then her argument commits a logical error. Thus, LIM must either have the capacity to make infinitary computations or else she commits a logical error. But, human beings do not have infinitary computational capacities.

4.2 Proof that DISJUNCTION has logical complexity Π_2^1

4.2.1 The first disjunct in DISJUNCTION

How many different methods are there for proving $\text{CON}(C)$ with less than mathematical certainty or in some other epistemic modality? Suppose that we consider statistical methods for proving $\text{CON}(C)$ with less than mathematical certainty. Among a host of statistical methods are Carnapian measure functions [26]. Since Carnapian inductive logics employ a caution parameter that has infinitely many values and which differentiates different logics, there are infinitely many different methods that can be used to prove $\text{CON}(C)$ with less than mathematical certainty. This just scratches the surface of methods, though. How many probabilistic logics are there? How many hybrid modal probabilistic logics? We could easily continue adding to the list of such methods.

Thus far, we have the following problem for LIM: look at each method for proving $\text{CON}(C)$ with less than mathematical certainty or in some other epistemic modality. Show it is Gödel susceptible. If there are infinitely many applicable methods, each of them must be enumerated and checked for being Gödel susceptible. We have shown, using Carnapian measure functions, that there are infinitely many such methods. But, there is an additional wrinkle. Suppose there is a program N_0 that can be used to mathematically prove $\text{CON}(M_i)$ with less than mathematical certainty or in some other epistemic modality, where M_i is a method for proving $\text{CON}(C)$ with less than mathematical certainty (or in some other epistemic modality) and which has been shown to be Gödel susceptible. LIM needs to verify that N_0 is also Gödel susceptible. But even if it is, there remains the possibility there is a program N_1 that can be used to mathematically prove $\text{CON}(N_0)$ with less than mathematical certainty or in some other epistemic modality. Suppose that N_1 is shown as Gödel susceptible. If so, that there is a possibility there is a program N_2 that can be used to mathematically prove $\text{CON}(N_1)$ with less than mathematical certainty or in some other epistemic modality. So, we have an infinite regress for each program or method M_i that we have shown to be Gödel susceptible. That is, the problem LIM now confronts is the following: look at each method M_i for proving $\text{CON}(C)$ with less than mathematical certainty or in some other epistemic modality. Try to prove that it is Gödel susceptible. If it is, look at each method $N_{<i,0>}$ for proving $\text{CON}(M_i)$ with less than mathematical certainty or in some other epistemic modality. Try to prove that it is Gödel susceptible. If it is, look at each method $N_{<i,1>}$ for proving $\text{CON}(N_{<i,0>})$ with less than mathematical certainty or in some other epistemic modality. Try to prove that it is Gödel susceptible. Continue in this

way until the index i in the ordered index pair $\langle i, j \rangle$ is infinite or until $N_{\langle i, j \rangle}$ is not limited by the Gödel incompleteness theorems.

Let us consider an objection that LIM might raise to the specter of the infinite regress. LIM tells us that there will be no infinite regress because of her dialectical situation in such arguments. Whenever those who argue against LIM propose a method N , all LIM has to do is to prove that N is Gödel susceptible. LIM plays a waiting game. LIM waits for anti-LIM to propose a method, and only then does LIM need to show that the proposed method is Gödel susceptible.

This objection fails because it is dead wrong about the dialectical situation of LIM. LIM does not play a wait and see game with anti-LIM. All LIM arguments are responsible to certain epistemic standards: if there are any possibilities that undermine those arguments, they must be examined. If it is possible that there is a Gödel-insusceptible method or program that proves $\text{CON}(C)$ with less than mathematical certainty or in some other epistemic modality, then that undermining possibility must be examined by LIM.

The methodological problem for LIM is that she must prove a negative existence claim: There are no Gödel-insusceptible means by which $\text{CON}(C)$ can be shown correct with less than mathematical certainty or in some other epistemic modality. Since there might be infinitely many possibilities for proving $\text{CON}(C)$ with less than mathematical certainty or in some other epistemic modality, each of them must be taken into account. If not, then the negative existence claim fails. LIM's claims are not epistemically justified if LIM does not prove the negative existence claim.

