PENULTIMATE DRAFT – PLEASE CITE THE PUBLISHED VERSION

To appear in: *A Critical Reflection on Automated Science*, Bertolaso, M., Sterpetti, F. (eds.), Springer.

Mathematical Proofs and Scientific Discovery

Fabio Sterpetti

Sapienza University of Rome

Department of Philosophy

fabio.sterpetti@uniroma1.it

Abstract

The idea that science can be automated is so deeply related to the view that the method of mathematics is the axiomatic method, that confuting the claim that mathematical knowledge can be extended by means of the axiomatic method is almost equivalent to confuting the claim that science can be automated. I argue that the axiomatic view is inadequate as a view of the method of mathematics and that the analytic view is to be preferred. But, if the method of mathematics and natural sciences is the analytic method, then the advancement of knowledge cannot be mechanized, since non-deductive reasoning plays a crucial role in the analytic method, and non-deductive reasoning cannot be fully mechanized.

Keywords

Automated Discovery; Analytic Method; Axiomatic Method; Formal Proofs; Gödel's Disjunction; Mathematical Knowledge

1. The Method of Mathematics and the Automation of Science

This chapter aims to disentangle some issues surrounding the claim that scientific discovery can be automated thanks to the development of Machine Learning, or some other programming strategy, and the increasing availability of Big Data (see e.g. Allen 2001; Colton 2002; Anderson 2008; King et al. 2009; Sparkes et al. 2010; Mazzocchi 2015). For example, according to Anderson, the "new availability of huge amounts of data, along with the statistical tools to crunch these numbers, offers a whole new way of understanding the world" (Anderson 2008). In this perspective, we "can throw the numbers into the biggest computing clusters the world has ever seen and let statistical algorithms find pattern where science cannot" (Ibidem). Sparkes and co-authors state that the advent of computers and computer science "in the mid-20th century made practical the idea of automating aspects of scientific discovery, and now computing is playing an increasingly prominent role in the scientific discovery process" (Sparkes et al. 2010). This increasing trend to automation in science has led to the development of so-called robot scientist, which allegedly, "automatically originates hypotheses to explain observations, devises experiments to test these hypotheses, physically runs the experiments by using laboratory robotics, interprets the results, and then repeats the cycle" (King at al. 2004). Some authors have gone, if possible, even further by claiming that the process of theory formation in pure mathematics can be automated. For example, Colton writes that theory formation in mathematics "involves, amongst other things, inventing concepts, performing calculations, making conjectures, proving theorems and finding counterexamples to false conjectures," and that computer programs "have been written which automate all of these activities" (Colton 2002, p. 1).

Unlike other researches that deal with automated discovery, I will not focus here on assessing some specific and recent achievements in computer science to verify whether computer machines were really able to autonomously make the scientific discoveries that are credited to them. Rather, I wish to underline how the very idea that scientific discovery can be automated derives, at least in part, from a confusion on the nature of the method of mathematics and science, and how this confusion affects the judgement over whether machines provide genuine discoveries. The origin of this confusion may be due to the fact that the mathematical tools developed to deal with computability originated in the first half of the XX century, in what can be called a formalist and axiomatic cultural 'environment'. From a formalist point of view, the method of mathematics is the axiomatic method, according to which, to demonstrate a statement one starts from some given premises, which are supposed to be true, and then deduces the statement from them. Hilbert, whose ideas were very influential at the time, viewed the axiomatic method as the crucial tool for mathematics and scientific inquiry more in general (Rathjen, Sieg 2018). For example, Hilbert writes that he believes that everything that can "be object of scientific thinking in general [...] runs into the axiomatic method and thereby indirectly to mathematics. Forging ahead towards the ever deeper layers of axioms [...] we attain ever deepening insights into the essence of scientific thinking itself" (Hilbert 1970, p.12). The mathematical agenda of the first decades of the XX century was set by Hilbert by fixing, in his famous address to the International Congress of Mathematicians held in Paris in 1900, the most relevant mathematical problems, and by putting forward his formalist program (Zach 2016). Hilbert's formalist program pursued two main goals: (1) the formalization of all of mathematics in axiomatic form, together with (2) a proof that this axiomatization of mathematics is consistent. In this view, the consistency proof itself was to be carried out using only finitary methods, in order to avoid any reference to infinitary methods, which were regarded as problematic because of their being related to a problematic and disputed concept such as that of infinity.

The theoretical success achieved by logicians and mathematicians in dealing with computability led someone to think that the view on the nature of mathematical knowledge implied by the axiomatic method was in some way vindicated by such extraordinary achievements. But the assessment of some results in computability theory and the inquiry on what is the most adequate way to describe what we do when we do mathematics are distinct issues, and they should be kept distinct (Longo 2011). And whether we keep those issues distinct affects how we assess the claim that science can be automated.

The way in which the confusion between mathematical successful developments in computability theory and the justification of axiomatic ideas originated, can be summarized as follows: the axiomatic view led to the rigorous study of computability. The study of computability led to the development of computer machines. Computer machines proved extraordinary useful and effective. This confirmed that computability theory was correct. And this led many to conclude that the axiomatic view, from which computability theory seemed to stem, should had been the correct view on what we do when we do mathematics, namely deducing consequences from given axioms. Now, mathematics has always been regarded as one of the highest achievements of human thought. So, the argument goes, if doing mathematics means to deduce consequences from given axioms, and machines are able to deduce consequences from given axioms thanks to the mathematical results developed by assuming that doing mathematics must be conceived of as deducing consequences from given axioms, this means both (1) that machines are able to do mathematics, and so that machines can in a sense think, and (2) that doing mathematics really means to deduce consequences from given axioms, i.e. that humans actually do mathematics by deducing consequences from given axioms. Now, if humans actually do mathematics by deducing consequences from given axioms, then human brains can be regarded as Turing Machines, at least in the sense that what we do when we think is, in the ultimate analysis, a kind of computation. Since mathematics and science are subsets of our thinking, this means that mathematics and science are computational activities. Indeed, a Turing Machine is a mathematical model of computation that defines an abstract machine which is capable of simulating any algorithm's behaviour. Since computational activities can be performed by computer machines, there is no compelling theoretical reason why machines might not be able to do mathematics and science. Nor is there any *fundamental* difference between Turing Machines and our brains. This is just a sketch of how the so-called Computational Theory of Mind (CTM)[1] is usually argued for. In this article, for the sake of convenience, I will refer to those who are broadly sympathetic to this line of reasoning and the idea that science can be automated as to 'the computationalists'.

It might be objected that those limitative results, such as Gödel's and Turing's, which paved the way to the development of computer machines, showed also that Hilbert's formalist program was unfeasible, and that this should had been enough to refuse the axiomatic view of mathematics. Indeed, according to Gödel's first incompleteness theorem, any sufficiently strong, consistent formal system *F* is incomplete, since there are statements of the language of *F* which are undecidable, i.e. they can neither be proved nor disproved in *F*. This theorem proves that the first one of the two main goals of Hilbert's program, i.e. the formalization of all of mathematics in axiomatic form, is unattainable. Moreover, according to Gödel's second incompleteness theorem, for any sufficiently strong, consistent formal system *F*, the consistency of *F* cannot be proved in *F* itself. This theorem proves that the second one of the two main goals of Hilbert's program, i.e. to give a proof that the axiomatization of all of mathematics is consistent, is unattainable.[2] Finally, Turing showed the relation between undecidability and incomputability, since it proved that there is no algorithm for deciding the truth of statements in Peano arithmetic. So, this objection is in a sense correct: limitative results showed that Hilbert's program was unfeasible. But it has not to be overlooked that the 'mathematical machinery' developed by mathematicians such as Gödel and Turing to adequately deal with Hilbert's problems and program was developed *within* a shared axiomatic perspective on mathematics and was perfectly suited to meet Hilbert's standard of rigor and formalization (Longo 2003). So, despite those results showed that Hilbert's program cannot be pursued, since they proved extremely useful and were developed within what can be called a Hilbertian theoretical framework, they contributed to perpetuate Hilbert's view of mathematics, according to which the method of mathematics is the axiomatic method.

Still today, many mathematicians and philosophers think that despite Hilbert's program cannot be entirely realized, the ideas conveyed by such program are appealing and not

---

1 On CTM, see Rescorla 2017.

2 For a survey on Gödel's theorems, see Raatikainen 2018.

completely out of track. In this view, the impact of limitative results such as Gödel's and Turing's on the formalist approach to mathematics has not to be overstated. For example, Calude and Thompson recently wrote that although it is not possible "to formalise all mathematics, it is feasible to formalise essentially all the mathematics that 'anyone uses'. Zermelo-Fraenkel set theory combined with first-order logic gives a satisfactory and generally accepted formalism for essentially all current mathematics" (Calude, Thompson 2016, p. 139). In their view, Hilbert program was not completely refuted by limitative results, rather it "can be and was salvaged by changing its goals slightly:" although it is not possible "to prove completeness for systems at least as powerful as Peano arithmetic," it is nevertheless "feasible to prove completeness for many weaker but interesting systems, for example, first-order logic [...], Kleene algebras and the algebra of regular events and various logics used in computer science" (Ibidem).[3] Examples of this kind testify that the axiomatic view, despite its inadequacy, is still the received view in the philosophy of mathematics.

1.1. *The Analytic View of the Method of Mathematics*

I argue that since the axiomatic view is inadequate, we should prefer an alternative view, namely the analytic view of the method of mathematics (Cellucci 2013, 2017), according to which knowledge is increased through the analytic method. According to the analytic method, to solve a problem one looks for "some hypothesis that is a sufficient condition for solving it. The hypothesis is obtained from the problem […], by some non-deductive

[3] This approach seems to forget that Hilbert did not seek formalization for the sake of formalization. Formalization was not an end, rather it was a means in Hilbert's view. His main aim was to give a secure foundation to mathematics through the formalization of a part of it. And this goal cannot be reached because of Gödel's results. So, it is difficult to overstate the relevance of those results for Hilbert's view. That some limited portion of mathematics can be formalized or shown to be complete, but it is not possible to formalize the all of mathematics or prove its completeness in general, it is not something that can salvage Hilbert's perspective "by changing its goals slightly," rather it is a complete defeat of Hilbert's view, since it shows the unfeasibility of its *main* goal. For example, according to Weyl, the relevance of Gödel's results cannot be overstated, since because of those results the "ultimate foundations and the ultimate meaning of mathematics remain an open problem [...]. The undecisive outcome of Hilbert's bold enterprise cannot fail to affect the philosophical interpretation" (Weyl 1949, p. 219).

rule, and must be plausible […]. But the hypothesis is in its turn a problem that must be solved," and is solved in the same way (Cellucci 2013, p. 55).[4]

The assessment of the *plausibility* of any given hypothesis is crucial in this perspective. But how has plausibility to be understood? The interesting suggestion made by the analytic view is that the plausibility of a hypothesis is assessed by a careful examination of the arguments (or reasons) for and against it. According to this view, in order to judge over the plausibility of a hypothesis, the following 'plausibility test procedure' has to be performed: (1) "deduce conclusions from the hypothesis"; (2) "compare the conclusions with each other, in order to see that the hypothesis does not lead to contradictions"; (3) "compare the conclusions with other hypotheses already known to be plausible, and with results of observations or experiments, in order to see that the arguments for the hypothesis are stronger than those against it on the basis of experience" (Ibidem, p. 56). If a hypothesis passes the plausibility test procedure, it can be temporarily accepted. If, on the contrary, a hypothesis does not pass the plausibility test, it is put on a 'waiting list', since new data may always emerge, and a discarded hypothesis may successively be re-evaluated. Thus, according to the analytic view of method, what in the ultimate analysis we really do in the process of knowledge ampliation, is to produce hypotheses, assess the arguments/reasons for and against each hypothesis, and provisionally accept or refuse such hypotheses.

In the last century, the dominance of a foundationalist perspective on scientific and mathematical knowledge, the influence of Hilbert's thought, and the diffusion of the idea that a logic of discovery cannot exist, led to the widespread conviction that the method of mathematics and science is (or should be) the axiomatic method.[5] Analysis, i.e. the search for new hypotheses by means of which problems can be solved, has been overlooked or neglected (Schickore 2014). Philosophers rejected the goal of "traditional epistemology from Plato to Boole: a theory of discovery" (Glymour 1991, p. 75). Indeed, since Plato and Aristotle philosophers "thought the goal of philosophy, among other goals, was to provide methods for coming to have knowledge" (Ibidem). But in the XX century, "there was in

---

[4] The origin of the analytic method may be traced back to the works of the mathematician Hippocrates of Chios and the physician Hippocrates of Cos, and was firstly explicitly formulated by Plato in *Meno*, *Phaedo* and the *Republic*. Here I can only give a sketch of the analytic method. For an extensive presentation of the analytic view, see Cellucci 2017; 2013.

