Understanding in Mathematics: The Case of Mathematical Proofs

Yacin Hamami* Rebecca Lea Morris†

Forthcoming in *Noûs*

Abstract

Although understanding is the object of a growing literature in epistemology and the philosophy of science, only few studies have concerned understanding in mathematics. This essay offers an account of a fundamental form of mathematical understanding: proof understanding. The account builds on a simple idea, namely that understanding a proof amounts to rationally reconstructing its underlying plan. This characterization is fleshed out by specifying the relevant notion of plan and the associated process of rational reconstruction, building in part on Bratman's theory of planning agency. It is argued that the proposed account can explain a significant range of distinctive phenomena commonly associated with proof understanding by mathematicians and philosophers. It is further argued, on the basis of a case study, that the account can yield precise diagnostics of understanding failures and can suggest ways to overcome them. Reflecting on the approach developed here, the essay concludes with some remarks on how to shape a general methodology common to the study of mathematical and scientific understanding and focused on human agency.

Contents

*Postdoctoral Researcher FNRS, Philosophy Department, University of Liège, Belgium. Email: yacin.hamami@gmail.com.

†Independent Scholar. Minneapolis, MN. Email: email@rebeccaleamorris.com.

1 Introduction

As of late, the nature of understanding is gaining increasing attention in epistemology, the philosophy of science, and the philosophy of mathematics.[1] One way to move forward on this issue is to identify and characterize general forms of understanding such as understanding *that*, understanding *why*, or understanding X where X is some object susceptible of being understood.[2] Another, complementary way is to conduct in-depth studies of particular, yet fundamental instances of understanding. Mathematics and the sciences offer a wealth of opportunities to progress on this latter front. Among the various entities that can be the object of understanding in scientific and mathematical practice, the case of mathematical proofs constitutes a promising candidate to pursue such investigations. For the capacity to understand pieces of mathematical reasoning is not only essential to anyone engaged with mathematics—from mathematics students to research mathematicians—it is also indispensable to anyone involved in scientific fields which rely on mathematical resources. Furthermore, the understanding of mathematical proofs is most often a precondition to, and a vehicle for, other forms of understanding such as understanding *why* a theorem is true or *why* a mathematicized scientific theory yields specific predictions. Thus, getting to grips with what it means to understand mathematical proofs would not only constitute a direct contribution to the philosophy of mathematics and the philosophy of science, it would also provide a privileged test case for the epistemology of understanding.

Our aim in the present work is to develop an account of the understanding of mathematical proofs. Our starting point is the common observation in mathematical practice that there is an important difference between *understanding* a proof and *verifying* it. Henri Poincaré put the distinction as follows:

> Does understanding the demonstration of a theorem consist in examining each of the syllogisms of which it is composed in succession, and being convinced that it is correct and conforms to the rules of the game? [...]
>
> Yes, for some it is ; when they have arrived at the conviction, they will say, I understand. But not for the majority. Almost all are more exacting; they want to know not only

[1] For general overviews of the philosophical literature on understanding, see Grimm (2011, 2012, 2021), Gordon (2017), Baumberger et al. (2017), and Kvanvig (2017).

[2] This latter form of understanding is often referred to as 'objectual' understanding.

> whether all the syllogisms of a demonstration are correct, but why they are linked together in one order rather than in another. As long as they appear to them engendered by caprice, and not by an intelligence constantly conscious of the end to be attained, they do not think they have understood. (Poincaré, 1908, p. 118)

Here is another statement of the distinction, this time formulated in the words of Bourbaki:[3]

> [E]very mathematician knows that a proof has not really been "understood" if one has done nothing more than verifying step by step the correctness of the deductions of which it is composed, and has not tried to gain a clear insight into the ideas which have led to the construction of this particular chain of deductions in preference to every other one. (Bourbaki, 1950, p. 223)

Although proof verification is a central topic of study in logic and the philosophy of mathematics, proof understanding has comparatively received much less attention in the literature. The situation has improved in the past few years with several authors investigating the relation(s) between mathematical proof and mathematical understanding.[4] Yet, providing a precise account of proof understanding remains an important philosophical challenge. As the mathematician Timothy Gowers pointed out: "It is not easy to say precisely what it means to understand a proof" (Gowers, 2007, p. 41).

Proof understanding falls under the category of understanding X or objectual understanding where the objects of understanding are mathematical proofs. It is useful to notice that objectual understanding can refer both to an *epistemic state* and to an *epistemic process*. When we say that an agent "understands X", "does not understand X", or "understands X to a certain degree", we are referring to an epistemic state. But when we say that an agent "is trying to understand X", "has managed to understand X", or "has failed to understand X", we are referring to an epistemic process. Of course, the two are intimately connected: it is by going through a certain process of understanding that an agent can end up into a certain state of understanding. This

[3]Other statements of the distinction can be found in Thurston (1994, p. 162), Gowers (2007, p. 41), Avigad (2008, p. 319), and Feferman (2012, p. 372).

[4]See, in particular, Thurston (1994), Resnik (1996), Heinzmann (1999), Robinson (2000), Tappenden (2005), Avigad (2008, 2010), Manders (2012), Macbeth (2012), Feferman (2012), Cellucci (2015), Folina (2018), Delarivière and Van Kerkhove (2021), Lehet (2021), Frans (2021), and Inglis and Mejía-Ramos (2021).

opens two possible paths of investigation into the nature of objectual understanding: one can attempt a direct characterization of "A understands X" by identifying a relevant set of necessary and sufficient conditions, in the same way as one may proceed for characterizing the epistemic states expressed by "A knows X" or "A believes X"; or one can attempt a characterization of the process of understanding itself, i.e., of what it takes for A to get into an epistemic state where A understands X. One advantage of the latter over the former is that, by characterizing the process of understanding X, one automatically gets a characterization of the state of understanding X, for an agent understands X whenever she has successfully completed the associated process of understanding X. In this work, we will adopt the latter approach and attempt to characterize the epistemic state of understanding a mathematical proof by characterizing the epistemic process of proof understanding.

What does it take for an agent to understand a mathematical proof? The proposal to be developed in this work is that understanding a proof requires the agent to *rationally reconstruct* the *plan* of the proof. In other words, our proposal is that the process of understanding a proof is a process of *rational reconstruction* of the proof's underlying plan. This leads to the following characterization:

An agent *understands* a mathematical proof P

$\Leftrightarrow$

She has been able to *rationally reconstruct* the plan underlying P.

The core of the paper will be devoted to specifying this characterization by spelling out the notion of proof plan in section 2 and the associated process of rational reconstruction in section 3. The resulting *plan-based account* of proof understanding will be evaluated in sections 4 and 5: in section 4, we will show that the plan-based account can diagnose failures of understanding and can suggest ways to overcome them—this will also be the occasion to demonstrate how the account works on a concrete example; in section 5, we will argue that the plan-based account can explain a significant range of distinctive phenomena that have been commonly associated with proof understanding in the literature. In section 6, we will compare the plan-based account with existing accounts of objectual understanding in epistemology and the philosophy of science. Reflecting on the present study, section 7 will provide a critical assessment of the plan-based account while 8 will conclude with some remarks on how to provide a general methodology common to the study of mathematical and scientific understanding and focused on human agency.

2 The Notion of Plan for Mathematical Proofs

In the most common situation, when one is trying to understand a mathematical proof, one is faced with a written text of the kind typically found in mathematical articles, textbooks, or monographs. This written text describes a sequence of *deductive inferences*—i.e., a sequence of *epistemic actions*—which, if carried out by an agent with the appropriate background knowledge, will bring her to an epistemic state in which she knows the theorem to be established. In this work, we will use the term 'mathematical proof' to refer to written mathematical proofs.[5] The sequence of deductive inferences described in a mathematical proof will be called a 'proof activity'.[6]

A proof activity is then a sequence of epistemic actions that one or more mathematical agents have found in order to achieve a specific goal, namely to establish the mathematical theorem at hand. Presumably, each of these epistemic actions is the result of a *rational choice* made on the basis of a *practical decision*—i.e., a decision regarding which action(s) to carry out in order to reach the goal. In this respect, proof activities are similar to many of our most ordinary activities that are goal-directed and that unfold over time such as travelling or cooking. Such activities necessarily require the agent to make numerous practical decisions that are essential for their concrete realization—for instance, any travelling activity will require the agent to decide which mode(s) of transport to use, and any cooking activity will require the agent to decide which ingredients to employ, among many other things. All these practical decisions are part of the process of *planning* the considered activity, that is, they are responsible for the construction of the *plan* underlying the considered activity.

Such a notion of plan for proof activities can be fleshed out by building on Michael Bratman's theory of planning agency.[7] For Bratman, a plan is a complex network of *intentions* related in

[5]This is a common use of the term, for instance when we say that "a mathematical proof of theorem X can be found in book Y or in article Z".

[6]Seeing mathematical proofs as describing sequences of deductive inferences is, arguably, an idealization since written mathematical proofs also report other types of actions such as assuming or computing (see, e.g., Tanswell, forthcoming). Following Poincaré and Mac Lane, we chose to focus on deductive inferences as the main building blocks of proofs. It would then be interesting to explore in further work how the account of proof understanding proposed here could be extended to a richer conception of proof activity.