4.2.2 Even more logical complexity

Suppose that C is so long it cannot be surveyed by any human agent, whether they are computationally describable or not. Although this is a distinct possibility where C is the program for human cognition, it is also a distinct possibility where C is the correct description of the human brain at the neurobiological level.

If that is the case, we will not know if there are any programs or methods that can be used to prove $\text{CON}(C)$ with less than mathematical certainty or in some other epistemic modality since we will not know what C consists in. In such a situation, LIM cannot rule out the possibility that there are ways of proving C with less than mathematical certainty or in the proof relation of some alternative epistemic modality. LIM could not, under such circumstances, maintain the limitative thesis. But, in fairness to LIM, there might be ways of compressing the length of C so that it can then be determined if there are methods that can be used to prove $\text{CON}(C)$ with less than mathematical certainty or in the proof relation of some alternative epistemic modality. One obvious way of doing this is to reduce C to some program C^* that is humanly surveyable. (One then looks at methods for proving $\text{CON}(C^*)$ with less than mathematical certainty or in some other epistemic modality.) There are three ways in which this can be done. One method is by a relative interpretation of C in C^* , another is by a translation of C into C^* , and the third is a reduction of C to C^* . There are logical differences between interpretations, translations, and reductions, which are the subject of reductive proof theory. What is common to all three is that the map from C into C^* is recursive and preserves negation. The latter condition ensures that logical consistency is preserved under the map.

The maps between C and C^* preserve consistency, provided that C^* is consistent. Since the assumption is that C is consistent, we need to find a short and consistent C^* . Suppose C^* is not short. It is possible there is a C^{**} that is consistent and short to which C^* can be reduced or translated or into which it can be interpreted. At each level of reduction for

which there is a consistent and nonshort $C^{(*n)}$, it is possible that reduction to the next level is by either a translation, reduction, or interpretation.

To avoid the logical error committed by Penrose [23], LIM will certainly have to consider the possibility that C is infeasibly long and then to consider how it might be compressed. The possibility of infinite chain reductions of length ω is a prospect that cannot be a priori ruled out. (The chain length could be ω since a reduction might not decrease the length of $C^{(*n)}$.) There are also other methods that can compress C , that are not the methods of reductive proof theory. For instance, C could be translated into another programming language in which compression devices called MACROS are available, or other higher-order programming constructs that facilitate program compression. There are infinitely many different programming systems, so there are infinitely many possibilities that might need to be examined in the search for a compressed C that is short enough to be humanly surveyable. There are also speed-up theorems in the theory of computability, which tell us that there is no recursive bound on the speed-up of some programs (over the initial program for which there is speed-up).

LIM can object to the preceding infinite regress generated by program compression considerations in the same way LIM objected to the first infinite regress above: anti-LIM must first present to LIM a short C . Once that is done, LIM can then see if there are methods or programs that prove $\text{CON}(C)$ with less than mathematical certainty or in some alternative epistemic modality. Once again, LIM badly misconceives of her epistemic situation in the dialectic of limitative results in the neurosciences. If it is possible that there is a short C , then LIM must examine the possibilities under which it can be obtained. Many of these possibilities (such as relative interpretability) might be dead-ends, might generate infinite regresses, or might create trade-off problems. But, LIM will not know that until a complete search has been made of all of the possibilities.

4.2.3 A three-dimensional infinite network of possibilities

We first noted that there might be infinitely many distinct methods for proving $\text{CON}(C)$ with less than mathematical certainty or in some other epistemic modality. For each such method, LIM must show it is Gödel susceptible. We then noted that for each method M_i that proves $\text{CON}(C)$ and is shown Gödel susceptible, there might be a method N_j that proves $\text{CON}(M_i)$. If so, LIM must show it is Gödel susceptible. But for each N_j that is shown Gödel susceptible, there might be an $N_{\langle i, j \rangle}$ that proves N_j 's correctness, for which it must be shown it is Gödel susceptible. After that, we saw that if C or any M_i or any N_j or any $N_{\langle i, j \rangle}$ is infeasibly long, we need to compress it to obtain a short C or short M_i or short N_j or short $N_{\langle i, j \rangle}$ and that LIM must show, of each of these methods, that it is Gödel susceptible.