[5] For a survey of the main conceptions of method that have been put forward so far, see Cellucci 2017; 2013. On the analytic method, see also Hintikka, Remes 1974, and Lakatos 1978, Vol. 2, Chap. 5. On the axiomatic method, see also Rodin 2014, part I.

philosophy almost nothing more of methods of discovery. A tradition that joined together much of the classical philosophical literature simply vanished" (Ibidem). For example, Popper famously stated that "there is no such thing as a logical method of having new ideas, or a logical reconstruction of this process" (Popper 2005, p. 8).

Contrary to this perspective, the analytic view maintains that there is a logic of discovery, and that one of the goals of philosophy is to provide methods for coming to have knowledge. Indeed, the analytic method "is a logical method" and from the fact that "knowledge is the result of solving problems by the analytic method, it follows that logic provides means to acquire knowledge" (Cellucci 2013, p. 284). Since logic is a branch of philosophy, this means that philosophy does provide methods for coming to have knowledge. This also means that according to the analytic view, logic has not to be understood as an exclusively deductive enterprise. If logic is understood as an exclusively deductive enterprise, then there cannot be a logic of discovery, since knowledge ampliation requires non-deductive reasoning. Indeed, deductive rules are usually regarded as non-ampliative, because the conclusion is contained in the premises, while non-deductive rules are usually regarded as ampliative, because the conclusion is not contained in the premises.[6] However, that there cannot be a deductive logic of discovery does not mean that there cannot be any logic of discovery. Indeed, logic does not need to be an exclusively deductive enterprise.[7] According to the analytic view, there can be a logic of discovery, but such logic cannot be exclusively deductive.

1.2. *The Analytic Method as a Heuristic Method*

That according to the analytic view there can be a logic of discovery, but such logic cannot be exclusively deductive, implies that the analytic method is not an algorithmic

---

[6] The claim that in deduction the conclusion is contained in the premises has to be understood as meaning that the conclusion either is literally a part of the premises or implies nothing that is not already implied by the premises. The claim that deduction is non-ampliative has been disputed by some philosophers. For example, Dummett famously objects that, if deductive rules were non-ampliative, then, "as soon as we had acknowledged the truth of the axioms of a mathematical theory, we should thereby know all the theorems. Obviously, this is nonsense" (Dummett 1991, p. 195). On this issue, which cannot be treated here for reason of space, and for a possible rejoinder to Dummett's objection, see Cellucci 2017, Sect. 12.7, and Sterpetti 2018, Sect. 6.

[7] This view is controversial. For a defense of the claim that a logic of discovery has to be deductive, see e.g. Jantzen 2015.

method. Methods can indeed be divided into algorithmic and heuristic (Newell, Shaw, Simon 1957). An algorithmic method "is a method that guarantees to always produce a correct solution to a problem. Conversely, a heuristic method is a method that does not guarantee to always produce a correct solution to a problem" (Cellucci 2017, p. 142). Algorithmic methods are closely associated with deductive reasoning and can be mechanized. Heuristic methods are instead closely associated with ampliative reasoning and cannot be mechanized. Algorithmic methods are regarded as closely associated with deductive reasoning, because algorithms are deeply related to computation, and computation can in a sense be regarded as a special form of deduction. For example, Kripke states that "computation is a deductive argument from a finite number of instructions," namely a "special form of mathematical argument" where one "is given a set of instructions, and the steps in the computation are supposed to follow – follow deductively – from the instructions as given. *So a computation is just another mathematical deduction, albeit one of a very specialized form*" (Kripke 2013, p. 80). Although it is not *really* possible to equate computation with deduction in a strict sense, since despite deductions can be regarded as isomorphic to computable functions (see below, Section 2), there are computable functions that are not isomorphic to any deduction, and so it cannot be claimed that *every* algorithmic method is deductive in character and, therefore, every algorithmic method is non-ampliative, there is in any case a close relation between the non-ampliativity of deductive reasoning and the mechanizability of algorithmic methods on the one hand, and the ampliativity of non-deductive reasoning and the non-mechanizability of heuristic methods on the other hand. Since according to the analytic method, discovery is pursued by forming hypotheses through non-deductive inferences, which are ampliative, the analytic method is a heuristic method, and so it cannot be mechanized.

Here some clarifications are in order. Indeed, both deductive and non-deductive inference rules can be formalized (Cellucci 2013), and so one might expect that both deductive and non-deductive reasoning could be mechanized. But things are more complicated. The crucial point is that there is no algorithmic method, i.e. no mechanizable method, to choose what inference rule to apply to what premise in order to produce the desired conclusion in a given context. For example, Gigerenzer states that although algorithms (i.e. formalized rules) for scientific inferences exist, "there is no 'second-order' algorithm for choosing among them," but, he continues, despite there is no algorithm for choosing among algorithms, "scientists nonetheless do somehow choose, and with considerable success" (Gigerenzer 1990, p. 663). How can scientist do that? They "argue with one another, offer reasons for" their choices, and "sometimes even persuade one another" (Ibidem). In other words, scientists choose which non-deductive inference rule to use in a given context by assessing the plausibility of such choice. This fits well with the analytic view, since which rule to use to find a hypothesis to solve a given problem is a hypothesis in its turn. So, also the process by which such hypothesis is evaluated has to be accounted for in terms of plausibility. And indeed, as Gigerenzer says, scientists provide and assess reasons for and against each hypothesis about which rule to use, i.e. they assess its plausibility. Gigerenzer says also that there is no algorithm for deciding what inference rule to use. This amounts to say that the process by which plausibility is assessed, i.e. the process by which reasons for and against a given hypothesis are evaluated, cannot be reduced to computation (more on this below). And this is true even for deductive reasoning. Indeed, even if a mathematical proof consists exclusively of deductions, there is no algorithm to automatically find out the 'correct' inferential path from a given set of premises to the desired conclusion, namely an algorithm to decide what deductive inference rule to apply to what premise and in what order. So, one should say that, in the strict sense, both deductive and non-deductive reasoning cannot be mechanized (Cellucci 2013).

But, it might nevertheless be objected that in the case of deductive reasoning there is at least an algorithm, i.e. a mechanizable procedure, for enumerating all deductions from given premises. By means of the so-called British Museum algorithm (Newell, Shaw, Simon 1957), one can deduce all possible consequences from a given set of premises. One can apply, for instance, Modus Ponens to all premises and derive all two-steps conclusions that are so derivable. Then, one can apply Modus Ponens to the set of two-steps conclusions previously derived and derive all three-steps conclusions that are so derivable, and so on. Thus, one might be confident that *if* the conclusion one wishes to reach is deductively derivable from a given set of premises, the desired conclusion will sooner or later be derived by such algorithm. So, one might claim that, despite there is no algorithmic method to prove that a given statement *C* can be deduced from a given set of premises *P*, deducing consequences from *P* and checking whether *C* is among such consequences is in principle a mechanizable enterprise. Thus, although deduction is not mechanizable in the strict sense, it can be conceded to computationalists that deduction can in a sense be mechanized, because an algorithm can, at least in principle, be developed to derive all deducible conclusions from given premises.

On the contrary, despite non-deductive inferences rules can be formalized as well as deductive rules, there is no algorithm for enumerating all the consequences that can be derived by means of non-deductive inference rules from a given set of premises. Non-deductive inference rules are ampliative, and it is ampliativity that makes a difference with regard to mechanizability. Indeed, if one has to mechanize an inferential process, the conclusion needs to be uniquely determined by its premises (Cellucci 2011). If it were otherwise, a mechanical procedure would be unable to decide what conclusion should be derived from a given premise. Analogously, when one deals with algorithms, each step needs to be determined by the previous one. In deductive inferences, conclusions are uniquely determined by premises. If, for instance, one applies the rule 'Modus Ponens' to premises '$A \rightarrow B$' and 'A', conclusion cannot but be 'B'. This is one of the reasons why many computer scientists regard computation as deeply related to deduction. Hayes and Kowalski, for example, define computation as 'controlled deduction' (Hayes 1973; Kowalski 1979). Computer programs are indeed algorithms, and algorithms are deeply related to computation. In order to develop algorithms that perform in *predictable* ways, computer programs need to be made of predictable elementary computation steps, i.e. steps whose effects are predictable. For example, Pereira states that a computation is a sequence of "atomic computing actions [...]. Without predictability of the effects of sequences of elementary actions, the art of programming as we know it would be impossible, as would the art of painting if pigments changed their colors randomly under the brush" (Wos et al. 1985, p. 9). If elementary computation steps are made of deductions, programs are predictable, since, as noted, in deduction conclusion is uniquely determined by premises. This explains why algorithmic methods, although they are not necessarily deductive in character, are usually regarded as closely associated with deductive reasoning.

But when one deals with non-deductive inference rules, which are ampliative rules, conclusions are generally not uniquely determined by premises. From the very same premise, different (and possibly incompatible) conclusions can be inferred by means of the same non-deductive inference rule. Consider induction. From the (true) premise 'all emeralds examined so far are green', at least two different and incompatible conclusions can be inductively inferred, namely 'all emeralds are green' and 'all emeralds so far examined are green, but all emeralds that will subsequently be examined will be blue' (Goodman 1983, Chap. 3; Cellucci 2011). The fact that when one deals with non-deductive inferences, conclusions are generally not uniquely determined by premises, makes heuristic methods, which rest on non-deductive inferences, non-mechanizable. Since conclusion is not uniquely determined by premises, in order to decide which is the conclusion that one has to draw from a given set of premises, one has to assess the plausibility of each possible conclusion. As stated, the process by which plausibility is assessed, i.e. the process by which reasons for and against a given hypothesis are evaluated, cannot be reduced to computation (more on this below). Therefore, the process of plausibility assessment cannot be made algorithmic. Since heuristic methods are based on non-deductive inferences, which in their turns are based on the process of plausibility assessment; and since the process of plausibility assessment cannot be made algorithmic, heuristic methods cannot be made algorithmic, i.e. they cannot *really* be mechanized.

Moreover, as research develops, new inference rules might always be added to the set of non-deductive inference rules that we use to find hypotheses, and so new conclusions, which previously were underivable from a given set of premises, might eventually become derivable from that very same set of premises by means of such new rules. Again, there is no algorithm to foresee what non-deductive inference rules will be added. As Bacon states, "the art of discovery may grow with discoveries" (Bacon 1961–1986, I, p. 223).

Since, contrary to deductive inferences, in the case of non-deductive inferences conclusions are not uniquely determined by premises and new rules might always be added, when one is dealing with non-deductive reasoning, there is no algorithm, not even in principle, to derive all possible consequences from given premises and check whether the desired conclusion is among them. Thus, in a sense, it can be said that non-deductive reasoning is non-mechanizable in an even stricter sense than deductive reasoning.[8]

Finally, it must be taken into consideration that deduction is usually regarded as truth-preserving, so that if premises are true, conclusion needs to be true. On the contrary, non-deductive inferences are usually regarded as non-truth-preserving, so that even if premises are true, conclusion might be false. So, when it is said that *there is* an algorithmic method to solve a given problem, it can safely be inferred that there is a solution to that problem and that that solution is 'correct', at least in the sense that it is true, provided that premises are true. On the contrary, when it is said that there is a heuristic method to solve a given problem, it cannot safely be inferred that the solution will be the 'correct' one, since non-deductive inference rules are not truth-preserving, and so conclusion might be false despite premises are true.

It should now be clearer why when it comes to discovery, one usually deals with heuristic methods. As already said, while deduction is non-ampliative, non-deductive inferences are ampliative. Since discovery is related to knowledge ampliation, it is an

[8] On this sort of asymmetry between deductive and non-deductive reasoning with respect to mechanizability, cf. Cellucci 2017, p. 306: "there is no algorithm for discovering hypotheses, and hence for obtaining the solution [of a given mathematical problem] by analysis, while [...] there is an algorithm for enumerating all deductions from given axioms, and hence for obtaining the solution [of that problem] by synthesis."

ampliative enterprise. So, usually it cannot be carried on by means of algorithmic methods, which usually rest on deduction, which is non-ampliative. Rather, discovery must be carried on by means of heuristic methods, since they rest on non-deductive inferences to extend our knowledge. And since non-deductive reasoning is ampliative but it cannot be mechanized, heuristic methods are ampliative, but they cannot be mechanized either.