[7]For a general presentation of Bratman's theory of planning agency, see Bratman (1987, 2010) and the subsequent literature. The notion of plan for proof activities presented in this section is the one we developed previously in Anonymized (forthcoming). A different, but related notion of proof plan has been proposed by Bundy and colleagues in the field of artificial intelligence (see, e.g., Bundy, 1988; Bundy et al., 2005). A comparison of the two notions can

specific ways. A plan usually starts from a single intention such as building a house, baking a cheesecake, going to Paris for the summer holidays, or proving Goldbach's conjecture. This intention needs to be turned into more specific intentions as the activity proceeds. For instance, an agent may turn her intention to build a house into more specific intentions such as digging the foundations, bringing specific materials to the construction site, establishing connections to the water and electricity networks, etc. In doing so, the agent is specifying further—i.e., *constructing*—her plan to build a house. In Bratman's theory, each step of plan construction is the result of an instance of *practical reasoning*. Practical reasoning is then the central process by which an agent can turn an intention in her current plan into more specific intentions organized in a subplan.

In the context of proof activities, intentions always have the same form: they are intentions *to show* (prove, establish) a given mathematical proposition *from* other mathematical propositions—we will refer to them as *proving intentions*. For instance, an agent who intends to show the Intermediate Value Theorem intends *to show* that a real function $f : [a, b] \to \mathbb{R}$ takes on each value between $f(a)$ and $f(b)$ *from* the hypothesis that f is a continuous function on the interval $[a, b]$. In this context, it is useful to distinguish between two types of proving intentions: proving intentions of type *'to show'* which are intentions that cannot be fulfilled directly, and so will always need to be turned into more specific proving intentions; proving intentions of type *'to infer'* which are intentions to carry out a deductive inference and which can be fulfilled directly by actually carrying out the considered inference. We will use the following sequent-style notations for proving intentions:

$$P_1, \ldots, P_n \quad \Rightarrow \quad C \quad \text{as a shortcut for the intention} \quad \textbf{to show} \quad C \quad \textit{from} \quad [P_1, \ldots, P_n],$$
$$P_1, \ldots, P_n \quad \hookrightarrow \quad C \quad \text{as a shortcut for the intention} \quad \textbf{to infer} \quad C \quad \textit{from} \quad [P_1, \ldots, P_n],$$

where $P_1, \ldots, P_n$ and C are placeholders for ordinary mathematical propositions—$P_1, \ldots, P_n$ and C are referred to as the *hypotheses* and the *conclusion* of the considered proving intention. Practical reasoning in the context of proof activities is then a process that takes as input a proving intention of type 'to show' and yields as output a subplan, that is, a specification of a possible course of action that may fulfill this proving intention.[8] As an illustration, consider an agent who intends to show

be found in Anonymized (forthcoming, section 8).

[8]Practical reasoning in this sense is closely related to the notion of *tactics* used in interactive and automated theorem proving, especially to the one developed by Ganesalingam and Gowers (2017) in their theorem prover

that a given relation $\sim$ is an equivalence relation on a set S from some propositions $P_1, \ldots, P_n$. This agent could reason from this intention to the following subplan:

To show that $\sim$ is an equivalence relation on S, *from* $[P_1, \ldots, P_n]$:

1. **Show** $\sim$ is reflexive, *from* $[P_1, \ldots, P_n]$,
2. **Show** $\sim$ is symmetric, *from* $[P_1, \ldots, P_n]$,
3. **Show** $\sim$ is transitive, *from* $[P_1, \ldots, P_n]$,
4. **Infer** $\sim$ is an equivalence relation, *from* $\sim$ is reflexive *and* $\sim$ is symmetric *and* $\sim$ is transitive.

Using our sequent-style notation, this instance of practical reasoning can be written as follows:

$$[P_1, \ldots, P_n] \quad \Rightarrow \quad \sim \text{ is an equivalence relation on } S$$

$$\downarrow$$

$$[P_1, \ldots, P_n] \quad \Rightarrow \quad \sim \text{ is reflexive}$$

$$[P_1, \ldots, P_n] \quad \Rightarrow \quad \sim \text{ is symmetric}$$

$$[P_1, \ldots, P_n] \quad \Rightarrow \quad \sim \text{ is transitive}$$

$$\begin{array}{r} \sim \text{ is reflexive} \\ \sim \text{ is symmetric} \\ \sim \text{ is transitive} \end{array} \quad \hookrightarrow \quad \sim \text{ is an equivalence relation on } S$$

This is an elementary form of practical reasoning, for it simply consists in unpacking the definition of an equivalence relation. In practice, practical reasoning can be much more complex, often involving long and tedious chains of trial and error in order to come up with a subplan to address the considered proving intention.

We now have all the elements to define the notion of plan for proof activities[9] (see Figure 2.1

---

designed to mimic human mathematical reasoning.

[9] This definition is the one provided in Anonymized (forthcoming).

for an illustration): an agent's *plan* for a proof activity is an *ordered tree*[10] such that (1) each *node* is a proving intention, (2) the *root* is the proving intention corresponding to the theorem at hand, and (3) each *set of ordered children* of a given parent node is a subplan that has been obtained from the parent node through an instance of practical reasoning. The *execution* of an agent's plan for a proof activity leads to the actual *realization* of the proof activity, i.e., to the carrying out of the associated sequence of deductive inferences. From this perspective, a mathematical proof is nothing more than a *report* of its associated proof activity. Our proposal is then to think of the plan of a mathematical proof simply as the plan underlying its associated proof activity.[11]

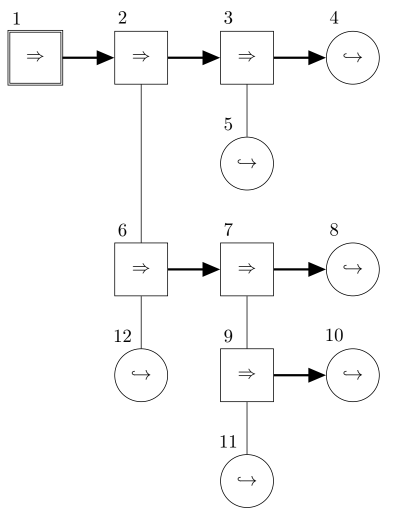

Figure 2.1: A schematic representation of a proof plan. Squared nodes represent proving intentions of type 'to show' and circled nodes represent proving intentions of type 'to infer'. The root of the tree is the proving intention to show the theorem at hand. In a complete proof plan such as this one, every intention of type 'to show' is decomposed into further intentions organized in a subplan. Executing the plan amounts to traverse this tree in the expected way (which is numbered in the figure) and for each proving intention of type 'to infer' encountered, to carry out the corresponding deductive inference. The execution of the plan leads then to the realization of a proof activity, i.e., a sequence of deductive inferences.

[10]In the mathematical sense of the term: an *ordered tree* is a rooted tree where each node comes equipped with an ordering of its children.

[11]Indeed, this is exactly what we do when we think about plans in the context of reports of human activities. Consider, for instance, a newspaper article reporting a bank robbery. In this case, the newspaper is a report of a particular activity—i.e., a sequence of actions—consisting in stealing money from a bank. If we are talking about the plan behind the bank robbery, we are talking about the plan of the agents who carried out the stealing activity. Describing this plan will involve reporting the various intentions of the bank robbers, how these intentions led to more specific ones, i.e., how they are hierarchically organized in the robbers' plan, and maybe some of the reasons behind the practical choices made by the protagonists.

3 The Plan-Based Account of Proof Understanding

Understanding a mathematical proof is a *process* that requires the *active* involvement of the agent that seeks to understand. An agent understands a mathematical proof whenever she has successfully completed this process. But what does this process consist in? As announced in the introduction, our proposal is that this process amounts to the rational reconstruction of the plan underlying the mathematical proof one aims to understand. Our proposed characterization is that an agent *understands* a mathematical proof P if and only if she has been able to *rationally reconstruct* the plan underlying P. Thanks to the notion of plan for mathematical proofs provided in the previous section, we can now specify further what this condition amounts to. In particular, we can decompose the process of rational reconstruction in two parts that we will refer to as the *tracking* and the *rationality* parts. The characterization becomes:

An agent *understands* a mathematical proof P

$\Leftrightarrow$

She has been able to:

1. Reconstruct the tree structure of the plan underlying P [tracking part]

2. Reproduce by herself all the instances of practical reasoning involved in the plan construction [rationality part]

We will now describe these two parts of the understanding process in turn.

Reconstructing the tree structure of the plan underlying a mathematical proof amounts to *tracking* the proving intentions of the author as they proceed from the initial intention to prove the theorem at hand to more and more specific intentions in the course of the plan construction. In this process, the agent must be able to realize whenever a proving intention has been turned into a subplan, and to identify the corresponding subplan. Equivalently, this means that the agent must be able to identify all the steps of plan construction, and for each step to see how the plan has been extended by the addition of one or more subplan(s). If the agent managed to successfully complete this process of tracking intentions and subplans, then she would have reconstructed the tree structure of the plan, and so would have successfully completed the tracking part.

This process is best illustrated with an example. Let's consider the following toy example:

Proposition. If n is an integer, then $n^2 + 3n$ is even.