Thus, there are infinitely many methods that might prove $\text{CON}(C)$ with less than mathematical certainty or in the proof relation of some other epistemic modality. For each of those methods M_i , if it is shown to be Gödel susceptible, we then look for a method or program that will prove $\text{CON}(M_i)$ with less than mathematical certainty or in the proof relation of some other epistemic modality. Call the method that proves $\text{CON}(M_i)$ N_j . If N_j is shown to be Gödel susceptible, then there might be an $N_{\langle i, j \rangle}$ that proves $\text{CON}(N_j)$ and which must then be shown to be Gödel susceptible. Obviously, for each of the infinitely many M_i 's, there are infinitely many $N_{\langle i, j \rangle}$'s. Finally, for every M_i, N_j , and $N_{\langle i, j \rangle}$, it is possible that each is infeasibly long, and so we must look for a compression of it into a short program. But for each M_i, N_j , and $N_{\langle i, j \rangle}$, there might be an infinite sequence of compression reductions R_i .

The three-dimensional infinite network of possibilities has the logical form: for each element in the $\{C, M_i\}$ infinite sequence and for each element in the $(M_i, N_j, N_{<i,j>})$ set of infinite sequences (attached to each M_i), there is an infinitely long element (R_i) in the infinite set of sequences $(M_i, N_j, N_{<i,j>}, R_i)F(x, y, z)$, (where F is a recursive predicate of the proof relation in question).

The effect of the infinite sequence hanging off each element is that the variable in the range of the existential quantifier (following the universal quantifier over the M_i, N_j and $N_{<i,j>}$ sequences) is not over unit elements, but over infinitely long elements. In which case, we need either an infinitary logic or else function quantifiers binding second-order variables (that is, the x, y, z in the predicate F). If we choose the latter, then the logical complexity of the problem jumps from Π_2^0 to Π_2^1 . This is a serious jump, since the recursive unsolvability of a Π_2^1 problem is much greater than the recursive unsolvability of a Π_2^0 problem.

4.3 The second disjunct in DISJUNCTION

Recall the second disjunct in DISJUNCTION: if a method or program for proving $\text{CON}(C)$ with less than mathematical certainty or in the proof relation of some other epistemic modality is shown to be Gödel insusceptible, then LIM must show that the proofs delivered by that method or program are not epistemically warranted. For each method or program examined by the procedure described in the first disjunct of DISJUNCTION that is found to be Gödel insusceptible, LIM must show it to be epistemically unwarranted.

Of course, there might well be other ways to prove $\text{CON}(C)$, which are not in the proof relation that confers less than mathematical certainty nor or in the proof relation of another epistemic modality. If those ways are Gödel susceptible (whatever that would mean, given that the connection to a proof relation in any kind of modality is surrendered), the claim LIM makes remains intact. If any of those ways are Gödel insusceptible, they *prima facie* refute LIM's claim. The only way to save it is for LIM to show the Gödel insusceptible methods are not epistemically warranted. (That is, all of the proofs delivered by those methods are not epistemically warranted.)

Since any method or program might turn out to be Gödel insusceptible, then any point in the three-dimensional infinite network of possibilities might need to be tested for epistemic adequacy. Of course, no point in the network might need to be tested, if every point represents a method or program that is Gödel susceptible. The question, however, is how we can show that a method or program is epistemically unwarranted without begging the question against anti-LIM (that is, begging it in favor of LIM). We do not want to choose a definition of epistemic warrant that unfairly favors LIM. Suppose that we choose a definition of epistemic warrant under which only methods that confer mathematical certainty on its deliverances are epistemically warranted. In that case, we have begged the question in favor of LIM in choosing a definition of epistemic warrant.