To recapitulate, algorithmic methods are methods by means of which it is possible to uniquely and automatically determine the solution of a problem. They are closely associated with deductive reasoning, because in deduction conclusion is uniquely determined by premises, and so deduction can be automated. Nevertheless, algorithmic methods need not to be deductive in character. There can well be algorithmic methods that are non-deductive in character. Consider the British Museum algorithm illustrated above. This algorithm is not deductive in character, despite it generates nothing but deductions. Nevertheless, it is able to uniquely and automatically determine the solution of a problem, i.e. it is an algorithmic method. Indeed, the British Museum algorithm can be mechanized, since it rests on deduction in order to solve a problem, and in deduction conclusion is uniquely determined by premises. Deduction can thus be automated. So, the British Museum algorithm, although it is not deductive in character, can be automated as well. Consider also “Baconian” induction, which is a kind of induction in which conclusion can be uniquely determined by given premises, and which can therefore be mechanized. This kind of inference is studied, for instance, by Inductive Logic Programming (Muggleton, De Raedt 1994). Nevertheless, this kind of “mechanical induction is an extremely limited one and in a sense can be reduced to deduction, so it is not really ampliative and hence is generally inadequate for discovery” (Cellucci 2013, p. 160). So, despite it cannot be claimed that algorithmic methods are deductive in character, it is fair to say that usually in non-deductive inferences conclusion is not uniquely determined by premises, and this makes those inferences non-mechanizable. So, those methods that rely on non-deductive inferences to solve a problem are usually non-mechanizable methods, i.e. they are heuristic methods. This is why algorithmic methods are usually closely associated with deductive reasoning, while heuristic methods are usually closely associated with ampliative reasoning.

According to the analytic view, the axiomatic method is inadequate to explain advancement in mathematics and natural sciences precisely because, since it conceives of logic as an exclusively deductive enterprise, and deduction is non-ampliative, the axiomatic method is unable to account for the process of *hypotheses production*.[9] This means that the axiomatic method cannot improve our understanding of how we acquire *new* mathematical knowledge. From an epistemological point of view, it is disappointing that the alleged *method* of mathematics is unable to say anything relevant about *how* new mathematical knowledge is acquired. In the light of the axiomatic method, knowledge ampliation remains a mystery. On the contrary, in the analytic view the path that has been followed to reach a result and solve a problem is not occulted, since in this view the context of discovery is not divorced from the context of justification. For instance, an analytic demonstration consists in a non-deductive derivation of a hypothesis from a problem and possibly other data, where the hypothesis is a sufficient condition for the solution of the problem, and is plausible (Cellucci 2017, Chap. 21).

It is important to underline that the analytic method involves both deductive and non-deductive reasoning. Indeed, to find a hypothesis we proceed from the problem by performing ampliative inferences, and then in order to assess the plausibility of such hypothesis we deduce conclusions from it. But the role that deduction plays in the analytic view is not the exclusive role that deduction is supposed to play in the axiomatic view. According to the analytic view, axioms are not the source of mathematical knowledge, and we shouldn't overestimate their role, which is limited to give us the possibility of presenting, for didactic or rhetorical purposes, some body of already acquired knowledge in deductive form. Axioms do not enjoy any temporal or conceptual priority in the development of mathematical knowledge, nor do they play any special epistemological role. As Hamming states, if "the Pythagorean theorem were found to not follow from postulates, we would again search for a way to alter the postulates until it was true. Euclid's postulates came from the Pythagorean theorem, not the other way" (Hamming 1980, p. 87).

Finally, it is worth noting that the concept of plausibility has not to be confused with the concept of probability. As Kant points out, "plausibility is concerned with whether, in the cognition, there are more grounds for the thing than against it" (Kant 1992, p. 331), while probability measures the relation between the winning cases and possible cases. Plausibility involves a comparison between the arguments for and the arguments against, so it is not a mathematical concept. Conversely, probability is a mathematical concept (see Cellucci 2013, Sect. 4.4). It may be objected that, although probability and plausibility

[9] This view can, at least in part, be traced back to Lakatos 1976, where Lakatos, by relying on the work of Pólya, strongly criticized the occultation of the heuristic steps that are crucial to the development of mathematics.

appear to be distinct concepts, we may account for plausibility-based considerations in terms of probability, because plausibility obeys the law of probability (Pólya 1941). But this objection is inadequate. To see that plausibility is not equivalent to probability, consider that, since probability is "a fraction whose numerator is the number of favorable cases and whose denominator is the number of all the cases possible" (Laplace 1951, p. 7), in order to effectively calculate the probability of a given hypothesis *h*, we have to know the denominator, i.e. the number of all the cases possible. But in many cases, we do not know (and perhaps we cannot even know) the number of all the cases possible. Thus, if plausibility were to be understood in terms of probability, we should not be able to evaluate the plausibility of all those hypotheses for which we are unable to determine in advance the set of all the possible rival alternatives. But, scientist routinely evaluate the plausibility of that kind of hypotheses, so it cannot be the case that probability is equivalent to plausibility.[10]

That plausibility is not a mathematical concept and cannot be reduced to probability is crucial. Indeed, this implies that plausibility assessment cannot be reduced to computation and so made algorithmic. Thus, computer machines cannot perform plausibility assessment. Since plausibility assessment is central to the analytic method, that computer machines cannot perform plausibility assessment means that *if* the analytic method is the method by which scientific discovery is pursued, contrary to what computationalists claim, scientific discovery cannot be automated.

1.3. *The Analytic View and the Automation of Science*

There are three main arguments against the claim that science can be automated. The first one is that since there is no logic of discovery, machines cannot be programmed to make genuine discoveries. To face this argument, computationalists should prove that there is in fact an algorithmic method for discovery. Many computationalists attempt to do this, i.e. to provide evidences that programs do make scientific discoveries (see e.g. King et al. 2009; Colton 2002). But to prove that machines made *genuine* scientific discoveries that are able to extend our knowledge is not an easy task, since, as this chapter tries to clarify, knowledge cannot *really* be extended by exclusively computational means.[11] The second

---

[10] For an extensive treatment of this issue, see Cellucci 2013, Sect. 20.4; Sterpetti, Bertolaso 2018.

[11] I cannot analyse here the discoveries allegedly made by computer programs. This is a topic for future research. Briefly, the main questions one has to address when dealing with this issue are: (1) whether hypotheses are really produced by programs, since often either a set of hypotheses

argument against the claim that science can be automated is that, since human minds outstrip Turing Machines, machines cannot equate human computational performances, and so discovery will remain a human enterprise. To face this argument, computationalists need to show that some of its assumptions is unjustified. In Section 3, I deal with this argument and illustrate how difficult might be to defend some of its assumptions. The third argument against the claim that science can be automated is that there is a logic of discovery, but this logic is not exclusively deductive. According to this line of reasoning, the method of mathematics and science is the analytic method, and this implies that machines cannot implement it, because it is not an algorithmic method, i.e. it is not mechanizable. Thus, neither mathematics nor natural sciences can be automated. My claim is that this last argument is the most difficult one for the computationalists to refute, since it represents the most serious theoretical objection to the claim that science can be automated. Indeed, if one maintains the axiomatic view, as the supporters of the first two arguments against the claim that science can be automated usually do, it is likely that computationalists will find a way to defend their position, because the axiomatic method is deductive in character, and thus every instance of it can be mechanized, at least in principle. So, if one adopts the axiomatic view, one can hardly find out a truly compelling theoretical reason to support the claim that science cannot be automated.

2. Proofs and Programs

To allow my argument to take off the ground, it is important to clarify the connection between computer science and mathematics, better the connection between the claim that scientific discovery can be automated and the idea that mathematics is theorem proving.[12] Indeed, those who think that the method of mathematics is the axiomatic method, usually also think that mathematics is theorem proving (Cellucci 2017, Chap. 20). According to them, mathematicians start from a set of axioms and deduce theorems from them. On the

or a set of heuristic strategies to routinely produce hypotheses from given inputs are already present in the so-called background knowledge of programs (see e.g. Marcus 2018); (2) whether programs can only produce results that can be obtained through a merely exploratory search of a well-defined space of possibilities, or they are also able to make innovative discoveries, i.e. discoveries that originate from the formulation of new concepts, i.e. concepts that cannot easily be derived from current ones and modify the very space of possibilities (see e.g. Wiggins 2006).

[12] On whether mathematics is theorem proving or problem solving, see Cellucci 2017, Chap. 20.

contrary, according to those who think that mathematics is problem solving, mathematicians analyze a problem and then try to infer, from the problem and other relevant available knowledge, a hypothesis that is able to solve it. As Mäenpää states, with his *Grundlagen der Geometrie*, Hilbert "reduced geometry to theorem proving" (Mäenpää 1997, p. 210). Then, "Hilbert's model has spread throughout mathematics" in the XX century, "reducing it to theorem proving. Problem solving, which was the primary concern of Greek mathematicians, has been ruled out" (Ibidem). I argue that mathematics is not theorem proving, because this view is unable to account for how mathematical knowledge is ampliated. For example, when Cantor demonstrated that to every transfinite cardinal there exist still greater cardinals, "he did not deduce this result from truths already known […] because it could not be demonstrated within the bounds of traditional mathematics. Demonstrating it required formulating new concepts and new hypotheses about them" (Cellucci 2017, p. 310). So, mathematical knowledge cannot *really* be extended by exclusively deductive means. Mathematics is thus better conceived of as problem solving, since this view allows us to appreciate the role played by ampliative reasoning in finding hypotheses to solve problems. But if mathematics is not theorem proving, the axiomatic view is inadequate as a view of the method of mathematics, and so it cannot be claimed that science can be automated. To see why this may be the case, we need firstly to consider the relation between computer programs and mathematical proofs. Indeed, by the Curry-Howard isomorphism, we know that there is a deep relation between computer programs and mathematical proofs.

The Curry-Howard isomorphism establishes a correspondence between systems of formal logic as encountered in proof theory and computational calculi as found in type theory (Sørensen, Urzyczyn 2006, p. v; Prawitz 2008). Proof theory is focused on formal proof systems. It was developed in order to turn "the concept of specifically mathematical proof itself into an object of investigation" (Hilbert 1970, p. 12). The λ-calculus was originally proposed as a foundation of mathematics around 1930 by Church and Curry, but it was a "somewhat obscure formalism until the 1960s," when its "relation to programming languages was [...] clarified" (Alama, Korbmacher 2018, Sect. 3). The λ-calculus is a model of computation. It was introduced few years before another model of computation was introduced, namely Turing Machines. In the latter model, "computation is expressed by reading from and writing to a tape, and performing actions depending on the contents of the tape" (Sørensen, Urzyczyn 2006, p. 1). Turing Machines resemble "programs in imperative programming languages, like Java or C" (Ibidem). In contrast, in λ-calculus one is concerned with functions, and "these may both take other functions as arguments, and return functions as results. In programming terms, λ-calculus is an extremely simple higher-order, functional programming language" (Ibidem). So, with the invention of computer machines, the λ-calculus proved to be a useful tool in designing and implementing programming languages. For instance, the λ-calculus can be regarded as an idealized sublanguage of some programming languages like LISP.

Proof theory and the λ-calculus finally met, since Curry and Howard realized that the programming language was a logic, and that the logic was a programming language. In 1934, Curry firstly observed that every type of a mathematical function ($A \rightarrow B$) can be read as a logical proposition ($A \supset B$), and that under this reading the type of any given function always corresponds to a provable proposition. Conversely, for every provable proposition there is a function with the corresponding type (Curry 1934). In later years, Curry extended the correspondence to include terms and proofs (Wadler 2015). In 1969, Howard circulated a manuscript, unpublished until 1980, where he showed that an analogous correspondence obtains between Gentzen's natural deduction and simply-typed λ-calculus (Howard 1980). In that paper, Howard made also explicit the deepest level of the correspondence between logic and programs, namely that simplification of proofs corresponds to evaluation of programs. This means that for each way to simplify a proof there is a corresponding way to evaluate a program, and vice versa (Wadler 2015).