Proof. Let n be an integer. If n is even, then $n = 2k$ for some integer k. We have:

$$\begin{aligned} n^2 + 3n &= (2k)^2 + 3(2k) \\ &= 2(2k^2 + 3k) \end{aligned}$$

and so $n^2 + 3n$ is even. If n is odd, then $n = 2k + 1$ for some integer k. We have:

$$\begin{aligned} n^2 + 3n &= (2k+1)^2 + 3(2k+1) \\ &= 2(2k^2 + 5k + 2) \end{aligned}$$

and so $n^2 + 3n$ is even. We conclude that, if n is an integer, then $n^2 + 3n$ is even. □

Here, the initial proving intention of the author is to show that n^2+3n is even from the hypothesis that n is an integer. How is this initial intention turned into more specific intentions? From the written proof, we can see that the plan of the author is first to show that n^2+3n is even when n is even, then to show that n^2+3n is even when n is odd, and finally to infer from this that n^2+3n is even when n is an integer. This means that the author has turned the initial proving intention into a subplan as follows:

$$n \text{ is an integer} \quad \Rightarrow \quad n^2 + 3n \text{ is even}$$

$$\downarrow$$

$$n \text{ is an integer}, n \text{ is even} \quad \Rightarrow \quad n^2 + 3n \text{ is even}$$

$$n \text{ is an integer}, n \text{ is odd} \quad \Rightarrow \quad n^2 + 3n \text{ is even}$$

$$\begin{array}{r} n \text{ is an integer} \\ n \text{ is even} \rightarrow n^2 + 3n \text{ is even} \\ n \text{ is odd} \rightarrow n^2 + 3n \text{ is even} \end{array} \quad \hookrightarrow \quad n^2 + 3n \text{ is even}$$

This corresponds to the topmost part of the tree structure of the plan, more precisely the root and its direct children. Then, the first and second proving intentions in the above subplan are themselves turned into more specific intentions. For instance, the first proving intention to show $n^2 + 3n$ is even from n is even leads to a subplan composed of two proving intentions: first to infer that $n = 2k$ for some integer k from n is even, and then to show from this that $n^2 + 3n$ is even. This latter proving intention is itself turned into two proving intentions: first to infer that $n^2 + 3n = (2k)^2 + 3(2k)$, and then to show from this that $n^2 + 3n$ is even. The last bit, which completes this part of the plan, is to infer that $n^2 + 3n = 2(2k^2 + 3k)$ from $n^2 + 3n = (2k)^2 + 3(2k)$, and finally to infer that $n^2 + 3n$ is even (from the definition of "even"). In the very same way, one can track how the second proving intention in the above subplan is turned into more specific intentions. Once an agent has completed this process of tracking intentions and subplans, she will have reconstructed the tree structure of the plan.

Rationally reconstructing the plan amounts, in addition to the reconstruction of its tree structure, to being able to reproduce all the instances of practical reasoning—i.e., all the practical decisions—involved in the plan construction. Thus, while the process of reconstructing the tree structure was concerned with identifying the different steps in the plan construction, the process of rational reconstruction is concerned with the practical decisions behind the choice of each of these steps. When we say that the agent has been able to reproduce by herself an instance of practical reasoning in the plan construction, we mean that the agent has been able to connect the input (a proving intention) to the output (a subplan) by one or more practical inferences of her own, i.e., by an exercise of her own practical reason. If the agent managed to successfully do so for all the construction steps in the plan, then she would have successfully completed the rationality part.

When we described the process of tracking intentions and subplans in the above toy example, we were only concerned with tracking when a proving intention has been turned into a subplan and with identifying the corresponding subplan. We can now ask about the practical decisions behind each of these transformations, i.e., behind each construction step in the plan. For instance, we may ask how the author came up with the decision to turn the initial proving intention into the subplan represented above. In this case, the practical reasoning behind this decision simply consists in applying a well-known mathematical method when it comes to proving a proposition for all integers, namely prove it separately for the even and for the odd numbers. The other steps of plan construction consist mainly in obvious algebraic manipulations to reach the desired goals. Thus, an agent with an elementary background in number theory who is familiar with these methods will be able to rationally reconstruct the plan underlying this toy example. Of course, more sophisticated proofs will require more practical reasoning expertise in order to reconstruct the underlying plans and thus, according to the present account, to understand them.

Maximal, Partial and Degrees of Proof Understanding Our characterization as stated above describes what may be called *maximal* or *full* understanding of a mathematical proof. In practice, however, mathematical agents rarely go through the whole process of rationally reconstructing the entire plan underlying a proof. Most often, depending on their epistemic goals and needs, they can be satisfied with a *partial* understanding of the proof under consideration—indeed, it is commonly acknowledged in the philosophical literature that understanding can be partial and can come in degrees. The plan-based account can straightforwardly accommodate such a notion of partial understanding in terms of partial completion of the rational reconstruction process. More specifically, there are two ways in which proof understanding can be partial according to the plan-based account. First, the agent may not have succeeded in reconstructing the totality of the tree structure of the plan. Second, the agent may not have succeeded in reproducing some of the instances of practical reasoning involved in the plan construction. In both cases, the agent will only have a partial understanding of the considered proof in the straightforward sense that there are parts of the proof that the agent does not understand. These two sources of partial understanding are not incompatible and can coexist with respect to the same proof: the agent may not have succeeded in reconstructing certain parts of the tree structure of the plan, and in the parts that she has successfully reconstructed, she may not have succeeded in reproducing some of the instances of practical reasoning.

A key feature of this conception of partial proof understanding is that it can be precisely *localized*: because of the tree structure of proof plans, it is possible to identify the parts of a plan that have been rationally reconstructed in a successful way from those that were not. The plan-based account yields then a precise notion of *degrees* of proof understanding: the degree with which an agent understands a mathematical proof can be measured in terms of the proportion of the plan that she has rationally reconstructed in a successful way, or equivalently in terms of the proportion of the process of rational reconstruction that she managed to complete.[12] This measure can be further refined by attributing specific weights to the different nodes depending on the difficulty of the associated practical reasoning. An agent possesses then a *full* or *maximal* understanding of mathematical proof whenever she understands it at the highest degree, i.e., when she has rationally reconstructed the entire plan underlying it—this corresponds to our main characterization. In practice, outright attributions of proof understanding to oneself or others usually do not require full or maximal understanding and are more likely governed by a threshold of degrees of understanding determined by the context.[13]

Clarificatory Remarks This completes the general description of the plan-based account of proof understanding. Here we shall provide further details on how the characterization works in order to account for (1) what happens on the side of the agent engaged in the process of understanding a proof, and (2) how attributions of proof understanding are made.

An assumption underlying our characterization is that, throughout the understanding process, the agent proceeds *as if* there is a rational plan underlying the proof P she aims to understand,

[12]This conception of partial understanding and degrees of understanding presents some similarities with accounts of these notions that involve networks of some sort. For instance, Kelp (2015) has advanced an account of understanding phenomena where degrees of understanding are measured in terms of distance or approximation to "fully comprehensive and maximally well-connected knowledge [of the phenomenon]" (Kelp, 2015, p. 3812). For Kelp, an agent possesses fully comprehensive and maximally well-connected knowledge of a phenomenon whenever "the basing relations that obtain between the agent's beliefs about P reflect the agent's knowledge about the explanatory and support relations that obtain between the members of the full account of P" (Kelp, 2015, p. 3810). According to this account, one can then represent knowledge of a phenomenon in terms of a graph with propositions as nodes and basing relations as connections between these nodes; degrees of understanding are then measured in terms of the proportion of these basing relations grasped by the agent. We find a similar idea expressed by Carter and Gordon (2014) when they say that: "Understanding wider subject matters will tend to be more cognitively demanding than understanding narrow subject matters because *more* propositions must be believed and their relations grasped in order for one to even attain the most minimal understanding" (Carter and Gordon, 2014, p. 8). The plan-based account yields a conception of degrees of understanding that is similar to such views in the sense that degrees of understanding are measured in terms of the extent to which the agent grasps a certain network or graph-theoretic structure, namely the tree structure of a proof plan.

[13]This idea has been developed in some depth by Kelp (2015, section 5.3) and can be directly adapted to the present context. Further work is required to determine exactly how this threshold is determined, and more specifically whether it is linked to specific parts of the plan, for instance its higher-level structure—we will come back to the notion of higher-level understanding in section 5.4.

i.e., *as if* P has been produced by a rational planning agent. In most cases, this is a genuine hypothesis entertained by the agent. This hypothesis is justified given that the vast majority of mathematical proofs in the mathematical literature are indeed produced by rational planning agents—this is often the main reason why it makes sense for the agent to try to reconstruct a proof's underlying plan in the first place. But we can imagine cases where the agent knows that the proof has not been produced by a rational planning agent, and yet wishes to understand it. For instance, we can imagine that an artificial intelligence system has been trained with many proofs produced by humans and was then asked to produce a proof by itself. In this case, we may expect the artificial intelligence system to display a human-like behavior, and more specifically a rational planning behavior. It then makes sense to try to understand such a proof *as if* it was produced by a rational planning agent. We can also imagine cases where this hypothesis is wrong but the agent does not know it. This could be the case for proofs that have been discovered (in part) by luck or that originate from a faulty plan. In those cases, nothing prevents other mathematical agents from understanding such proofs at a later stage, and to rationally reconstruct plans for them. In fact, it is very common in mathematical practice that agents come to understand a proof much better than the author herself. This happens, for instance, in cases where a proof has been produced through a strategy that anticipates notions and methods to be developed later on—a typical example here would be some of Euler's and Gauss' proofs in number theory that anticipate group-theoretic methods (see, e.g., Wussing, 1984, part I, chapter 3). Such strategies may have been discovered initially by trial and error, but they may be systematized and generalized at a later stage in the development of mathematics. In these cases, it will be possible for future agents to reconstruct rational plans for these proofs that were unthinkable by the original authors.