Of course, there are obvious lower bounds for probabilities used to score epistemic warrant claims. If what is proved by a method has a 50% chance of being true, we can conclude it is not epistemically warranted. However, what do we say when the probability of being true is greater than 0.5? What is the cut-off point? What if we do not have sufficient statistics for showing the likelihood of what a method proves? What epistemological theory do we employ in assessing the epistemic adequacy of a method? Even if we are guided by statistical methods used in the sciences, those methods still make philosophical presuppositions about the nature of probabilities. What does it mean to say we search the

space of epistemologies for various construals of epistemic warrant? Notice that all LIM arguments, although they attempt to show mathematical limitative results, quickly progress to requiring philosophical arguments to establish their limitative claim. It easily follows that any elucidation of the notion “epistemic justification of C ”, must be philosophically respectable and that ostensibly “mathematically pure” limitative results in the sciences will entail making philosophical claims.

These issues about the choice of a definition of epistemic warrant are indeed critical problems for LIM. LIM must show any Gödel insusceptible method in the three-dimensional network of possibilities is epistemically unwarranted. Without establishing that, LIM arguments unravel. We can put the point more strongly: anyone who wishes to use LIM must be prepared to decide what counts as epistemic justification of the correctness of C . The point is, anyone who wishes to use the Gödel incompleteness theorems to establish limitative results in the neurosciences (or in molecular biology, system biology, or synthetic biology) must be prepared to decide what counts as the proper standard of epistemic justification of in those sciences without begging the question.

Elucidating the epistemic warrant of methods that prove $\text{CON}(C)$ with less than mathematical certainty or in the proof relation of another epistemic modality is a necessary condition for LIM arguments. When a method in the three-dimensional network of possibilities is Gödel insusceptible, we must look at whether it is epistemically unwarranted. If a method is Gödel susceptible, we need to assess whether proofs of $\text{CON}(C)$ with less than mathematical certainty or in the proof relation of another epistemic modality are epistemically warranted. An important philosophical project, then, for the neurosciences is elucidation of the notion “epistemic warrant of proofs of $\text{CON}(C)$ with less than mathematical certainty or in the proof relation of another epistemic modality”.

5 Using the Gödel incompleteness theorems to strengthen three philosophical limitative arguments—the Explanatory Gap, Mysterian, and Zombie arguments

One of the goals of the cognitive neurosciences is to provide a neurophysiological explanation of consciousness. The elusive quarry is the neural substrate of consciousness. Philosophers have examined various issues concerning consciousness: what is its nature and phenomenology and whether it is reducible to some wholly physical state such as a neurophysiological state. One argument—the Explanatory Gap argument [27]—claims we cannot explain how consciousness arises from neurophysiology. Another argument—the Mysterian argument [28]—is that we will never have the right concepts for understanding consciousness and, thus, how consciousness either does arise or does not arise from the physical. A third argument—the Zombie argument [29, 30]—is that consciousness is nonphysical and so cannot be described by any science. These three limitative arguments are philosophical. The first two are epistemic: they purport to show that we will never have the right arsenal of concepts to either understand or to explain how consciousness can arise from a purely physical substrate. The Zombie argument is metaphysical: it purports to establish that consciousness is not wholly physical and thus falls outside of the province of the neurosciences since its treatment is not amenable to any of the sciences. (That is not to say that, if this philosophical argument is correct, that there is not a mathematical description of consciousness since even if consciousness is nonphysical, it still is possible that it has a mathematical description.)