The Curry-Howard isomorphism is based on the so-called 'propositions-as-sets' principle. In this perspective, a proposition is thought of as its set of proofs. Truth of a proposition corresponds to the non-emptiness of the set. To illustrate this point, we will stick to an example made by Dybjer and Palmgren (2016). Consider a set $E_{m,n}$, depending on m, $n \in \mathbb{N}$, which is defined by:

$$E_{m,n} = \begin{cases} \{0\} & if\ m = n \\ \emptyset & if\ m \neq n \end{cases}$$

$E_{m,n}$ is nonempty when $m = n$. The set $E_{m,n}$ corresponds to the proposition $m = n$, and the number 0 is a proof-object inhabiting the sets $E_{m,n}$. Consider now the proposition: *m is an even number*, expressed as the formula $\exists n \in \mathbb{N}.\ m = 2n$. A set of proof-objects can be built which corresponds to this formula by using the general set-theoretic sum operation. Suppose that A_n ($n \in \mathbb{N}$) is a family of sets. Then its disjoint sum is given by the set of pairs

$$(\Sigma n \in \mathbb{N})\ A_n = \{(n, a) : n \in \mathbb{N}, a \in A_n\}.$$

If we apply this construction to the family $A_n = E_{m,2n}$ we can see that $(\Sigma n \in \mathbb{N})E_{m,2}$ is nonempty when there is an $n \in \mathbb{N}$ with $m = 2n$. By using the general set-theoretic product operation $(\Pi n \in \mathbb{N})A_n$ we can similarly obtain a set corresponding to a universally quantified proposition.

So, in this context proofs of $A \supset B$ are understood as "functions from (proofs of) A to (proofs of) B and $A \supset B$ itself the set of such functions" (von Plato 2018, Sect. 6). If we take, for instance, "$f : A \supset B$ and $a : A$, then functional application gives $f(a) : B$. The reverse, corresponding to the introduction of an implication, is captured by the principle of functional abstraction of [...] Church's λ-calculus" (Ibidem). As von Plato states, the Curry-Howard isomorphism made intuitionistic natural deduction crucial to computer science. Indeed, the Curry-Howard isomorphism "gives a computational semantics for intuitionistic logic in which computations, and the executions of programs more generally, are effected through normalization" (Ibidem). A proof of an implication $A \supset B$, for instance, is a "program that converts data of type A into an output of type B. The construction of an object (proof, function, program) f of the type $A \supset B$ ends with an abstraction" (Ibidem). When an object a of type A is put into f as an argument, the resulting expression is not normal, "but has a form that corresponds to an introduction followed by an elimination. Normalization now is the same as the execution of the program f" (Ibidem).

For the aim of this chapter, the relevance of the Curry-Howard isomorphism lies in that it shows that computer programs are strictly equivalent to formalized mathematical proofs. Indeed, we have that for each proof of a given proposition, there is a program of the corresponding type, and vice versa. But the correspondence is even deeper, in that for each way to simplify a proof there is a corresponding way to evaluate a program, and vice versa. This means that "we have not merely a shallow bijection between propositions and types, but a true isomorphism preserving the deep structure of proofs and programs" (Wadler 2015, p. 75). In other words, we can understand programs as proofs and proofs as programs.

Why the fact that programs are proofs is relevant to the discussion of the claim that science can be automated? Because to claim that science can be automated amounts to claim that computer machines are able to contribute to knowledge ampliation. Now, if one wishes to claim that machines are able to contribute to knowledge ampliation, and programs are equivalent to mathematical proofs, and mathematical proofs are chains of deductions from given axioms, then one needs to claim that mathematical knowledge can be ampliated by means of deductions from given axioms, or by means that can be shown to be equivalent to deduction. And this amounts to claim that the method by which mathematical knowledge is ampliated is the axiomatic method. More generally, to claim that machines are able to contribute to knowledge ampliation amounts to claim that the process of knowledge ampliation can be entirely reduced to computation, i.e. that it is possible to account for the process of knowledge ampliation in exclusively computational terms. Indeed, if according to the axiomatic view, mathematical proofs are crucial to the ampliation of mathematical knowledge, and mathematical proofs are equivalent to programs, so that proofs can be mechanized, the process of knowledge ampliation can be accounted for in computational terms, i.e. it can be reduced to computation. In other words, if one wishes to claim that machines are able to contribute to knowledge ampliation, one commits oneself to the claim that the axiomatic view is an adequate view of how knowledge is extended. So, the claim that scientific knowledge can be extended by computer machines is equivalent to the claim that the method of mathematics is the axiomatic method and mathematical knowledge can be extended by means of that method. This implies that confuting the claim that mathematical knowledge can be extended by the axiomatic method is equivalent to confuting the claim that science can be automated. This is why I focus here on the issue of the method of mathematics in order to assess the claim that scientific discovery can be automated.

3. Mathematical Knowledge

In order to support the claim that, since the analytic method is the method of mathematics, neither mathematics nor natural sciences can be automated, in this section I illustrate one of the main reasons that can be provided to show that the axiomatic view is inadequate as a view of the method of mathematics.

According to the traditional view of mathematics, mathematical knowledge is acquired by exclusively deductive means, namely by deductive proofs from previously acquired mathematical truths. For example, Prawitz states that “mathematics [...] is essentially a deductive science, which is to say that it is by deductive proofs that mathematical knowledge is obtained” (Prawitz 2014, p. 78). This view gives raise to several problems, but here I focus on the problem of accounting for how we acquired the initial body of mathematical truths from which mathematics originated. More precisely, I argue that since the axiomatic method is unable to account for how we acquired the initial body of mathematical truths from which mathematics originated, the axiomatic view is unable to secure the epistemic superiority of mathematical knowledge over scientific knowledge and provide a secure foundation to mathematics, and that this fact undermines one of the main reasons for why the axiomatic view was so appealing to many philosophers in the first place.

3.1. *Mathematical Starting Points*

The problem of accounting for how we acquired the initial body of mathematical truths from which mathematics originated is deeply related to the issue of whether mathematics is distinct from other sciences. Indeed, the degree of certainty of mathematical knowledge is usually thought to be higher than that of scientific knowledge. Still today, mathematics is regarded as "the paradigm of certain and final knowledge" (Feferman 1998, p. 77) by most mathematicians and philosophers. According to many mathematicians and philosophers, the degree of certainty that mathematics is able to provide is one of its qualifying features. For example, Byers states that the certainty of mathematics is "different from the certainty one finds in other fields [...]. Mathematical truth has [...] [the] quality of inexorability. This is its essence" (Byers 2007, p. 328). The higher degree of certainty and justification displayed by mathematical knowledge is usually supposed to be due to the method of mathematics, which is commonly taken to be the axiomatic method. In this view, the method of mathematics differs from the method of investigation in the natural sciences: whereas "the latter acquire general knowledge using inductive methods, mathematical knowledge appears to be acquired [...] by deduction from basic principles" (Horsten 2015). So, it is the deductive character of mathematical demonstrations that confers its characteristic certainty to mathematical knowledge, since demonstrative "reasoning is safe, beyond controversy, and final" (Pólya 1954, I, p. v), precisely because it is deductive in character.

Now, a deductive proof "yields categorical knowledge [i.e. knowledge which is independent of any particular assumptions] only if it proceeds from a secure starting point and if the rules of inference are truth-preserving" (Baker 2016, Sect. 2.2). Let us concede, for the sake of the argument, that it can safely be asserted that deduction is truth-preserving, and so that if premises are true, conclusion needs to be true.[13] Then, if one embraces the axiomatic view, to prove that it is true that mathematical knowledge displays a higher degree of justification because of its deductive character, one has to prove that the mathematical starting points of mathematical reasoning, i.e. axioms, are known by some means that guarantees a degree of justification higher than the degree of justification

[13] For a defense of the claim that the axiomatic view is inadequate also because there is no non-circular way of proving that deduction is truth-preserving, see Cellucci 2006.

provided by the means by which in the natural sciences the non-mathematical starting points of inductive inferences are known. Otherwise, the claim that mathematics is epistemically superior to natural sciences would be ungrounded. Indeed, according to the axiomatic method, mathematical knowledge is extended by deductions. So, the certainty of mathematical results depends on whether the axioms from which they are derived are known to be true with certainty. Mathematical results derived from axioms can in their turn become starting points for other deductions, and so on. At any stage of this process of ampliation of mathematical knowledge, the newly obtained results will be as certain as the initial starting points, since deduction is truth-preserving. This point has been clearly illustrated by Williamson:

> At any given time, the mathematical community has a body of knowledge, including both theorems and methods of proof. Mathematicians expand mathematical knowledge by recursively applying it to itself [...]. Of course, present mathematical knowledge itself grew out of a smaller body of past mathematical knowledge by the same process. Since present mathematical knowledge is presumably finite, if one traces the process back far enough, one eventually reaches 'first principles' of some sort that did not become mathematical knowledge in that way. (Williamson 2016, p. 243).

The difficult question is: How such 'first principles' became mathematical knowledge? There is no clear and undisputed answer to this question. And yet answering to this question it is crucial, since the epistemic status of mathematical knowledge depends on the epistemic status of those first principles.

The problem is that the axiomatic view is unable to account for *how* such 'first principles' became mathematical knowledge, nor is it able to justify their alleged epistemic superiority. Indeed, if mathematical knowledge is knowledge of the most certain kind, and the method of mathematics is the axiomatic method, in order to claim that knowledge produced by that method is certain, the starting points by which such knowledge is derived have to be known to be true with certainty. In an axiomatic context, this amounts to claim that one can know with certainty that the axioms that constitute our mathematical starting points are consistent. But, it is almost uncontroversial that it is generally impossible to mathematically prove that axioms are consistent, because of Gödel's results. Indeed, as already noted, by Gödel's second incompleteness theorem, for any consistent, sufficiently strong deductive theory *T*, the sentence expressing the consistency of *T* is undemonstrable in *T*. So, according to many philosophers, there must be some *other* way to know with certainty that those axioms which constitute our mathematical starting points are consistent, otherwise the claim that mathematical knowledge displays an epistemic status which is superior to that of scientific knowledge cannot be justified.

This issue deserves a careful examination, since it is related to an important discussion on the philosophical consequences of Gödel's results which took place in the last decades, namely the discussion on whether Gödel's results imply that CTM is untenable (Horsten, Welch 2016; Raatikainen 2005). That discussion is useful to illustrate how even those philosophers who reject the idea that mathematical reasoning can be automated, failed to recognize the inadequacy of the axiomatic view, and so were led astray in their attempts to argue against the computationalist perspective.

3.2. *Gödel's Disjunction*

Gödel regarded it as clear that the incompleteness of mathematics demonstrated that mathematical reasoning cannot be mechanized. In his view, it "is not possible to mechanize mathematical reasoning, i.e., it will never be possible to replace the mathematician by a machine, even if you confine yourself to number-theoretic problems" (Gödel *193?, p. 164). In order to support his view, Gödel famously formulated the so-called Gödel's Disjunction (GD), according to which either the human mathematical mind cannot be captured by an algorithm, or there are absolutely undecidable problems of a certain kind. More precisely, he states that "either [...] the human mind [...] infinitely surpasses the powers of any finite machine, or else there exist absolutely unsolvable diophantine problems" (Gödel 1951, p. 310), where 'absolutely' means that those problems would be undecidable, "not just within some particular axiomatic system, but by any mathematical proof that the human mind can conceive" (Ibidem). In other words, either absolute provability outstrips all forms of relative provability or there are absolutely undecidable sentences of arithmetic. That there are absolutely undecidable sentences implies that the mathematical world "is independent of human reason, insofar as there are mathematical truths that lie outside the scope of human reason" (Koellner 2016, p. 148). So, GD can be rephrased as follows. At least one of these claims must hold: either minds outstrip machines or mathematical truth outstrips human reason.

A crucial notion in Gödel's argument against the claim that minds are equivalent to machines is the notion of 'absolute provability'. Relative provability is mechanizable, since it is provability within some particular axiomatic (i.e. formal) system, and any axiomatic system can be represented by a Turing Machine, i.e. it can be represented by an algorithm. The problem is that by Gödel's first incompleteness theorem, for any sufficiently strong axiomatic system F, there are statements which are undecidable within F, i.e. they can neither be proved nor disproved in *F*. Gödel shows how the issue of the decidability of statements in an axiomatic system *F*, is equivalent to the issue of the solvability of a certain kind of problems in arithmetic, namely diophantine problems. So, the argument goes, if there are absolutely undecidable sentences in mathematics, i.e. sentences that cannot be decided algorithmically, and if the human mind is equivalent to a Turing Machine, i.e. it is able to decide only algorithmically decidable sentences, then there are mathematical truths that cannot be known by human minds (Horsten, Welch 2016).