However, in the vast majority of cases, the plan to be rationally reconstructed will just be the plan of the author[14] who produced the corresponding proof activity. In practice, there is often only one obvious way of reconstructing a proof's underlying plan. Furthermore, the majority of

[14] As we did above, we will often refer to the agent who produced the proof as the "author" of the proof. In the context of this paper, we will always talk about "the" author of the proof in the singular. When a proof has been produced by a group of mathematical agents working in collaboration, "the" author refers to the group taken as a collective agent. This subtlety is not directly relevant for the plan-based account, for what matters for understanding is the process of plan construction itself, not the specifics of the agent(s) who came up with the plan in the first place.

plan construction steps just reflect knowledge that is shared by the relevant community. The cases mentioned above where a proof has not been generated from a rational plan or has been generated by a faulty plan are more the exception than the rule. As we just argued, the plan-based account can handle these special cases insofar as it is possible for an agent to understand a proof that was not generated by a plan that has been rationally constructed. In these cases, we will still speak of "rationally reconstructing the plan underlying P" because this reflects what is happening from the perspective of the agent who aims to understand insofar as the agent proceeds *as if* there is an underlying plan to be reconstructed.[15] But from an external perspective, the agent is simply constructing a rational plan.

How do attributions of proof understanding work according to the plan-based account? Consider a situation where an agent A aims to judge whether an agent B understands a mathematical proof P—here B can be another agent or A herself, in which case A is judging her own understanding of P. We should distinguish between two cases depending on whether agent A has no or some information about the author. In the former case, the best A can do is judge whether B has succeeded in rationally constructing a plan for P, that is, whether the plan B has (re)constructed is indeed a rational plan for P. In the second case, A can judge not only whether B has succeeded in rationally (re)constructing a plan for P, she can also judge to what extent this plan is similar or different from that of the author. In the case where it is different, A may judge that B understands P although not in the way intended by the author. This situation would be roughly similar to the comprehension of a text that admits multiple interpretations where we may say that an agent understands a text, though not in the way the author intended.

This picture of proof understanding attributions should be refined by taking into account the above considerations on partial understanding. As discussed above, mathematical agents in practice rarely reconstruct the entire plan of a proof. This means that outright attributions of proof understanding are more likely made on the basis of a certain degree of proof understanding and are thus governed by a threshold to be determined by the context. This also means that it is possible to attribute proof understanding with qualifiers depending on the degree of understanding, for in-

[15] As a reviewer suggested, it may be more appropriate to speak here of "grasping a viable plan for P", thereby avoiding to speak of reconstructing a plan when no plan has been constructed in the first place. This alternative formulation possesses a number of advantages, one of them being to establish a natural connection with other accounts of understanding formulated in terms of grasping.

stance when judging whether an agent has a "poor" or "good" understanding of a proof. Finally, it is possible to judge that an agent has a better understanding of a proof than another agent by comparing their degrees of proof understanding. Equipped with the notions of partial proof understanding and degree of proof understanding, the plan-based account possesses then the resources to handle different cases of proof understanding attributions. However, determining precisely the values that govern outright proof understanding attributions, qualifiers such as "poor" and "good", and comparison between agents is an empirical question that calls for a dedicated study.

Verifying vs Understanding Proofs From the above description, we can already see that the process of understanding a proof as conceptualized in the plan-based account is quite different in nature from that of verifying a proof. In the verification process, the agent evaluates one inference after the other in a linear way, and all that matters is that each inference be valid. To verify a proof, there is no need to track the intentions and the plan construction of the author. There is no need to locate the practical decisions that lead to the construction of this particular plan, and a fortiori, there is no need to reproduce these practical decisions. In other words, there is no need to see how each inference contributes to the main objective, namely to prove the theorem at hand. The only thing to check is that the chain of valid inferences does end up with the mathematical proposition to be proved. Another important difference between the processes of understanding and verifying a proof lies in their respective temporal structures. The two processes are concerned with the inferences of the considered proof, but the order in which the inferences are processed is different. The toy proof example above illustrates this: the last inference to be made in the proof activity already figures in the first construction step of the plan. This temporal difference is the manifestation of the familiar fact that the order in which actions are planned in a given activity is often not the same as the order in which the actions are carried out when the plan is executed. Take the banal example of a flight trip: the choice of which bus to take to go to the airport is likely to be one of the last practical decisions to be made in planning a trip, and yet this is going to be one of the first actions in the trip, i.e., the activity resulting from the execution of the plan. From the perspective of the plan-based account, one may say that verifying is more concerned with the legitimacy of the actions in the proof activity, while understanding is more concerned with the rational thinking that leads to this particular sequence of actions in the first place.

Comparison with Avigad's Account It is instructive to compare the plan-based account of proof understanding with that advanced by Avigad (2008, 2010). The core idea of Avigad's account is that "ascriptions of understanding are best understood in terms of the possession of certain abilities" (Avigad, 2008, p. 318). Accordingly, providing an account of proof understanding amounts to identifying and characterizing the relevant set of abilities.[16] The main difference between Avigad's account and the plan-based account lies then in the key distinction stated in the introduction between understanding as an epistemic state and understanding as an epistemic process. The plan-based account aims to characterize the epistemic process of proof understanding, i.e., what it takes for an agent to understand a proof. By contrast, Avigad's account aims to characterize the epistemic state of proof understanding, without necessarily specifying what an agent should do to end up in such a state. For this reason, the two accounts are perfectly compatible, shedding light on two sides of the same coin. Indeed, there is an intimate connection between the two, for the process of rational reconstruction of a proof's plan requires itself the exercise of specific abilities, the main one being the ability for practical reasoning in the context of proving, an ability which itself can be decomposed into more specific abilities. An agent who understands a proof according to the plan-based account will then necessarily possess a certain set of abilities. Indeed, several of the abilities identified by Avigad as associated with proof understanding appear to be directly connected with planning abilities, for instance, "the ability to give a high-level outline, or overview of the proof", "the ability to indicate 'key' or novel points in the argument, and separate them from the steps that are 'straightforward'", or "the ability to 'motivate' the proof, that is, to explain why certain steps are natural, or to be expected" (Avigad, 2008, p. 328).[17] These two accounts are then best thought as complementary than opposite.

In the next two sections, we will be concerned with evaluating the plan-based account of proof understanding. In section 4, we will argue that the plan-based account can correctly diagnose failures of proof understanding and can propose treatments to overcome them. In section 5, we will argue that the plan-based account can explain a significant range of different phenomena commonly associated with proof understanding.

[16] Avigad (2008, pp. 327–328) provides an extensive list of some of the abilities commonly associated with proof understanding. We will come back to it in section 7.1.

[17] Section 5 will say a bit more on how these different abilities are related to proof understanding as conceived by the plan-based account.

4 Evaluating the Account: Diagnosing and Treating Understanding Failures

One way to evaluate an account of proof understanding is to assess its capacity to correctly diagnose failures of understanding and to prescribe successful cures. In other words, an account of proof understanding should be able to identify the cause of an understanding failure—the obstacles that block the understanding process—and to propose a way to overcome them. According to the plan-based account, an agent has failed to understand a mathematical proof P whenever she did not manage to rationally reconstruct the plan underlying P. This means that she may have failed to reconstruct the tree structure of the plan underlying P (the tracking condition), and/or she may have failed to reproduce by herself all the instances of practical reasoning involved in the plan construction (the rationality condition). Such understanding failures can then be treated by providing the agent with additional information enabling her to overcome the obstacles she encountered in trying to meet the tracking and/or the rationality conditions. The examples that most dramatically illustrate understanding failures are mathematical proofs that are easy to verify but hard to understand. In this section, we will examine one mathematical proof of this kind. We will argue that our account provides correct diagnoses and prescribes successful cures for the potential failure(s) of understanding highlighted by this example. This case study will also be the occasion to illustrate further the concrete functioning of the plan-based account on a specific example.

4.1 A Case Study: McKay's Proof of Cauchy's Group Theorem

Cauchy's group theorem is a classic of a first course in group theory, usually paving the way to the more general Sylow theorems. In its most common formulation, the theorem states that, given a finite group G and a prime number p, if p divides the order of G, then G has an element of order p.[18] The following concise proof is due to James H. McKay, here reproduced verbatim from McKay (1959) with his own formulation of the theorem:

[18]The order of a finite group is the cardinality of its underlying set. The order of an element $a \in G$ is the smallest positive integer p such that $a^p = 1$ where 1 is the identity element of G.

Theorem 1 (Cauchy's Group Theorem)**.** *If the prime p divides the order of a finite group G, then G has kp solutions to the equation $x^p = 1$.*

Proof. Let G have order n and denote the identity of G by 1. The set

$$S = \{(a_1, \ldots, a_p) \mid a_i \in G,\ a_1 a_2 \ldots a_p = 1\}$$

has n^{p-1} members. Define an equivalence relation on S by saying two p-tuples are equivalent if one is a cyclic permutation of the other.

If all components of a p-tuple are equal then its equivalence class contains only one member. Otherwise, if two components of a p-tuple are distinct, there are p members in the equivalence class.

Let r denote the number of solutions to the equation $x^p = 1$. Then r equals the number of equivalence classes with only one member. Let s denote the number of equivalence classes with p members. Then $r + sp = n^{p-1}$ and thus $p|r$. □

McKay's proof is often presented with more details in group theory or algebra textbooks,[19] but its verification only requires rudimentary notions of group theory and discrete mathematics. That S has n^{p-1} elements comes from the fact that there are as many elements in S as $(p-1)$-tuples $(a_1, \ldots, a_{p-1})$ with $a_i \in G$. The reason is that, for any $(p-1)$-tuples $(a_1, \ldots, a_{p-1})$, there is one and only one p-tuple $(a_1, \ldots, a_{p-1}, a_p) \in S$ for a_p would have to be the unique inverse of $a_1 a_2 \ldots a_{p-1}$ in G. It is not hard to verify that the relation defined in terms of cyclic permutation is indeed an equivalence relation. Once it has been established that the resulting equivalence classes on S have either 1 or p elements,[20] the rest of the proof unfolds straightforwardly.