Each of these arguments can be strengthened by employing the Gödel incompleteness theorems. Since Explanatory Gap and Mysterian arguments against the view that consciousness is reducible to a neurophysiological substrate are epistemic arguments, they do not rule out that consciousness is identical with brain states, but only that the means for showing how this is done cannot, in principle, be achieved. Here is how the Gödel incompleteness theorems can be used to strengthen these two arguments. Take first an Explanatory Gap argument. This is the view that we will never have an adequate scientific explanation of how consciousness arises from wholly physical items. If it turns out that the logical requirements of such an explanatory scheme are constrained by the Gödel theorems, then we have a reason for why the explanation will never be available. Similarly, if it turns out that the concepts necessary for understanding how consciousness either is reducible or is not reducible to wholly physical items are constrained by the Gödel theorems, we then we have a reason for why it is that the concepts are beyond human comprehension. In both cases, the application of the Gödel theorems replaces various philosophical assumptions that have been argued about in the literature and thus constitutes a strengthening of those arguments.

The Zombie Argument—advanced by David Chalmers [29, 30]—argues that if it is possible for there to be an exact physical duplicate of myself that lacks consciousness, then consciousness is not wholly physical. Here is how the Gödel incompleteness theorems can be used to strengthen the Zombie argument. Suppose showing that an exact physical duplicate of me must also be conscious involves proving mathematical theorems, which are constrained by the Gödel incompleteness theorems. If so, then we will not be able to prove that my physical duplicate must be conscious and thus not be able to show that it is not possible for my physical duplicate not to be conscious. If so, this leaves open the possibility that my physical duplicate is not conscious. If we could show that it is logically impossible for my physical duplicate not to be conscious, then we would close the door on it being possible that it is not conscious.

6 Conclusion

The Gödel incompleteness theorems provide mathematical limitative results in the neurosciences (and in philosophical arguments about the nature of consciousness), molecular biology, systems biology, and synthetic biology only if a criterion of demonstrative success in those sciences is that the individual claims be made with mathematical certainty. I think such a demand is epistemically unwarranted and, thus, that the Gödel incompleteness theorems pose no mathematical limits on the success of these sciences. The same holds for philosophical arguments about the nature of consciousness that are strengthened by application of the Gödel incompleteness theorems. The Gödel theorems can only be used in these arguments if the demand is that the claims be made with mathematical certainty. If that demand is relaxed, then the theorems cannot purchase a foothold in those arguments.

References

1. Perez Velazquez, J.L.: Brain, behavior and mathematics: are we using the right approaches? *Physica D* **212**, 161–182 (2005). This paper provides an encyclopediac treatment of the literature on mathematical methods in the neurosciences.