Why did Gödel focus on diophantine problems? Because in order to support his idea that mathematical reasoning cannot be mechanized, Gödel aimed to show that, despite some portions of mathematics can be completely formalized, the all of mathematics cannot be formalized. He was thus interested in finding the smallest portion of mathematics which cannot be formalized (Gödel *193?), and diophantine problems are among core problems in number theory. Moreover, the tenth problem in the famous list of Hilbert asked for "a procedure which in a finite number of steps could test a given (polynomial) diophantine equation for solvability in integers," and it "is easy to see that it is equivalent to ask for a test for solvability in natural numbers" (Davis 1995, p. 159). Gödel's result on the unsolvability of certain diophantine problems was not able to prove that Hilbert's tenth problem is unsolvable. Nevertheless, the relation that Gödel highlighted between a set being computable and the solvability of a diophantine equation is deeply related to the hypothesis that Hilbert's tenth problem is unsolvable, since to prove that Hilbert's tenth problem is unsolvable amounts to show that "every recursively enumerable [...] relation is diophantine, or, equivalently, that every primitive recursive relation is diophantine" (Ibidem), as Matiyasevič proved in 1970 (Matiyasevič 2003).

To better see the relationship between absolutely unsolvable diophantine problems and the issue of whether human minds are equivalent to Turing Machines, consider that Gödel subscribes to the iterative conception of sets. According to this conception, in order to construct ever larger sets, one begins with the integers and iterate the power-set operation through the finite ordinals. This iteration "is an instance of a general procedure for obtaining sets from a set *A* and well-ordering R" (Boolos 1995, p. 291). Axioms can be formulated to describe the sets formed at various stages of this process, but "as there is no end to the sequence of operations to which this iterative procedure can be applied, there is none to the formation of axioms" (Ibidem). Gödel observes that higher-level set-theoretic axioms will entail the solution of certain problems of inferior level left undecided by the preceding axioms; those problems take a particularly simple form, namely to determine the truth or falsity of some diophantine propositions. Diophantine propositions are sentences of the form:

$$(\forall x_1, \ldots, x_n \in \mathbb{Z})(\exists y_1, \ldots, y_m \in \mathbb{Z}) p(x_1, \ldots, x_n, y_1, \ldots, y_m) = 0$$

where p is a diophantine polynomial, i.e. a polynomial with integer coefficients.

Now, it can be proved that the question of whether a given Turing Machine produces a certain string as an output is equivalent to the question of whether a certain diophantine proposition P is true: the decision problem for diophantine propositions is essentially the decision problem for Turing Machines, under another description (Leach-Krouse 2016). Along this line of reasoning, Gödel proves that Gödel's first theorem "is equivalent to the fact that there exists no finite procedure for the systematic decision of all diophantine problems" (Gödel 1951, p. 308). Thus, the decision problem for diophantine propositions is absolutely unsolvable, i.e. it is impossible to find a mechanical procedure for deciding every diophantine proposition.

Now, recall that according to GD, either the human mind surpasses the powers of any finite machine, or else there exist absolutely unsolvable diophantine problems. We have seen that there are absolutely unsolvable diophantine problems. Does this mean that Gödel endorses the second disjunct of GD? The answer is in the negative.

In Gödel's view, that there is an *algorithmically* unsolvable decision problem for diophantine propositions does not mean that the answer to the question about whether there are diophantine unsolvable problems has to be in the affirmative. According to Gödel, it is true that we have "found a problem which is absolutely unsolvable [...]. But this is not a problem in the form of a question with an answer Yes or No, but rather something similar to squaring the circle with compass and ruler" (Gödel *193?, p. 175). Gödel's idea is that the decision problem for diophantine propositions is undecidable because "the problem of finding a mechanical procedure restricts the types of possible solutions, just as the problem of squaring the circle with compass and straightedge restricts possible solutions" (Leach-Krouse 2016, p. 224). In Gödel's view, mechanical solutions aren't the only intelligible solutions that can be offered to mathematical problems. In other words, Gödel believes that it is possible to decide diophantine propositions in some non-mechanical way. Indeed, Gödel credits Turing with having established beyond any doubt that the recursive functions are exactly the functions that can actually be computed. So, Gödel accepted the Church-Turing thesis. But Gödel understood the thesis as stating that the "recursive functions are exactly the *mechanically* computable functions, *not* the functions computable by a humanly executable method" (Ibidem).

Why Gödel thinks that some mathematical problems are solvable even if there are no mechanical solutions to such problems? One of the reasons is that Gödel thinks "in a somewhat Kantian way that human reason would be fatally irrational if it would ask questions it could not answer (Raatikainen 2005, p. 525). It is because he subscribes to this kind of 'rational optimist' (Shapiro 2016) that Gödel rejects the second disjunct of GD, i.e. that there are absolutely unsolvable problems in mathematics. Gödel shares Hilbert's belief that "for any precisely formulated mathematical question a unique answer can be found" (Gödel *193?, p. 164), and so that there cannot be absolutely unsolvable problems in mathematics. Hilbert was "so firm in this belief that he even thought a mathematical proof could be given for it, at least in the domain of number theory" (Ibidem). A mathematical proof of Hilbert's idea that there aren't absolutely unsolvable problems in mathematics can be understood as the proof of the following statement (H): given "an arbitrary mathematical proposition *A* there exists a proof either for *A* or for not-*A*, where by 'proof' is meant something which starts from evident axioms and proceeds by evident inferences" (Ibidem). But "formulated in this way the problem is not accessible for mathematical treatment because it involves the non-mathematical notion of evidence" (Ibidem), where 'evidence' can be understood as 'justification'. But if one tries to reduce this informal notion of proof to a formal one, so that it can be accessible for mathematical treatment, then it comes out that it is not possible to prove (H), because, as Gödel himself proved, it is impossible to prove that there are not unsolvable mathematical problems, i.e. it is impossible to prove that for any mathematical proposition *A* there exists a proof either for *A* or for not-*A*. According to Gödel, this negative result can have two different meanings: "(1) it may mean that the problem in its original formulation has a negative answer, or (2) it may mean that through the transition from evidence to formalism something was lost" (Ibidem). In his view, it "is easily seen that actually the second is the case" (Ibidem).

Why Gödel thinks that (2) is more plausible than (1)? Because according to him, "the number-theoretic questions which are undecidable in a given formalism are always decidable by evident inferences not expressible in the given formalism" (Ibidem). In other words, undecidability results are a reflection of the inadequacy of our *current* axioms. But new and better axioms can always be produced in order to ask questions left unanswered, because, as noted above, there is no limit to the process of axioms formation. It is for this reason that in Gödel's view, questions which are undecidable in a given formalism are always decidable by evident inferences not expressible in the given formalism. For example, consider Peano Arithmetic (PA). According to Gödel, if one climbs the hierarchy of types, the axiom system for second order arithmetic PA^2 decides the statement left undecided at the lower level, namely the consistency of PA, i.e. *Con*(PA). Now one has to prove *Con*(PA^2). But, for Gödel's second incompleteness theorem, PA^2 does not decide *Con*(PA^2). However, the axiom system for third-order arithmetic PA^3 settles the statement left undecided at the lower level, namely *Con*(PA^2). And so on. Thus, in Gödel's view, questions which are undecidable in a given formalism are always decidable by evident inferences not expressible in the given formalism. This can be done by introducing new and more powerful axioms and by working with such a more powerful formalism.[14] What about the 'evidence' of those new 'evident inferences'? According to Gödel, those new evident inferences "turn out to be exactly as evident as those of the given formalism" (Gödel *193?, p. 164).

But this view is unsatisfactory, since it does not answer the very question we started with: How is it that the axioms of a given formalism are justified? If more powerful axioms can always be formed from inferior axioms, and if those more powerful axioms are as justified as the less powerful axioms from which they are derived, it is crucial to know to what extent initial axioms are justified. It cannot simply be answered that to prove what is not provable in a given formalism, one can always resort to some stronger axioms, which are as justified as the weaker axioms, because unless initial weaker axioms are known to be true with certainty, the fact that stronger axioms are as justified as weaker ones does not tell anything about the degree of certainty to which stronger axioms are justified. To prove something is to provide justification for something. If to prove something in system *A*, one has to rely on axioms of system *A′*, which are stronger than axioms of system *A*, and so on, then a regression is lurking.

3.3. *Intrinsic and Extrinsic Justification*

So, we need to focus again on the following question: How can mathematical starting points be justified? According to Gödel, axioms can be justified either intrinsically or extrinsically. Axioms that are intrinsically justified are those "new axioms which only

14 In fact, things are more complicated. When one climbs the hierarchy of sets, the stronger axioms that become available lead to "more intractable instances of undecidable sentences" (Koellner 2011, Sect. 1). For example, at the third infinite level one can formulate Cantor's Continuum Hypothesis. These instances of independence "are more intractable in that no simple iteration of the hierarchy of types leads to their resolution" (Ibidem). I will not address this issue here.

unfold the content of the [iterative] concept of set" (Gödel 1964, p. 261). On the contrary, axioms are extrinsically justified if even "disregarding the intrinsic necessity of some new axiom, and even in case it has no intrinsic necessity at all, a probable decision about its truth is possible also in another way, namely, inductively by studying its 'success'" (Ibidem), where by 'success' Gödel means 'fruitfulness in consequences' (Koellner 2014, Sect. 1.4.2). Both these accounts of how axioms can be justified are unsatisfactory.

As regard intrinsic justification, according to Gödel by "focusing more sharply on the concepts concerned" (Gödel *1961/?, p. 383) one clarifies the meaning of those concepts. By such procedure, "new axioms, which do not follow by formal logic from those previously established, again and again become evident" (Ibidem, p. 385). In Gödel's view, this explain why minds and machines are not equivalent, since "it is just this becoming evident of more and more new axioms on the basis of the meaning of the primitive notions that a machine cannot imitate" (Ibidem). But it is not easy to determine with certainty whether a new axiom is merely the unfolding of the content of the iterative concept of set, which is supposed to be sufficiently evident and unambiguous to be easily graspable by human minds. For example, against the widespread idea that the more familiar axioms of Zermelo-Fraenkel set theory with the axiom of Choice (ZFC) follow directly from the iterative conception of set, i.e. that they are intrinsically justified, while other stronger axioms, such as e.g. large cardinal axioms, are only supported by extrinsic justification, Maddy writes that even "the most cursory look at the particular axioms of ZFC will reveal that the line between intrinsic and extrinsic justification, vague as it might be, does not fall neatly between ZFC and the rest" (Maddy 1988, p. 483). According to Maddy, that the more familiar axioms of ZFC are "commonly enshrined in the opening pages of mathematics texts should be viewed as an historical accident, not a sign of their privileged epistemological or metaphysical status" (Ibidem).

Besides the difficulty of demarcating intrinsic justification from extrinsic justification, there is also the difficulty of adjudicating between different possible but incompatible intrinsic justifications. For example, Cellucci criticizes Gödel's idea that we can extend our knowledge of the concepts of set theory by focusing more sharply on the concept of set as follows:

> Suppose that, by focusing more sharply on the concept of set Σ, we get an intuition of that concept. Let S be a formal system for set theory, whose axioms this intuition ensures us to be true of Σ. So Σ is a model of S, hence S is consistent. Then, by Gödel's first incompleteness theorem, there is a sentence A of S which is true of Σ but is unprovable in S. Since A is unprovable in S, the formal system S′ = S ∪ {¬A} is consistent, and hence has a model, say

> Σ′. Then ¬A is true of Σ′ and hence A is false of Σ′. Now, Σ and Σ′ are both models of S, but A is true of Σ and false of Σ′, so Σ and Σ′ are not equivalent. Suppose next that, by focusing more sharply on the concept of set Σ′, we get an intuition of this concept. Then we have two different intuitions, one ensuring us that Σ is the concept of set, and the other ensuring us that Σ′ is the concept of set, where the sentence A is true of Σ and false of Σ′. This raises the question: Which of Σ and Σ′ is the genuine concept of set? Gödel's procedure gives no answer to this question. (Cellucci 2017, p. 255).

This scenario cannot be easily dismissed, because Gödel does not require intuition to be infallible (Williamson 2016), so we cannot exclude that we can find ourselves in the situation described above, where we have two different and incompatible intuitions of the concept of set, one ensuring us that Σ is the concept of set, and the other ensuring us that Σ′ is the concept of set. More generally, although it is often claimed that "axioms do not admit further justification since they are *self-evident*" (Koellner 2011, Sect. 1), it is very difficult to neatly distinguish what is self-evident and what is not. Indeed, "there is wide disagreement in the foundations of mathematics as to which statements are self-evident" (Koellner 2014, Sect. 1.4.1).[15] As Hellman and Bell write, contrary to the "popular (mis)conception of mathematics as a cut-and-dried body of universally agreed-on truths and methods, as soon as one examines the foundations of mathematics, one encounters divergences of viewpoint […] that can easily remind one of religious, schismatic controversy" (Bell, Hellman 2006, p. 64).