Although one may have succeeded in verifying the correctness of the proof, one may remain puzzled by it. In particular, in a first encounter with McKay's proof, one may not directly see how the different steps contribute to the main goal—i.e., to prove Cauchy's group theorem. It is only when one reaches the equation $r + sp = n^{p-1}$ towards the end of the proof that one can

[19] The proof is usually recast using the notion of group action (see, e.g., Fraleigh, 2003, pp. 322–323; Rotman, 1995, p. 74; Kurzweil and Stellmacher, 2004, pp. 62–63).

[20] The only step that is not straightforward here is to show that if two components of a p-tuple are distinct, then its equivalence class contains p members. It is in this part of the proof that the hypothesis that p is prime is used.

reconnect with the main objective. This may trigger a number of questions on the part of the reader, for instance: "Why introduce this particular set S of p-tuples in G?," "Why define this specific equivalence relation on S in terms of cyclic permutations?," "How did one think of counting the elements of S in these two different ways?" If a reader entertains questions of this kind after having successfully verified the proof and is not able to answer them, then this is an indicator that she does not fully understand it.

Our account would diagnose such a failure of understanding as a failure on the part of the reader to rationally reconstruct the plan underlying McKay's proof. To see what this means, let's put ourselves in such a reader's shoes and let's try to rationally reconstruct the plan underlying it. The root of the plan must be the proving intention to show the theorem at hand, that is, to show that G has kp solutions to the equation $x^p = 1$ from the hypotheses that G is a finite group and p is a prime that divides the order of G. How is this initial intention turned into more specific intentions organized in a subplan? As we have just noted, it is only when reaching the last two steps of the proof that we can see the connection with the main goal. From this, we can figure out that the last two proving intentions of the initial subplan are intentions to infer $n^{p-1} = r + sp$ from $\#S = n^{p-1}$ and $\#S = r + sp$, and then to infer $p|r$ from $n^{p-1} = r + sp$ and $p|n$. Working backward, we can realize that, for this plan to succeed, one would need to show that $\#S = n^{p-1}$ and $\#S = r + sp$, and this is exactly what McKay does in the proof.[21] From this, we can then see that the author has turned the initial proving intention into a subplan as follows:

$$
\begin{array}{rcll}
G \text{ is a finite group of order } n & \Rightarrow & p|r & \text{where } r \text{ is the number of solutions} \\
p \text{ is a prime that divides } n & & & \text{to } x^p = 1 \\
& \downarrow & & \\
\ldots & \Rightarrow & \#S = n^{p-1} & \text{where } S \text{ is defined as above} \\
\ldots & \Rightarrow & \#S = r \cdot 1 + s \cdot p & \text{where } r, s \in \mathbb{N} \\
\#S = n^{p-1} \text{ and } \#S = r + sp & \hookrightarrow & n^{p-1} = r + sp & \\
n^{p-1} = r + sp \text{ and } p|n & \hookrightarrow & p|r &
\end{array}
$$

[21] Addressing these two proving intentions requires further planning, and this is where we find the bulk of the proof.

At this stage, we have succeeded in tracking how the initial intention has been transformed into more specific intentions organized in a subplan. But can we also reproduce the instance of practical reasoning that takes the initial intention as input and the above subplan as output? If we were originally puzzled by the proof in the way described above, then it is unlikely that we could reproduce this instance of practical reasoning. For if we were able to do so, then presumably we would possess a reason to introduce the set S, to consider the equivalence relation in terms of cyclic permutations, and to count the elements of S in two different ways. In other words, we would be able to answer the questions mentioned above that were indicating a failure to understand the proof. This provides evidence for the hypothesis that the reader's failure to understand the proof originates in a failure to rationally reconstruct the plan underlying it, which here amounts to a failure to reproduce the first instance of practical reasoning in the plan construction.

What should the reader do in order to overcome this understanding failure? Our account makes a precise proposal on this issue: the reader needs to be in a position to reproduce the first instance of practical reasoning in the plan construction. Our account would then prescribe giving the reader relevant information enabling her to reproduce the above instance of practical reasoning. Interestingly, this seems to be exactly the information provided by McKay in the introduction to his proof:

> Since $ab = 1$ implies $ba = b(ab)b^{-1} = 1$, the identities are symmetrically placed in the group table of a finite group. Each row of a group table contains exactly one identity and thus if the group has even order, there are an even number of identities on the main diagonal. Therefore, $x^2 = 1$ has an even number of solutions.
>
> Generalizing this observation, we obtain a simple proof of Cauchy's theorem. (McKay, 1959, p. 119)

Let's see how this information may help the reader to reproduce the first instance of practical reasoning. We start again from the initial intention to prove the theorem:

$$p \text{ is a prime that divides the order of a finite group } G \quad \Rightarrow \quad G \text{ has } kp \text{ solutions to the equation } x^p = 1$$

McKay reminds us in the above passage that, when one wants to prove a general theorem of this kind, a rational move is to try to prove it for some simple cases. The simplest case here would be to try to prove the theorem for $p = 2$, that is:

G is a finite group of even order $\quad \Rightarrow \quad$ G has an even number of solutions to the equation $x^2 = 1$

Since we are working with finite groups and are interested in products of two elements, another rational move, also recalled by McKay, is to see what happens when we consider finite groups through their group tables (also called multiplication or Cayley tables). To make things even more concrete, we can look at the group table of a particular finite group. Here is the group table of $\mathbf{D_3}$, the dihedral group of order 6:

*	**1**	**a**	**b**	**c**	**d**	**f**
1	**1**	a	b	c	d	f
a	a	**1**	d	f	b	c
b	b	f	**1**	d	c	a
c	c	d	f	**1**	a	b
d	d	c	a	b	f	**1**
f	f	b	c	a	**1**	d

Table 1: Multiplication table of $\mathbf{D_3}$ the dihedral group of order 6

As McKay points out, looking at group tables such as the one above invites a number of observations. First, we can note that the solutions to the equation $x^2 = 1$ are those elements for which the identity element lies on the diagonal. If we add, as McKay noticed, that there is one and only one identity per row because of the unicity of the inverse in a group, then we can already realize that the number of identities on the diagonal is exactly the number of solutions to the equation $x^2 = 1$. Second, we can ask what happens when the identity is *not* on the diagonal. This is the first observation of McKay, namely that if $ab = 1$ then $ba = b(ab)b^{-1} = 1$, that is, the identities not on the diagonal come by pairs in the group table. Third, because we are considering finite groups of even order, we can see that there are an even number of rows in their group tables. Bringing these three observations together, we get that there is an even number of rows with exactly one identity per row, an even number of identities not on the diagonal, and so an even number of identities on the diagonal. Generalizing this sequence of observations to an arbitrary prime number p, we obtain the above subplan.

We have now established a sequence of rational moves that allowed us to connect the initial proving intention to the above subplan, that is, we have succeeded in reproducing the first instance of practical reasoning in the plan construction. According to our account, we have now overcome the previous understanding failure. And there is good evidence that this is indeed the case, for the agent now possesses answers to the questions raised above that were indicating her understanding failure. More specifically, the agent is now aware of a reason to introduce the set S, for S is simply the generalization of the idea to locate the identities in the multiplication tables—i.e., to consider the pairs of elements whose product yields the identity element—transposed from the case of pairs to the general case of p-tuples. The agent is also aware of a reason to introduce an equivalence relation on S in terms of cyclic permutations, for this comes from generalizing the first observation of McKay, namely that identities not on the diagonal come in pairs, which, in the case of products of p-tuples, leads to the hypothesis that p-tuples whose elements are not identical and whose product yields the identity element come in sets of p distinct members, namely sets of p-tuples which are all cyclic permutations of each other. Finally, the agent is now aware of a reason to count the elements of S in two different ways, for this is the direct generalization of the strategy employed in the case $p = 2$.

4.2 Wrapping Up

The objective of the previous case study was to evaluate the capacity of our account to diagnose failures of proof understanding—i.e., to pinpoint their cause or origin—and to identify what it takes to overcome them. According to the plan-based account, a failure in proof understanding originates in a failure to rationally reconstruct a proof's underlying plan, and overcoming this failure requires providing the agent with the relevant information enabling her to complete the process of rational reconstruction. The case study provides evidence in support of these predictions. First, we have seen that the failure of proof understanding highlighted in the previous example triggers the very same questions as a failure of rationally reconstructing the proof's underlying plan, which is to be expected if the former process boils down to the latter. Second, we have seen that our account can explain the origin of these questions, for those questions are exactly the ones that arise for an agent trying to rationally reconstruct the proof's underlying plan. Third, we have seen that this failure of proof understanding can be overcome by providing the agent with information enabling her to complete the process of rational reconstruction, as predicted by our account. The previous case study provides then an initial set of evidence that the process of understanding a proof amounts to that of rationally reconstructing its underlying plan.

5 Evaluating the Account: What the Account can Explain

Another important way to evaluate an account of proof understanding is in terms of its capacity to explain different phenomena commonly associated with understanding proofs. In this section, we will argue that the plan-based account can explain a significant range of such phenomena reported in the literature.