2. Earman, J., Glymour, C., Norton, J.: *Introduction to the Philosophy of Science: A Text by Members of the Department of History and Philosophy of Science at the University of Pittsburgh*. Prentice Hall (1992)
3. Glymour, C.: *Thinking Things Through*. MIT Press (1992)
4. Glymour, C., Scheines, R., Spirtes, P.: *Causation, Prediction and Search*, 2nd revised edn. MIT Press, Cambridge (2001)
5. Landweber, L.F., Kari, L.: The evolution of cellular computing: nature's solution to a computational problem. *BioSystems* **52**, 3–13 (1999) This paper presents a fascinating example of a biological computation performed in ciliates (which are unicellular eukaryotes) that was done artificially (by L. M. Adleman in 1994, ushering in the birth of DNA computing) before its discovery in ciliates
6. Kochen, S., Specker, E.P.: The problem of hidden variables in quantum mechanics. *Journal of Mathematics and Mechanics* **17**, 59–87 (1967)
7. Chermiak, C.: Component placement optimization in the brain. *J. Neurosci.* **14**(4), 2418–2427 (1994). This paper contains important—and remarkable—work on resource-boundedness limitative results in the neurosciences
8. Chermiak, C., Mokhtazadeh, Z., Rodriguez-Esteban, R., Changizi, K.: Global optimization of cerebral cortex layout. *PNAS* **101**(4), 1081–1086 (2004)
9. Chermiak, C.: The bounded brain: toward quantitative neuroanatomy. *J. Cogn. Neurosci.* **2**(1), 58–68 (1990). In this paper, Chermiak shows how unrealistic assumptions about the computational resources of the human brain led to incorrect estimates of such brain structures as cortical sheet area. These unrealistic assumptions are not limitations on what we can, in principle, know in the neurosciences (unless it is also shown that the assumptions are innate and impossible to refute—hardly likely to be the case, however)
10. Paris, J., Harrington, L.: A mathematical incompleteness in Peano arithmetic. In: Barwise, J. (ed.) *Handbook of Mathematical Logic*, pp. 1133–1142. North-Holland Publishing Company, Amsterdam (1977)
11. Kripke, S.: *Lectures on Fulfillability and its Use in Proving Independence Results in Peano Arithmetic*, given at Princeton University. Spring, (1979). Notes of the lectures were taken by N. Salmon (University of California, Santa Barbara) and an edited version of these notes will be published by Oxford University Press in the series Publications of the Saul Kripke Center.
12. Gödel, K.: On formally undecidable propositions of Principia mathematica and related systems I. In: *Collected Works. Publications 1929–1936*, vol. 1. Oxford University Press (1986)
13. Chaitin, G.: Gödel's theorem and information. *Int. J. Theor. Phys.* **21**, 941–954 (1982)
14. Friedman, H.: Boolean relation theory and the incompleteness phenomena, manuscript of forthcoming book (available on Friedman's webpage)
15. Friedman, H.: Limitations on our understanding of the behavior of simplified physical systems, p. 36. 30 October 2008, (available on Friedman's webpage)
16. Russell, B., Whitehead, A.N.: *Principia Mathematica*, vols. 1, 2, 3. Cambridge University Press, Cambridge, (1910–1913)
17. Newman, M.E.J., Strogatz, S.H., Watts, D.J.: Random graphs with arbitrary degree distributions and their applications. *Phys. Rev. E* **64**, 026118 (2001)
18. Lamora-Lopez, G.: Structural characterization of networks using the cat cortex as an example. In: Graben, P., Zhou, C., Thiel, M., Kurths, J. (eds.) *Lectures in Supercomputational Neuroscience: Dynamics in Complex Brain Networks*, pp. 77–106. Springer, Heidelberg (2008)
19. Friedman, H.: My forty years on his shoulders. Gödel Centenary Lecture, p. 55. 29 April 2007 (available on Friedman's webpage)
20. Buechner, J.: Gödel, Putnam and Functionalism. MIT Press, (2008), for a discussion of the structure of the arguments—pro and con—using the Gödel theorems in psychology and the philosophy of mind
21. Gödel, K.: Some basic theorems on the foundations of mathematics and their implications. In: *Collected Works. Unpublished Essays and Lectures*, vol. 3. Oxford University Press (1995)
22. Lucas, J.R.: Minds, machines and Gödel. *Philosophy* **36**, 112–127 (1961)
23. Putnam, H.: Review of Roger Penrose's *Shadows of the Mind*. *Bull. Am. Math. Soc.* **32**, 370–373 (1995)
24. Penrose, R.: *Shadows of the Mind*. Oxford University Press, UK (1994)
25. Kreisel, G.: Which number theoretic problems can be solved in recursive progressions on $\pi(1,1)$ paths through $O^?$ *J. Symb. Log.* **37**(2), 311–334 (1972)
26. Carnap, R.: *Studies in inductive logic and probability*, vols. 1 and 2. Jeffrey, R. (ed.) University of California Press, Berkeley (1971 and 1980)
27. Levine, J.: Materialism and qualia: the explanatory gap. *Pac. Philos. Q.* **64**, 354–361 (1983)
28. McGinn, C.: Problems in philosophy: the limits of inquiry. In: *Chapter Two: Consciousness*. Blackwell, New York (1993)

29. Chalmers, D.: *The Conscious Mind: In Search of a Fundamental Theory*. Oxford University Press, (1996), Part II, *The Irreducibility of Consciousness*. I omit details of the argument, but do note that there are in the physics literature no cloning theorems that are relevant to assessing Chalmers' argument (where a clone is an exact physical duplicate, not to be confused with the notion of a clone that is used in biology)
30. Chalmers, D.: Materialism and the metaphysics of modality. *Philos. Phenomenol. Res.* **59**, 473–493 (1999)