Moreover, one should also keep in mind that "even such distinguished logicians as Frege, Curry, Church, Quine, Rosser and Martin-Löf have seriously proposed mathematical theories that have later turned out to be inconsistent" (Raatikainen 2005, p. 523). As Davis states, in all those cases insight "didn't help" (Davis 1990, p. 660).

Finally, by Gödel's results we know that we cannot define once and for all a set of axioms, and then try to justify those axioms by claiming that they are self-evident because they are so simple and elementary that they cannot fail to appear as self-evident to anyone. Rather, we know that we will always need to introduce new ever stronger axioms. And those axioms are increasingly less simple than the simple ones we started with. So, even if one concedes, for the sake of the argument, that simplest axioms might appear as self-evident to the majority of mathematicians, the fact that we need to introduce new stronger axioms raises the question of how one is to justify these new axioms, for "as one continues to add stronger and stronger axioms the claim that they are [...] self-evident [...] will grow

[15] On disagreement in mathematics, see Sterpetti 2018.

increasingly more difficult to defend" (Koellner 2011, Sect. 1), since it is usually perceived that the more one moves along the hierarchy of sets, the less axioms are self-evident, and rather they become increasingly disputable.

As regard extrinsic justification, Gödel famously writes that:

> There might exist axioms so abundant in their verifiable consequences, shedding so much light upon a whole field, and yielding such powerful methods for solving problems [...] that, no matter whether or not they are intrinsically necessary, they would have to be accepted at least in the same sense as any well-established physical theory. (Gödel 1964, p. 261).

Now, the problem is that this kind of justification makes mathematical knowledge on a par with scientific knowledge with respect to epistemic justification. In other words, this kind of justification of the axioms is unable to support the claim that mathematical knowledge is epistemically superior to scientific knowledge because it is deductively derived from premises which are certain. For example, Davis states that in this perspective new "axioms are just as problematical as new physical theories, and their eventual acceptance is on similar grounds" (Davis 1990, p. 660). But if mathematical starting points are justified because of their consequences, things do not go in the way predicted by the axiomatic view. Rather, things go the other way around. It is no more the case that conclusions are justified because they are deductively inferred from premises which are certain, rather premises are regarded as justified because it is possible to derive from them consequences which have so far proved interesting or useful. This kind of reasoning is not deductive in character, rather it is inductive, and so it seems inadequate to support the axiomatic view of the method of mathematics and the alleged epistemic superiority of mathematical knowledge over scientific knowledge.

Moreover, extrinsic justification makes mathematics susceptible to those criticisms that are usually reserved to natural sciences. For example, it might be objected that even if consequences derived from some new axiom *A* are useful and interesting, and no contradiction has so far been derived from *A*, one cannot know that things will continue this way. A contradiction may always emerge, unless *A* is known to be true with certainty. If we do not know *A* to be true with certain in advance, how can we know that a contradiction will not be derived from *A* within one hundred years? So, if mathematics is extrinsically justified, i.e. it is inductively justified, the justification of mathematical knowledge is prone to the same challenges to which the justification of scientific knowledge is prone.

3.4. *Lucas' and Penrose's Arguments*

The fact that there is no way to show that the all of mathematics can be derived from some self-evident axioms or by merely elaborating on the concept of set, and so that mathematics can be intrinsically justified, implies that the axiomatic view is unable to support the claim that mathematical knowledge is epistemically superior to scientific knowledge. It also implies that the 'inductive challenge' just illustrated can be moved against those who wish to support the claim that scientific discovery cannot be mechanized because humans outstrip machines, if they maintain the axiomatic view of mathematics.

To see this point, consider Lucas' (1961) and Penrose's (1989) arguments in support of the first disjunct of GD, i.e. the claim that minds outstrip machines, and so mathematics cannot be mechanized. It is worth noting that Gödel did not think that either one of the disjuncts of GD could be established solely by appeal to the incompleteness theorems. He thought instead that the disjunction as a whole, i.e. GD, was a "mathematically established fact" (Gödel 1951, p. 310), and that it was implied by his incompleteness theorems. In contrast, Lucas and Penrose argued that the incompleteness theorems imply the first disjunct of GD, i.e. the claim that minds outstrip machines (Koellner 2016).

According to Lucas (1961), while any recursively enumerable system *F* can never prove its Gödel sentence *G*, i.e. the sentence saying of itself that it is not provable in *F*, human minds can know that *G* is true. In this view, *G* is absolutely provable, although it is not provable in *F*. According to Lucas this shows that minds outstrip machines. However, in order to draw such conclusion, one needs to assume that human minds can know that *F* is consistent, i.e. that it is absolutely provable that *F* is consistent. Indeed, to claim that *G* is absolutely provable amounts to claim that the consistency of the axioms of *F* is absolutely provable. It is worth recalling that we refer here to 'absolute provability', because it is instead algorithmically provable that the consistency of the axioms of *F* is not provable in *F*, because of Gödel's second incompleteness theorem.

To see why for Lucas' argument to work one needs to assume that it is absolutely provable that *F* is consistent, recall that Gödel's second incompleteness theorem has a conditional form (Raatikainen 2005). Indeed, Gödel showed that for any sufficiently strong formal theory *F*, if *F* is consistent, a sentence *G* in the language of *F*, which is equivalent in *F* to the sentence expressing the consistency of *F*, cannot be proved in *F*. Thus, if *F* proves only true sentences, i.e. it is consistent, then *G* cannot be proved in *F*.

But how can one support the claim that it is absolutely provable that *F* is consistent? Penrose, following Lucas, claims that although *G* is unprovable in *F*, we can always 'see' that *G* is true by means of the following argument. If *G* is provable in *F*, then *G* is false, but that is impossible, because our formal system "should not be so badly constructed that it actually allows false propositions to be proved!" (Penrose 1989, p. 140). If it is impossible to prove *G* in *F*, then *G* is unprovable, and therefore it is true.

So, again, in order to 'see' the truth of *G* one has to be able to 'see' the consistency of the axioms of *F*, i.e. that it is not possible to derive contradictions from the axioms of our formal system. As Davis states, in this view if some form of "*insight* is involved, it must be in convincing oneself that the given axioms are indeed consistent, since otherwise we will have no reason to believe that the Gödel sentence [i.e. *G*] is true." (Davis 1990, p. 660). In this line of reasoning, to 'see' the consistency of the axioms of *F* is thus precisely what humans can do, but machines cannot. Thus, humans must be able to see the consistency of axioms in some non-algorithmic way. This means that for Lucas' and Penrose's arguments to work, one should be able account for how it is possible to see the consistency of axioms in some non-algorithmic way. If it cannot be proved that human minds are able to know with certainty that the axioms of *F* are consistent in some non-algorithmic way, Lucas' and Penrose's argument would amount to a merely conditional statement asserting that 'the Gödel sentence of *F*, i.e. *G*, is true, if *F* is consistent'. But this conditional statement is provable in *F*, and therefore to claim that humans can see the truth of such conditional statement is not sufficient for establishing the first disjunct of GD, i.e. the claim that minds outstrip machines.

Now, neither Lucas nor Penrose give us an adequate and detailed account of what it is a non-algorithmic way of proving the consistency of the axioms of *F*. For instance, they do not provide any account of how it is that a human brain can 'see' the consistency of the axioms of *F* which may be compatible with what we know about human cognitive abilities.[16] They simply assume that it is evident that (at least some portion of) mathematics *is* consistent, since no falsities can be derived in it. But, as we already noted, even if one concedes, for the sake of the argument, that there are certain limited formal theories of which the set of provable sentences can be seen to contain no falsehoods, such as e.g. Peano Arithmetic (PA); and even if one concedes, for the sake of the argument, that the Gödel sentence for PA is true and unprovable in PA, this does not amount to concede "that we can see the truth of Gödel sentences for more powerful theories such as ZF set theory, in which almost the whole of mathematics can be represented" (Boolos 1990, p. 655). Such an 'extrapolation' is ungrounded, and neither Lucas nor Penrose give us compelling reason to think that instead we should rely on it. Rather we have reason to be skeptical about such

[16] On this issue, see Sterpetti 2018.

an extrapolation. Indeed, in the absence of an adequate justification of the certainty by which we are supposed to know the consistency of the axioms of *F*, a sort of pessimistic meta-induction over the history of mathematics can be raised, which is analogous to the one raised by Laudan (1981) over the history of science. If, as already noted, distinguished logicians as Frege, Curry, Church, Quine, Rosser and Martin-Löf have proposed mathematical theories that have later turned out to be inconsistent, how can we be sure that our 'seeing' that *F* is consistent is reliable? Consider Zermelo-Fraenkel set theory, ZF. According to Boolos, there is no way to be certain that "we are not in the same situation vis-a-vis ZF that Frege was in with respect to naive set theory [...] before receiving, in June 1902, the famous letter from Russell, showing the derivability in his system of Russell's paradox" (Boolos 1990, p. 655).

It should now be clearer why if one does not have a compelling justification for the claim that the axioms of *F* are consistent, then mathematical knowledge cannot be claimed to be epistemically superior to scientific knowledge. As Boolos says, are "we really so certain that there isn't some million-page derivation of '$0 = 1$' that will be discovered some two hundred years from now? Do we know that we are really better off than Frege in May 1902?" (Ibidem).

3.5. *Lucas's and Penrose's Arguments and the Axiomatic View*

The discussion above shows that those who support the idea that minds are not equivalent to machines along the lines of Lucas and Penrose fail because, even if they wish to show that mathematical reasoning, and scientific discovery more generally, cannot be reduced to mechanical computation, their view of the method of mathematics is basically equivalent to the axiomatic view, according to which to do mathematics is to provide axiomatic proofs, i.e. to deductively prove theorems. This is the reason why they are committed to show that human minds can somehow 'see' something which is uncomputable by machines, i.e. something that cannot be deductively derived by machines. Their formalist and foundationalist approach to knowledge and mathematics leads them to think that in order to show that mathematical knowledge cannot be mechanized, it should be proved that humans can realize the axiomatic ideal that it has been proved machines cannot realize. But they do not really put into question the axiomatic view of mathematics. The problem is that if one claims that the method of mathematics is the axiomatic method, which relies exclusively on deductive inferences, since, as noted above in Section 1, in deduction conclusions are uniquely determined by premises, so that deduction can be made algorithmic and thus mechanized, it is then difficult for one to justify the claim that mathematical knowledge cannot be mechanized. So, one of the reasons why Lucas' and Penrose's arguments fail is that they both remain within the boundary of the traditional view of mathematics, according to which we "begin with self-evident truths, called 'axioms', and proceed by self-evidently valid steps, to all of our mathematical knowledge" (Shapiro 2016, p. 197).

I think that the failure of Lucas' and Penrose's attempts invites us to rethink such a conception of the method of mathematics. Their failure does not show that minds are equivalent to machines, because it cannot *really* be proved that minds make computations that machines cannot make. Rather, their failure suggests that their view of the method of mathematics is inadequate, since it is unable to account for why it is that minds and machines are not equivalent. Since we have good reason to think that it is not so easy to prove that minds are equivalent to machines, if assuming the traditional view of mathematics is an impediment in showing why this is the case, we should reject such view of the method of mathematics and search for another view of mathematics that allows us to adequately account for that fact. In my view, such a different view of mathematics is already available, and it is the analytic view.

What good reason do we have to think that it is not so easy to prove that minds are equivalent to machines? If minds were equivalent to machines, we would expect mathematical knowledge be advanced by deductions from given axioms. But the axiomatic view is challenged by the *actual* development of mathematics. Basically, mathematical knowledge is not advanced exclusively by means of deductions from given axioms. For example, Portoraro states that mathematicians "combine induction heuristics with deductive techniques when attacking a problem. The former helps them guide the proof-finding effort while the latter allows them to close proof gaps" (Portoraro 2019, Section 4.8). Shapiro writes that the standard way to establish new theorems is "to embed them in a much richer structure, and take advantage of the newer, more powerful expressive [...] resources. The spectacular resolution of Fermat's theorem, via elliptical function theory, is not atypical in this respect" (Shapiro 2016, p. 197–198). This 'holistic' epistemology of mathematics is compatible with the analytic view of the method of mathematics.[17] Indeed, in this view, we do not solve problems because we are able to derive the solution from a fixed set of axioms known to be consistent from the start, rather mathematical knowledge is ampliated by formulating new and richer hypotheses, which need not necessarily be new

[17] For an account of the resolution of Fermat's Last Theorem inspired by the analytic view, see Cellucci 2017, Sect. 12.13.

axioms, and which can lead to the solution of the problems we are dealing with. So, it is not the case that complex mathematical problems are solved by starting from simple and consistent axioms. Rather, it is the case that some aspects of mathematics are illuminated by expanding our theoretical resources, i.e. by creating new and more complex mathematics, which can shed light on the mathematical problems that arise at a 'lower' level of complexity and give raise to new problems to be solved.