5.1 Understanding vs Verifying Proofs

Our investigations started with the contrast between understanding a proof and verifying it which we illustrated in the introduction with two paradigmatic quotes by Poincaré and Bourbaki. At the end of section 3, we already explained how the understanding process, as conceived by the plan-based account, differs from the verification process. It is nonetheless instructive to come back to Poincaré's and Bourbaki's quotes since both authors point out distinctive features of proof understanding. First, let us recall Poincaré's quote:

> Does understanding the demonstration of a theorem consist in examining each of the syllogisms of which it is composed in succession, and being convinced that it is correct and conforms to the rules of the game? [...]
>
> Yes, for some it is ; when they have arrived at the conviction, they will say, I understand. But not for the majority. Almost all are more exacting; they want to know not only whether all the syllogisms of a demonstration are correct, but why they are linked together in one order rather than in another. As long as they appear to them engendered by caprice, and not by an intelligence constantly conscious of the end to be attained, they do not think they have understood. (Poincaré, 1908, p. 118)

Poincaré provides here a general feature of the agent who understands a proof, namely that such an agent should know "why [the deductive steps of the proof] are linked together in one order rather than in another". Poincaré also provides a feature of the agent who does not understand, namely that, for such an agent, the chain of deductive steps constituting the proof appears as being "engendered by caprice, and not by an intelligence constantly conscious of the end to be attained".

The plan-based account of proof understanding can explain these two features. If an agent understands a proof, this means that she has been able to rationally reconstruct its underlying plan.

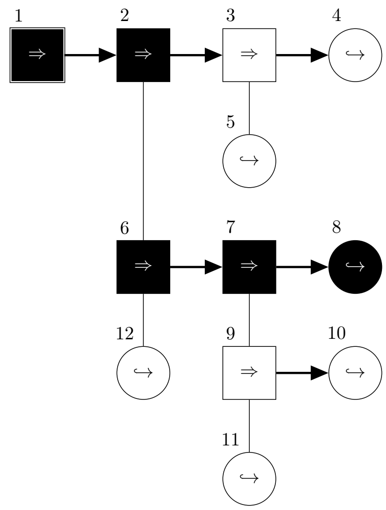

Figure 5.1: For Poincaré, the agent who understands must be able to see how each deductive inference in the proof is connected to "the end to be attained" through a path "engendered" by an "intelligence". In the plan-based account, the agent who understands is able to connect each deductive inference to the goal of proving the theorem at hand through an unbroken chain of practical decisions. As an example, we materialized this chain in the figure for the deductive inference corresponding to the proving intention number 8. Shall this chain be broken, the agent could not provide a complete explanation as to why this deductive step occurs in the proof.

For such an agent, all the deductive steps in the proof appear as the result of a chain of rational practical decisions going back to the initial intention of proving the theorem at hand. For each deductive step, the agent can then explain why this deductive step occurs at this particular stage of the proof by describing the chain of practical reasoning going from the initial proving intention to the subplan where the deductive step occurs. By contrast, when an agent does not understand, this means that she did not manage to rationally reconstruct the proof's underlying plan. This means, in particular, that for some or all deductive steps in the proof, the agent cannot locate them as part of a subplan in the overall plan, or the agent did not manage to reproduce some of the instances of practical reasoning leading to them in the plan construction. In both cases, the agent cannot relate them rationally to the main objective of proving the theorem at hand. In other words, the agent cannot trace back these deductive steps to the main objective through an uninterrupted chain of rational practical decisions. That the chain of rational practical decisions is broken for these deductive steps explains why they do not appear as being "engendered [...] by an intelligence constantly conscious of the end to be attained". The "intelligence" of Poincaré corresponds, in the plan-based account, to the agent's faculty of practical reason, the term "engendered" refers to the fact that the proof is nothing more than the execution of its underlying plan and so that each deductive step is "engendered" by practical decisions in the plan construction, and the expression "constantly conscious of the end to be attained" refers to the chain of practical decisions connecting each deductive step to the main objective of proving the theorem at hand. Whenever this chain is broken in the sense that the agent did not manage to reproduce it in its entirety, the agent is then blocked in her capacity to produce a full explanation as to why this step occurs in the proof (see Figure 5.1). In this case, such a deductive step will appear as mysterious, arbitrary, or "engendered by caprice".[22]

5.2 Understanding and Proof Ideas

Bourbaki's quote points out another important aspect of proof understanding:

> [E]very mathematician knows that a proof has not really been "understood" if one has done nothing more than verifying step by step the correctness of the deductions of which it is composed, and has not tried to gain a clear insight into the ideas which have led to the construction of this particular chain of deductions in preference to every other one. (Bourbaki, 1950, p. 223)

Bourbaki provides here another feature of the agent who understands a proof, namely that such an agent should possess "a clear insight into the ideas which have led to the construction of this particular chain of deductions in preference to every other one". How should the notion of "idea" be interpreted in this context? Ordinarily, when one describes an idea in a proof, one is led to describe a certain course of actions. For instance, in the case of McKay's proof (cf. section 4.1), one may say that the main idea is to count the set S in two different ways, once directly using the group-theoretic property that every element has a unique inverse, and once through its equivalence classes. When we describe a proof idea in this way, we are simply describing a subplan or a part thereof in the overall plan, that is, a course of actions meant to fulfill a specific proving intention. This interpretation of proof ideas as subplans, or parts of subplans, seems confirmed by the fact that, for Bourbaki, "ideas" are what "led to the construction of this particular chain of deductions in preference to every other one". Insofar as the "chain of deductions" is the sequence of actions that results from the execution of the plan, the various subplans in the overall plan are indeed the entities that specify the "deductions" or deductive inferences reported in the written proof.

[22] Pólya (1949) refers to such steps as "deus ex machina" and argues that they require "motivations" in mathematical communications. For a philosophical analysis of motivations in the context of mathematical proofs, see Morris (2020).

The plan-based account can then explain why the agent who understands a proof possesses a "clear insight" into the ideas of the proof. When an agent understands a proof, this means that she has been able to rationally reconstruct its underlying plan. This means that, for any subplan in the overall plan, the agent has been able to reproduce the instance of practical reasoning that led to this subplan. The agent possesses then what we may call a *rational* insight as to why this particular subplan occurs in the overall plan, and thus as to why the "chain of deductions" resulting from the execution of this particular subplan occurs in the proof. If a proof idea is nothing more than a subplan or a part thereof, then the agent who understands a proof indeed possesses a "clear insight" into the ideas of the proof.

5.3 Understanding and Proof Architectures

When one understands a mathematical proof, one is able to grasp the connections between the different parts of the proof and to see how each of them contributes to the main objective—i.e., to establish the mathematical proposition at hand. By contrast, the agent who does not understand only sees in the proof a monotonous sequence of deductive steps ending miraculously with the proposition to be established. This dimension of proof understanding is a common place in mathematical practice. Janet Folina identifies it in the views of several authors:

> Poincaré and Feferman both draw attention to the importance of the *plan* of the proof; the unity that is found in its "architecture": how the parts are related and connected to one another. Emphasizing unity, and relationships between parts, resonates with Haylock's emphasis on making connections as central to mathematical understanding. In addition, the importance of seeing the whole, or the blueprint, of an argument also calls to mind Michener's view that mathematical understanding enables one to avoid getting lost in details.[23] (Folina, 2018, p. 136)

Thus, the agent who understands a proof sees an order or a structure in the sequence of deductive steps. She is able to see how the proof is organized into parts and how the different parts of the proof are connected—this organization or structure is what Poincaré calls the "architecture" of a proof. The agent who understands a proof is then able to grasp its architecture. As Folina noted,

[23] The references associated with the authors mentioned in this passage and cited in Folina (2018) are Poincaré (1900/1996, 1908), Feferman (2012), Haylock and Cockburn (2008), and Michener (1978).

this notion of proof architecture is intimately connected with the notion of proof plan. This is in direct line with our perspective: it is precisely in virtue of being produced by rational planning agents, and thus of possessing a rational plan, that mathematical proofs can be said to possess an architecture.

Because the plan-based account of proof understanding establishes a direct connection between understanding and proof architecture via the notion of proof plan, it can readily explain why understanding a proof is related to grasping connections between its different parts and to seeing how each part contributes to the main objective. First, it should be noted that the plan-based account yields a precise notion of proof part: a part of a proof is simply a subtree in the overall plan (see Figure 5.2).[24] A proof part is then a plan in itself taking care of a specific proving intention in the overall plan. For instance, in the toy example described in section 3, one part of the proof consists in showing that if n is even, then $n^2 + 3n$ is even, while another part of the proof consists in showing that if n is odd, then $n^2 + 3n$ is even. Here, the key observation is this: because different proof parts are always subtrees of the same tree, they are always connected in the overall plan. More specifically, given any two proof parts, one can be contained in the other, or they may be distinct. If they are distinct—i.e., if they do not have any node or proving intention in common—then there is always a bigger subtree containing both parts as subtrees (see Figure 5.2). Thus, whatever connections there are to grasp between different parts of the proof, they are to be found within the proof's underlying plan.

Now, if an agent understands a mathematical proof, this means that she has been able to rationally reconstruct its underlying plan. This means, in particular, that the agent has succeeded in reconstructing the tree structure of the plan. It follows that given two or more different parts of the proof, the agent can always tell how they relate to each other. For instance, the agent who understands McKay's proof will know that the two distinct proof parts consisting in showing that $\#S = n^{p-1}$ and that $\#S = r \cdot 1 + s \cdot p$ are both parts of the same bigger proof part which aims to show that $p|r$.

But there is more: because the agent has *rationally* reconstructed the plan, she not only knows how the different parts of the proof are connected, she also knows *why* they are connected in these specific ways. That is, the agent can *explain* any connection by describing the corresponding chain

[24]This notion of proof part is intimately connected with the notion of *module* as introduced by Avigad (2020) in the context of mathematical proofs.