It is important to stress that if one conceives of how mathematics advances in this way, one can more clearly see why mathematical reasoning cannot be mechanized. Basically, machines are fixed devices, which need to be built to produce desired outputs. In order to do that, one needs to know in advance almost everything about the problem the machine will be asked to solve. On the contrary, when we humans try to extend our knowledge, we do not know in advance what we will have to face, what shape the space of possibilities may have (Sterpetti, Bertolaso 2018). So, it is quite difficult to program a machine to make discoveries in the same way it happens to us humans to make discoveries. Since we do not know in advance how mathematical knowledge will be developed, we do not know how to program a machine that can really advance mathematical knowledge on its own. For example, Shapiro states that since "we do not know, in advance, just what rich theories we will need to prove future theorems about the natural numbers, we are not in a position to say what" the available theoretical resources are if we try to deal with issues such as whether "there are arithmetic sentences unknowable in principle" (Shapiro 2016, p. 198). The point is that while a Turing Machine has to have a fixed alphabet, a single language, and a fixed program for operating with that language, real "mathematicians do not have that. There is no fixed language, no fixed set of expressive resources, and there is no fixed set of axioms, once and for all, that we operate from" (Ibidem).

It is because the axiomatic view is unable to satisfactorily account for how knowledge is ampliated that I think that there are good reasons to think that it is not easy to prove that machines are equivalent to minds. I claimed that the analytic view is instead able to provide a more satisfactory account of how knowledge is ampliated. But the analytic view does that at a cost. According to the analytic view, knowledge is advanced by non-deductive means. Non-deductive inferences rest on plausibility assessment. Plausibility is not a mathematical concept, so plausibility assessment cannot be reduced to computation, i.e. it cannot be mechanized. This implies that discovery cannot be mechanized and that there is no epistemic difference between scientific knowledge and mathematical knowledge. So, according to the analytic view minds are not equivalent to machines independently from whether machines and minds can make the same computations. The axiomatic view and computationalism fall together. And the alleged epistemic superiority of mathematics falls as well. In this perspective, "there is no difference in kind between a mathematical proof, an entrenched scientific thesis, and a well-confirmed working hypothesis. Those are more matters of degree. In principle, nothing is unassailable in principle" (Ibidem).

3.6. *Absolute Provability and the Axiomatic View*

We have seen that it is not easy to claim that minds outstrip machines and that mathematics is epistemically superior to natural sciences, because it is not easy to adopt the axiomatic view *and* defend the claim that axioms are known to be consistent. Someone might try to rescue the axiomatic view and Gödel's and Hilbert's belief that "for any precisely formulated mathematical question a unique answer can be found," by both (1) defending the claim that the all of mathematical knowledge is absolutely provable and (2) weakening the claim that mathematics is epistemically superior to natural sciences. The problem is that this strategy to rescue the axiomatic view leads to the explicit acceptance of the claim that minds are equivalent to machines. So, there is a sort of dilemma here: either one accepts CTM in order to rescue the axiomatic view, or one rejects the axiomatic view in order to defend the claim that minds are not equivalent to machines.

For example, Williamson (2016) elaborates on Gödel's view in order to shed light on the concept of *absolute provability*. He suggests applying the term 'normal mathematical process' to all "those ways in which our mathematical knowledge can grow. Normal mathematical processes include both the recursive self-application of pre-existing mathematical knowledge and the means, whatever they were, by which first principles of [...] mathematics originally became mathematical knowledge" (Williamson 2016, p. 243). In this view, a mathematical hypothesis is "*absolutely provable* if and only if it can in principle be known by a normal mathematical process" (Ibidem). So far, so good. If one follows Williamson's argument, one can reach the conclusion that "every true formula of mathematics is absolutely provable, and every false formula is absolutely refutable" (Ibidem, p. 248). The problem is precisely that there is no satisfactory account of the means by which first principles of mathematics originally became mathematical knowledge. So, Williamson is not able to provide a justification of the claim that mathematical starting points became mathematical knowledge in a way which makes them epistemically superior to scientific knowledge, nor is he interested in providing such a justification, since he rejects the thesis that mathematical knowledge is certain in a way which is intrinsically different from the way in which scientific knowledge is certain. According to Williamson, that all arithmetical truths are absolute provable, "does not imply that human [...] minds can somehow do more than machines" (Ibidem). It merely means that for "every arithmetical truth *A*, it is possible for a finitely minded creature [...] to prove *A*" (Ibidem). But there is nothing in this view that excludes the hypothesis that "for every arithmetical truth *A*, some implementation of a Turing machine can prove *A*" (Ibidem). Rather, in this view "the implementation of the Turing machine is required to come to know *A* by a normal mathematical process, and so to be minded" (Ibidem), precisely because Williamson conceives of mathematical reasoning exclusively in terms of computability. As we noted, if all of mathematical reasoning is accounted for in terms of computational activity, no room is left in this view for plausibility assessment, i.e. for non-deductive inferences and for those heuristic methods that rest on such inferences, since plausibility assessment cannot be reduced to computation. This means that, in this view, the method of mathematics is the axiomatic method, which rests exclusively on deduction. As already noted, deduction can be regarded as reducible to computational activity. This seems to imply that in this view, the axiomatic method is an algorithmic method and can therefore be mechanized. So, in Williamson's view, there is no radical difference between machines and minds. If it is true that we know that no "possible implementation of a Turing machine can prove all arithmetical truths" (Ibidem), we got no reason to support the claim that finitely minded creatures like us can instead prove all arithmetical truths. Thus, minds do not outstrip machines.

Why Williamson rejects the idea that mathematical knowledge is more certain than scientific knowledge? Precisely because there is no satisfactory way to support such a claim of superiority, and so he thinks that such claim should be rejected. The point is that there is no way to convincingly argue for the epistemic superiority of mathematical starting points. According to Williamson, for anyone who objects that his approach is unable to guarantee the mathematical certainty of the new stronger axioms that we need to introduce to prove statements left unsettled by weaker axioms, the "challenge is to explain the nature of this 'mathematical certainty' we are supposed actually to have for the current axioms, but lack for the new one" (Ibidem, p. 251).

To recapitulate, Williamson's view is unable to support the claim that mathematical knowledge is epistemically superior to scientific knowledge, because although Williamson conceives of mathematical reasoning in terms of computability, i.e. almost exclusively in deductive terms, he, unlike Gödel, does not believe that there is a way to prove that mathematical starting points are true which is distinct from the way by which scientific hypotheses are known to be true. And it is because of the supposed computational character of mathematical reasoning that in this view machines and minds are equivalent.

Williamson's view is thus unable to support the claim that machines and minds are not equivalent because, as Lucas' and Penrose's, it endorses the axiomatic view of mathematics, according to which to do mathematics is to make deductions from given axioms. Since deductive reasoning can, at least in a non-strict sense, be mechanized, in this view it cannot be adequately supported the claim that machines are not equivalent to human minds.

3.7. *The Debate on Gödel's Disjunction and the Axiomatic View*

The debate just illustrated can be summarized by describing it in a dilemmatic form: either (1) mathematics is epistemically superior to natural sciences, and so there must be a way to prove the consistency of axioms by some non-mechanical means, and this implies that human minds outstrip machines (Lucas' and Penrose's arguments), so that discovery cannot be automated; or (2) mathematics is not epistemically superior to natural sciences, and so there is no way to prove the consistency of axioms by some non-mechanical means, and this implies that human minds do not outstrip machines (Williamson's argument), so that it cannot be excluded that discovery can be automated.

I wish to stress again that what is taken for granted by both those who argue for (1) and those who argue for (2) is the axiomatic view, according to which mathematical knowledge is advanced by purely *deductive* means. In this perspective, whether minds and machines are equivalent depends on whether they are able to make the same computations. But this means that it has already been accepted that mathematical reasoning can really be reduced to computation.

On the contrary, the analytic view does not have to face this dilemma. Indeed, according to the analytic view, mathematics and other sciences share the very same method, namely the analytic method, so there is no difference between mathematical knowledge and scientific knowledge with respect to epistemic justification. Both mathematics and sciences are advanced in the same way, i.e. by forming hypotheses through non-deductive inferences that are able (at least provisionally) to solve some given problems. In this view, deductive and non-deductive inferences are on a par with respect to their justification. So, the analytic view is able to account for the continuity between mathematics and sciences that it asserts we should expect, while the axiomatic view is not really able to justify its claim that mathematics and sciences display different degrees of epistemic justification. According to Cellucci, "the fact that generally there is no rational way of knowing whether primitive premises" of axiomatic proofs "are true [...] entails that primitive premises of axiomatic proofs [...] have the same status as hypotheses" (Cellucci 2008, p. 12) in both mathematics and natural sciences.

According to the analytic view, that there is no difference in the epistemic status of mathematical and scientific knowledge does not mean that machines are equivalent to minds, since in this view mathematical reasoning cannot be reduced to computation. Rather, in this view, non-deductive reasoning is essential to the ampliation of mathematical knowledge. So, in this view, that machines are not equivalent to minds does not imply that minds perform *computations* that machines cannot perform. In this view, machines are not equivalent to minds because non-deductive reasoning is crucial to the ampliation of knowledge. Since non-deductive reasoning cannot really be mechanized,[18] because it rests on the process of plausibility assessment, which cannot be mechanized, machines are unable to perform non-deductive reasoning. And it is *this* fact that implies that mathematical knowledge cannot be ampliated by machines alone.

4. Conclusions

In this chapter, I showed how the idea that science can be automated is deeply related to the idea that the method of mathematics is the axiomatic method, so that confuting the claim that mathematical knowledge can be extended by the axiomatic method is almost equivalent to confuting the claim that science can be automated. I defended the thesis that, since the axiomatic view is inadequate to account for how mathematical knowledge is extended, the analytic view should be preferred. To do that, I analysed whether the axiomatic view can adequately account for two aspects of its own conception of mathematical knowledge, namely (1) how we acquired the initial body of mathematical truths from which the all of mathematics is supposed to be originated and (2) the alleged epistemic superiority of mathematics over natural sciences. Then, I developed an argument that can be summarized as follows. If the method of mathematics and science is the analytic method, the advancement of knowledge cannot be mechanized, since ampliative reasoning, i.e. non-deductive reasoning, plays a crucial role in the analytic method, and non-deductive reasoning cannot be fully mechanized.

[18] This claim might appear disputable to those who claim that ampliative reasoning is actually performed by machines. As already said, for reason of space, I have to leave the analysis of the claim that machines do autonomously perform ampliative reasoning for future work.

Acknowledgements

I wish to thank Carlo Cellucci for careful reading and commenting on an earlier draft of this chapter.

References

Alama, J., Korbmacher, J. (2018). The Lambda Calculus. In: E.N. Zalta (ed.), *The Stanford Encyclopedia of Philosophy*, URL = <https://plato.stanford.edu/archives/fall2018/entries/lambda-calculus/>.

Allen, J.F. (2001). In Silico Veritas. Data-Mining and Automated Discovery: The Truth Is in There. *EMBO reports*, 2: 542–544.

Anderson, C. (2008). The End of Theory: The Data Deluge Makes the Scientific Method Obsolete. *Wired Magazine*, 23 June.

Bacon, F. (1961–1986). *Works*. Stuttgart Bad Cannstatt: Frommann Holzboog.

Baker, A. (2016). Non-Deductive Methods in Mathematics. In: E.N. Zalta (ed.), *The Stanford Encyclopedia of Philosophy*, URL = <https://plato.stanford.edu/archives/win2016/entries/mathematics-nondeductive/>.

Bell, J., Hellman, G. (2006). Pluralism and the Foundations of Mathematics. In: S.H. Kellert, H.E. Longino, C.K. Waters (eds.), *Scientific Pluralism*. Minneapolis: University of Minnesota Press: 64–79.

Boolos, G. (1990). On “Seeing” the Truth of the Gödel Sentence. *Behavioral and Brain Sciences*, 13: 655–656.

Boolos, G. (1995). Introductory Note to *1951. In: K. Gödel (1995), *Collected Works. Volume III*. Oxford: Oxford University Press: 290–304.

Byers, W. (2007). *How Mathematicians Think*. Princeton: Princeton University Press.