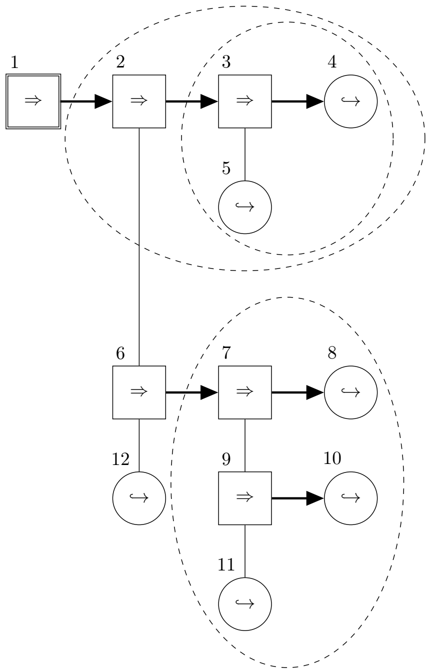

Figure 5.2: A proof part is a subtree in the overall plan. The three subtrees circled in the figure are examples of three different proof parts. Given any two proof parts, they are either distinct or one is included in the other. When two proof parts are distinct, they are always both parts of a bigger proof part. In the plan-based account, the architecture of a proof is the structure composed of all the proof parts together with the connections between them.

of practical reasoning.[25] If a proof part is contained in another, the agent can explain the different instances of practical reasoning that connect the root of the smaller part to that of the bigger part. If two proof parts are distinct, the agent can explain the instances of practical reasoning that connect the roots of the two parts to the root of the bigger part they are both contained in. In particular, because any proof part is always a subtree of the overall plan, the agent can always explain the connection between any proof part and the main or initial intention of proving the theorem at hand. In this sense, the agent knows how each part contributes to the main objective.

In sum, being able to grasp the connections between the different parts of a proof and seeing how they contribute to the main objective appear as a direct consequence of having succeeded in rationally reconstructing the proof's underlying plan, that is, as a direct consequence of having understood the proof as conceived by the plan-based account of proof understanding.

[25]This ability resonates directly with "the ability to 'motivate' the proof, that is, to explain why certain steps are natural, or to be expected" that Avigad (2008, p. 328) associated with proof understanding. On what it means to 'motivate' a proof, see also Morris (2020).

5.4 Understanding and Proof Details

Written mathematical proofs, especially in published articles, most often contain a wealth of technical details necessary to establish the associated mathematical claims. Understanding a proof requires somehow abstracting from the details in order to identify the organizational structure of the proof, to see the forest from the trees. The logician and computer scientist John Alan Robinson put the point as follows:

> When we survey a real proof in [a] high-level, outline way, and fix our attention on its main idea or ideas, we can better intuitively appreciate its overall plan. We understand the proof as an explanation, in a sense, even though our view of it is neither rigorous nor complete. With only the overall plan before it, the mind is not concerned with the details. For the purpose of obtaining a (higher-level) understanding, it even seems essential that the (lower-level) details should be ignored. If too many details enter into the primary sketch, we simply lose sight of the main architecture of the proof—we are unable to see the big picture.[26] (Robinson, 2000, p. 292)

The plan-based account can explain why managing and ignoring the details is important for proof understanding and why proof understanding can be more or less difficult depending on the quality of the proof presentation, and especially on the amount of details and the way they are organized. The main point here is that the plan-based account describes proof understanding as a top-down process: in the rational reconstruction of the plan, the most natural way to proceed is to work downward from the root of the tree, that is, to figure out how proving intentions are turned into more specific intentions starting from the initial intention of proving the theorem at hand. But to reconstruct the topmost part of the plan, one needs to identify the topmost proving intentions from the written proof. In this respect, all the deductive steps reported in the written proof that contribute to lower-level proving intentions are irrelevant and should be ignored. These are precisely the "details" of the proof, namely the deductive steps whose corresponding proving intentions 'to infer' are situated in the lower part of the tree—the plan-based account provides then a precise

[26] The aspects of proof understanding highlighted in this passage resonate strongly with two of the abilities that Avigad associated with proof understanding, namely "the ability to give a high-level outline, or overview of the proof" and "the ability to indicate 'key' or novel points in the argument" (Avigad, 2008, p. 328). On this point, see also section 7.1 below.

characterization of what the “details” of a proof are. Thus, obtaining what Robinson calls a “higher-level” understanding of the proof is, from this perspective, the very first stage of the understanding process which indeed requires ignoring the “lower-level” details. If one has obtained such a “higher-level” understanding, this means that one has succeeded in reconstructing the topmost part of the plan, and so that one has grasped the “main architecture” of the proof, i.e., how the main parts of the proof are connected to each other.

This task can be made more or less difficult depending on the quality of the proof presentation. If the written proof contains a lot of details but does not present them in an organized or structured way, this can make the understanding process particularly tedious. By contrast, a well-organized or well-structured proof can strongly facilitate the understanding process. J.A. Robinson particularly emphasized this point:

> [F]ew written mathematical expositions are carefully designed to assist the mind in this respect. Written expositions of a proof are too often merely dumps of the details, in all their complexity, with little or no guidance as to the conceptual organization underlying them, let alone the informal intuitions from which they may have sprung. (Robinson, 2000, p. 292)

According to the plan-based account, the best way to present a mathematical proof in order to ease understanding is to make the rational reconstruction of the underlying plan as easy as possible. This means, in particular, to facilitate the localisation of each deductive step within the tree structure of the plan. Clearly, written proofs that are “dumps of the details” will make this localisation process particularly tedious. By contrast, written proofs that make the organization or the structure of the proof salient—for instance by making the different parts of the proof easily identifiable and by highlighting the connections between them—will make it easier to reconstruct the tree structure of the plan.[27]

In sum, from the perspective of the plan-based account, managing and ignoring the details is essential to the understanding process since it is essential to reconstruct the topmost part of the plan—the very first stage of the understanding process. The agent who succeeded in reconstructing

[27] Determining the best ways to present a proof in order to facilitate understanding is an interesting issue with important practical applications for mathematical communication and mathematics education. On this theme, see the work of Leron (1983) and Lamport (1995, 2012).

the topmost part of the plan possesses then a “higher-level” understanding of the proof: she sees how the main parts are connected to each other, that is, she grasps the “main architecture” of the proof. This process can be made more or less difficult depending on the way the proof details are organized in the proof presentation.

6 Comparison with other Philosophical Accounts of Objectual Understanding

In this work, we have delved in some depth into one particular, yet fundamental instance of objectual understanding. How does the resulting plan-based account of proof understanding relates to existing accounts of objectual understanding?

It should first be observed that objectual understanding is a big category. One can understand many kinds of things of very different nature. This is acknowledged by most authors writing on objectual understanding, for instance:

> We understand rules and reasons, actions and passions, objectives and obstacles, techniques and tools, forms, functions, and fictions, as well as facts. We also understand pictures, words, equations, and patterns. Ordinarily these are not isolated accomplishments; they coalesce into an understanding of a subject, discipline, or field of study. (Elgin, 1996, p. 123)

> Understanding has a multitude of appropriate objects, among them complicated machines, people, subject disciplines, mathematical proofs, and so on. (Riggs, 2003, p. 217)

However, although philosophical accounts of objectual understanding usually aim at a maximum of generality, they often adopt a specific conception of what sort of thing X is in “understanding X”.

In analytic epistemology, accounts of understanding X often consider that X is some *subject matter* (e.g., Elgin, 1996, 2009; Kvanvig, 2003; Carter and Gordon, 2014). Typical examples of subject matters are the American Civil War, the European Union law, Quantum mechanics, or the Comanche dominance of the southern plains in the nineteenth century. In this sense, a subject matter is a body or network of propositions with specific relations between them. Mathematical proofs can be conceived as subject matters in a narrow and a broad sense. In a narrow sense, one would only consider the mathematical propositions that figure in the proof and the relations among them. One type of relation between propositions in a proof are logico-deductive relations between the conclusion and premisses of deductive inferences. Grasping these relations amounts to verifying the proof, i.e., seeing that the proof is correct or rigorous. Another type of relation comes from the fact that all the propositions in a proof are related in the proof's underlying plan through the associated proving intentions. Grasping these relations precisely amounts to grasping the underlying plan, that is, to understanding the proof according to the plan-based account. In this respect, the plan-based account bears a strong resemblance to epistemological accounts that consider that understanding a subject matter amounts to grasping *coherence* relations (e.g., Kvanvig, 2003; Carter and Gordon, 2014). In a broad sense, one would also consider the propositions that are about the proof itself, for instance propositions related to its author(s), its discovery, the context in which it emerges, the relations with other mathematical theories, concepts, proofs, etc. If one takes any book on Wiles' proof of Fermat's last theorem (e.g., Singh, 1997), it will be about the proof in precisely this sense, that is, the proof in this broader sense is the subject matter of the book. It is likely that many of these additional propositions will be related to the proof's conception and planning, and this could be accounted for by the plan-based account. But many other propositions may not—one example would be biographical information on the author(s). A promising starting point to develop an account of proof understanding where proofs are taken in this broader sense would be to combine the plan-based account with existing epistemological accounts of understanding subject matters.

In philosophy of science, accounts of understanding X often consider that X is some *phenomenon* (e.g., de Regt and Dieks, 2005; Lipton, 2009; Strevens, 2013; Kelp, 2015; de Regt, 2017; Khalifa, 2017; Dellsén, 2020). Most of these accounts invoke a specific proxy for the phenomenon, for instance a model (Dellsén, 2020), a theory (de Regt and Dieks, 2005; de Regt, 2017),[28] or an explanation (Lipton, 2009; Strevens, 2013; Khalifa, 2017; de Regt, 2017) of the phenomenon. A typical example is de Regt's account of scientific understanding:

> A phenomenon P is understood scientifically if and only if there is an explanation of P that is based on an intelligible theory T and conforms to the basic epistemic values of empirical adequacy and internal consistency. (de Regt, 2017, p. 92)

[28]For de Regt and Dieks (2005) and de Regt (2017), the notion of model also plays a key role as a proxy for achieving scientific understanding.