Calude, C.S., Thompson, D. (2016). Incompleteness, Undecidability and Automated Proofs. In: V. Gerdt et al. (eds.). *Computer Algebra in Scientific Computing. CASC 2016*. Cham: Springer: 134–155.

Cellucci, C. (2006). The Question Hume Didn’t Ask: Why Should We Accept Deductive Inferences? In: C. Cellucci, P. Pecere (eds.), *Demonstrative and Non-Demonstrative Reasoning*. Cassino: Edizioni dell’Università degli Studi di Cassino: 207–235.

Cellucci, C. (2008). Why Proof? What is a Proof? In: R. Lupacchini, G. Corsi (eds.), *Deduction, Computation, Experiment. Exploring the Effectiveness of Proof*. Berlin: Springer: 1–27.

Cellucci, C. (2011). Si può meccanizzare l'induzione? In: B. Continenza et al. (a cura di), *Vittorio Somenzi. Antologia e Testimonianze 1918-2003*. Mantova: Fondazione Banca Agricola Mantovana: 362–364.

Cellucci, C. (2013). *Rethinking Logic*. Dordrecht: Springer.

Cellucci, C. (2017). *Rethinking Knowledge*. Dordrecht: Springer.

Colton, S. (2002). *Automated Theory Formation in Pure Mathematics*. London: Springer-Verlag.

Curry, H.B. (1934). Functionality in Combinatory Logic. *Proceedings of the National Academy of Science*, 20: 584–590.

Davis, M. (1990). Is Mathematical Insight Algorithmic? *Behavioral and Brain Sciences*, 13: 659–660.

Davis, M. (1995). Introductory Note to *193? In: K. Gödel, *Collected Works. Volume III*. Oxford: Oxford University Press: 156–163.

Dummett, M. (1991). *The Logical Basis of Metaphysics*. Cambridge (MA): Harvard University Press.

Dybjer, P., Palmgren, E. (2016). Intuitionistic Type Theory. In: E.N. Zalta (ed.), *The Stanford Encyclopedia of Philosophy*, URL = <https://plato.stanford.edu/archives/win2016/entries/type-theory-intuitionistic/>.

Feferman, S. (1998). *In the Light of Logic*. Oxford: Oxford University Press.

Gigerenzer, G. (1990). Strong AI and the Problem of "Second-Order" Algorithms. *Behavioral and Brain Sciences*, 13: 663–664.

Glymour, C. (1991). The Hierarchies of Knowledge and the Mathematics of Discovery. *Minds and Machines*, 1: 75–95.

Gödel, K. (*193?). Undecidable Diophantine Propositions. In: K. Gödel (1995), *Collected Works. Volume III*. Oxford: Oxford University Press: 164–175.

Gödel, K. (1951). Some Basic Theorems on the Foundations of Mathematics and Their Implications. In: K. Gödel (1995), *Collected Works. Volume III*. Oxford: Oxford University Press: 304–323.

Gödel, K. (*1961/?). The Modern Development of the Foundations of Mathematics in the Light of Philosophy. In: K. Gödel (1995), *Collected Works. Volume III*. Oxford: Oxford University Press: 374–387.

Gödel, K. (1964). What Is Cantor's Continuum Problem? In: K. Gödel (1990), *Collected Works. Volume II*. Oxford: Oxford University Press: 254–270.

Goodman, N. (1983[4]). *Fact, Fiction, and Forecast*. Cambridge (MA): Harvard University Press.

Hamming, R.W. (1980). The Unreasonable Effectiveness of Mathematics. *The American Mathematical Monthly*, 87: 81–90.

Hayes, P.J. (1973). Computation and Deduction. In: *Proceedings of the 2nd Mathematical Foundations of Computer Science Symposium*. Prague: Czechoslovak Academy of Sciences: 105–118.

Hilbert, D. (1970). Axiomatic Thinking. *Philosophia Mathematica*, ser. 1, 7: 1–12 [1st ed.: 1918].

Hintikka, J., Remes, U. (1974). *The Method of Analysis*. Dordrecht: Reidel.

Horsten, L. (2015). Philosophy of Mathematics. In: E.N. Zalta (ed.), *The Stanford Encyclopedia of Philosophy*, URL = <http://plato.stanford.edu/archives/spr2015/entries/philosophy-mathematics/>.

Horsten, L., Welch, P. (2016). Introduction. In: L. Horsten, P. Welch (eds.), *Gödel's Disjunction.* Oxford: Oxford University Press: 1–l5.

Howard, W.A. (1980). The Formulae-as-Types Notion of Construction. In: J.R. Hindley, J.P. Seldin (eds.), *To H.B. Curry. Essays on Combinatory Logic, Lambda Calculus and Formalism*. New York: Academic Press: 479–490.

Jantzen, B.C. (2015). Discovery Without a 'Logic' Would Be a Miracle. *Synthese*, DOI: 10.1007/s11229-015-0926-7.

Kant, I. (1992). *Lectures on Logic*. Cambridge: Cambridge University Press.

King, R.D. et al. (2009). The Automation of Science. *Science*, 324: 85–89.

Koellner, P. (2011). Independence and Large Cardinals. In: Zalta, E.N. (ed.), *The Stanford Encyclopedia of Philosophy*, URL = <https://plato.stanford.edu/archives/sum2011/entries/independence-large-cardinals/>.

Koellner, P. (2014). Large Cardinals and Determinacy. In: E.N. Zalta (ed.), *The Stanford Encyclopedia of Philosophy*, URL = <https://plato.stanford.edu/archives/spr2014/entries/large-cardinals-determinacy/>.

Koellner, P. (2016). Gödel's Disjunction. In: L. Horsten, P. Welch (eds.), *Gödel's Disjunction.* Oxford: Oxford University Press: 148–188.

Kowalski, R.A. (1979). Algorithm = Logic + Control. *Communications of the ACM*, 22: 424–436.

Kripke, S.A. (2013). The Church-Turing 'Thesis' as a Special Corollary of Gödel's Completeness Theorem. In: B.J. Copeland, C.J. Posy, O. Shagrir (eds.), *Computability*. Cambridge (MA): MIT Press: 77–104.

Lakatos, I. (1976). *Proofs and Refutations*. Cambridge: Cambridge University Press.

Lakatos, I. (1978). *Philosophical Papers*. 2 Voll. Cambridge: Cambridge University Press.

Laplace, P.S. (1951). *A Philosophical Essay on Probabilities*. New York: Dover Publications. [1st French ed.: 1814]

Laudan, L., (1981). A Confutation of Convergent Realism. *Philosophy of Science*, 48: 19–49.

Leach-Krouse, G. (2016). Provability, Mechanism, and the Diagonal Problem. In: L. Horsten, P. Welch (eds.), *Gödel's Disjunction.* Oxford: Oxford University Press: 211–242.

Longo, G. (2003). Proofs and Programs. *Synthese*, 134: 85–117.

Longo, G. (2011). Reflections on Concrete Incompleteness. *Philosophia Mathematica*, 19: 255–280.

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

Maddy, P. (1988). Believing the Axioms. I. *The Journal of Symbolic Logic*, 53: 481–511.

Mäenpää, P. (1997). From Backward Reduction to Configurational Analysis. In: M. Otte, M. Panza (eds.), *Analysis and Synthesis in Mathematics: History and Philosophy*. Dordrecht: Springer: 201–226.

Marcus, G. (2018). Innateness, AlphaZero, and Artificial Intelligence. arXiv:1801.05667v1.

Markov, A.A. (1971). On Constructive Mathematics. In: American Mathematical Society Translations, Ser. 2, Vol. 98. Providence: Rhode Island: 1–9. [1st Russian ed.: 1962]

Matiyasevič, Y. (2003). Enumerable Sets Are Diophantine. In: G.E. Sacks (ed.), *Mathematical Logic in the 20th Century*. Singapore: Singapore University Press: 269–273.

Mazzocchi, F. (2015). Could Big Data Be the End of Theory in Science? A Few Remarks on the Epistemology of Data-Driven Science. *EMBO reports*, 16: 1250–1255.

Muggleton, S., De Raedt, L. (1994). Inductive Logic Programming. Theory and Methods. *Journal of Logic Programming*, 19–20: 629–679.

Newell, A., Shaw, J.C., Simon, H.A. (1957). Empirical Explorations of the Logic Theory Machine: A Case Study in Heuristic. In: *Proceedings of the 1957 Western Joint Computer Conference*. New York: ACM: 218–230.

Penrose, R. (1989). *The Emperor's New Mind*. Oxford: Oxford University Press.

Pólya, G. (1941). Heuristic Reasoning and the Theory of Probability. *The American Mathematical Monthly*, 48: 450–465.

Pólya, G. (1954). *Mathematics and Plausible Reasoning*. Princeton: Princeton University Press.

Popper, K.R. (2005). *The Logic of Scientific Discovery*. London: Routledge.

Portoraro, F. (2019). Automated Reasoning, In: E.N. Zalta (ed.), *The Stanford Encyclopedia of Philosophy*, URL = <https://plato.stanford.edu/archives/spr2019/entries/reasoning-automated/>.

Prawitz, D. (2008). Proofs Verifying Programs and Programs Producing Proofs: A Conceptual Analysis. In: R. Lupacchini, G. Corsi (eds.), *Deduction, Computation, Experiment. Exploring the Effectiveness of Proof*. Berlin: Springer: 81–94.

Prawitz, D. (2014). The Status of Mathematical Knowledge. In: E. Ippoliti, C. Cozzo (eds.), *From a Heuristic Point of View. Essays in Honour of Carlo Cellucci*. Newcastle upon Tyne: Cambridge Scholars Publishing: 73–90.

Raatikainen, P. (2005). On the Philosophical Relevance of Godel's Incompleteness Theorems. *Revue internationale de philosophie*, 4: 513534.

Rathjen, M., Sieg, W. (2018). Proof Theory. In: E.N. Zalta (ed.), *The Stanford Encyclopedia of Philosophy*, URL = <https://plato.stanford.edu/archives/fall2018/entries/proof-theory/>.

Rescorla, M. (2017). The Computational Theory of Mind. In: E.N. Zalta (ed.), *The Stanford Encyclopedia of Philosophy*, URL = <https://plato.stanford.edu/archives/spr2017/entries/computational-mind/>.

Rodin, A. (2014). *Axiomatic Method and Category Theory*. Berlin: Springer.

Schickore, J. (2014). Scientific Discovery. In: Zalta, E.N. (ed.), *The Stanford Encyclopedia of Philosophy*, URL = <http://plato.stanford.edu/archives/spr2014/entries/scientific-discovery/>.

Shapiro, S. (2016). Idealization, Mechanism, and Knowability. In: L. Horsten, P. Welch (eds.), *Gödel's Disjunction.* Oxford: Oxford University Press: 189–207.

Sørensen, M.H., Urzyczyn, P. (2006). *Lectures on the Curry-Howard Isomorphism*. Amsterdam: Elsevier.

Sparkes, A. et al. (2010). Towards Robot Scientists for Autonomous Scientific Discovery. *Automated Experimentation*, 2:1, DOI: 10.1186/1759-4499-2-1.

Sterpetti, F. (2018). Mathematical Knowledge and Naturalism. *Philosophia*, DOI: 10.1007/s11406-018-9953-1.

Sterpetti, F., Bertolaso, M. (2018). The Pursuit of Knowledge and the Problem of the Unconceived Alternatives. *Topoi. An International Review of Philosophy*, DOI: 10.1007/s11245-018-9551-7.

von Plato, J. (2018). The Development of Proof Theory. In: E.N. Zalta (ed.), *The Stanford Encyclopedia of Philosophy*, URL = <https://plato.stanford.edu/archives/win2018/entries/proof-theory-development/>.

Wadler, P. (2015). Propositions as Types. *Communications of the ACM*, 58: 75–84.

Weyl, H. (1949). *Philosophy of Mathematics and Natural Science*. Princeton: Princeton University Press.

Wiggins, G.A. (2006). Searching for Computational Creativity. *New Generation Computing*, 24: 209–222.

Williamson, T. (2016). Absolute Provability and Safe Knowledge of Axioms. In: L. Horsten, P. Welch (eds.), *Gödel's Disjunction.* Oxford: Oxford University Press: 243–253.

Wos, L., Pereira, F., Hong, R. et al. (1985). An Overview of Automated Reasoning and Related Fields. *Journal of Automated Reasoning*, 1: 5–48.

Zach, R., (2016). Hilbert's Program. In: E.N. Zalta (ed.), *The Stanford Encyclopedia of Philosophy*, URL = <https://plato.stanford.edu/archives/spr2016/entries/hilbert-program/>.