It is, of course, very common for agents in mathematical practice at all levels to request explanations of mathematical proofs—this could be from a mathematician to a colleague or from a mathematics student to its teacher. In this sense, a mathematical proof could be conceived as a phenomenon that can be explained. From the perspective of the plan-based account, such explanations would consist in further describing the plan and planning underlying the proof under consideration. This type of explanation will then provide information on how to rationally reconstruct the proof's underlying plan, thereby allowing the agent who requested the explanation to understand the proof—we recover in this way a natural correspondence between proof explanation and proof understanding. Further work is called for to properly articulate this notion of proof explanation and to see how it relates to existing accounts of mathematical and scientific explanations.

Some accounts of understanding X have adopted a more abstract stance by considering that X is a *system* or *structure* of some sort (e.g., Grimm, 2011, 2017; Zagzebski, 2001, 2019). In this context, a system or structure is an entity which is composed of multiple parts that relate to and depend upon one another in various ways. According to Grimm, understanding a system or structure amounts precisely to grasping these dependence relations in the sense that one needs to "appreciate how the structure 'works', or how changes in its various parts will lead, or fail to lead, to changes in other parts" (Grimm, 2021, section 2.2). For Zagzebski, we understand an object when we grasp its structure, which allows us to see the object *as* an object—an object's unity is then given by its structure. The core idea of Zagzebski's account is that "understanding is the grasp of structure" (Zagzebski, 2019, p. 124).[29] Thinking of objectual understanding in terms of grasping systems or structures resonates strongly with our previous discussion on understanding and proof architecture (see section 5.3). There, we pointed out a common observation on proof understanding, namely that when one understands a proof, one is able to grasp the connections between the different parts of the proof and to see how they contribute to the main objective of establishing the mathematical proposition at hand. In the plan-based account, we were able to explain this observation in terms

[29]Because of the focus on structures, there is here a direct connection between the views of Grimm and Zagzebski and the conceptions of mathematical understanding advanced by Cellucci (2015) and Folina (2018) in the philosophy of mathematics. More specifically, Cellucci holds that understanding a "piece of mathematics" involves "the recognition of the fitness of the parts to each other and to the whole" (Cellucci, 2015, p. 344). Folina has explicitly adopted what she called a "structuralist perspective" to mathematical understanding where understanding requires to "comprehend [. . .] wholes" or "structures" (Folina, 2018, p. 136). This structuralist perspective is then fleshed out by building on the structuralist view of mathematics.

of the modular and hierarchical structure of proof plans. An account of proof understanding in the vein of Grimm's and Zagzebski's conceptions of objectual understanding would directly explain this common observation. The interesting question is whether such an account would also be able to explain all the other phenomena associated with proof understanding discussed in section 5.

7 Critical Assessment of the Account

In this section, we provide a critical assessment of the plan-based account. We will discuss the scope of the account and what the account can say on the understanding of proofs generated by artificial intelligence systems.

7.1 Assessing the Scope of the Plan-Based Account

The plan-based account is an attempt to capture the notion of proof understanding hinted at by Poincaré and Bourbaki in the quotes reported in the introduction. It is entirely possible that there is more to this notion of proof understanding and/or that there are other notions of proof understanding present in mathematical practice. For this reason, it is important to assess the scope of the plan-based account as an analysis of proof understanding and to identify what may or may not be accommodated by it. A quick test to initiate this assessment is to start from the list of abilities advanced by Avigad (2008) as being commonly associated with proof understanding and to evaluate the extent to which the plan-based account can explain them:

> In ordinary circumstances, when we say, for example, that someone understands a particular proof, we may take them to possess any of the following:
>
> 1. the ability to respond to challenges as to the correctness of the proof, and fill in details and justify inferences at a skeptic's request;
> 2. the ability to give a high-level outline, or overview of the proof;
> 3. the ability to cast the proof in different terms, say, eliminating or adding abstract terminology;
> 4. the ability to indicate 'key' or novel points in the argument, and separate them from the steps that are 'straightforward';

5. the ability to 'motivate' the proof, that is, to explain why certain steps are natural, or to be expected;
6. the ability to give natural examples of the various phenomena described in the proof;
7. the ability to indicate where in the proof certain of the theorem's hypotheses are needed, and, perhaps, to provide counterexamples that show what goes wrong when various hypotheses are omitted;
8. the ability to view the proof in terms of a parallel development, for example, as a generalization or adaptation of a well-known proof of a simpler theorem;
9. the ability to offer generalizations, or to suggest an interesting weakening of the conclusion that can be obtained with a corresponding weakening of the hypotheses;
10. the ability to calculate a particular quantity, or to provide an explicit description of an object, whose existence is guaranteed by the theorem;
11. the ability to provide a diagram representing some of the data in the proof, or to relate the proof to a particular diagram;

and so on. (Avigad, 2008, pp. 327–328)[30]

The abilities listed here can be divided into three categories. The first category are abilities that have to do with proof plans and their rational reconstruction. The abilities 2, 4, 5, 7 and 8 fall under this category. The ability to "give a high-level outline, or overview of the proof" (number 2) is precisely what we discussed in section 5.4, while the abilities to "indicate 'key' or novel points in the argument, and separate them from the steps that are 'straightforward'" (number 4) and to "'motivate' the proof, that is, to explain why certain steps are natural, or to be expected" (number 5) are discussed in section 5.2. Ability 7 is also concerned with the decisions behind the choice of each proof step but this time with a focus on logico-deductive relations and the necessity or not of certain hypotheses for the validity of individual steps. It reminds us of the obvious fact that proof planning aims at the production of valid proofs and thus that practical reasoning requires a

[30]In Avigad (2008), the list was presented with bullet points. Here it is numbered to ease subsequent references to the different abilities.

dedicated attention to the hypotheses that are needed or not for a proof step to be valid.[31] Ability 8 is as well directly connected with the capacity for practical reasoning in the context of proof planning (see section 2). All these abilities can be explained by the plan-based account.

The second category are abilities connected with proof verification. The abilities 1 and 9 fall under this category. Although it is sometimes said that an agent "understands" a proof when she has managed to verify it and see that it is correct, from a methodological point of view it seems better to analyze separately the processes of understanding and verification. The abilities 1 and 9 are thus better accommodated within an account of proof verification rather than an account of proof understanding.

The third category are abilities connected with general mathematical proficiency. The abilities 3, 6, 10 and 11 fall under this category. One potential reason why these abilities are commonly associated with proof understanding is that understanding is likely to be a pre-condition for exercising these abilities. For instance, it is hard to see how an agent could be able to "cast the proof in different terms" (number 3) or "give natural examples of the various phenomena described in the proof" (number 6) without understanding the proof under consideration in the first place. But these abilities seem to mainly depend and reflect the level of mathematical proficiency and literacy of the agent. The main question here is whether these abilities should be accommodated within an account of proof understanding, or whether they are better analyzed in a separate way as part of a general account of mathematical proficiency and literacy.

[31] Interestingly, Avigad's formulation of ability 7 in terms of what *would* happen "when various hypotheses are omitted" resonates with Grimm's modal account of understanding in terms of grasping structures where "'[g]rasping' a structure [. . .] seem to bring into play something like a modal sense or ability—that is, an ability not just to register how things are, but also to anticipate how certain elements of the system would behave, were other elements different in one way or another" (Grimm, 2011, p. 89). In the present case, the structure under consideration would be the logico-deductive structure of a mathematical proof. Although we did not put as a necessary requirement of proof understanding that the agent be able to see what would happen if certain aspects of a proof and its underlying plan had been different, mathematical agents with a certain level of expertise would usually be able to anticipate the consequences of such changes. For instance, a mathematical agent may be able to evaluate whether a given subplan may or may not work if a given proving intention would be different—e.g., by evaluating whether a given mathematical method can also be applied to a variation of the proposition to be proved. So a mathematical agent with sufficient expertise who understands a mathematical proof will usually grasp its underlying plan in the modal sense proposed by Grimm, i.e., will possess the additional ability to anticipate what would happen if the proof and its underlying plan had been different.

The quick test performed here already yields useful observations on the scope of the plan-based account. In particular, it reminds us of the obvious fact that the process of proof understanding is connected with many other epistemic processes (verification, explanation, communication, etc.) and that the capacity of proof understanding is related to many other mathematical abilities. This raises two important questions: What exactly should be part of an account of proof understanding? What are the relations between the processes and capacity of proof understanding and the many other epistemic processes and capacities at play in mathematical practice? Although we are not answering these questions here, the plan-based account offers a starting point to address them.

7.2 Understanding Proofs Generated by Artificial Intelligence Systems

The fast-growing development of powerful artificial intelligence systems and their potential capacity to produce mathematical proofs raises the following interesting question: Can human mathematical agents understand these proofs, and if yes what would be the nature of this understanding process? Probably the most interesting cases to look at are mathematical proofs generated by large language models (LLMs) such as the OpenAI's GPT models—the ones used, e.g., in ChatGPT. In the same way that LLMs can generate texts that are very much human-like—e.g., in the context of a conversation for ChatGPT—it is perfectly plausible that they could also generate mathematical proofs that look as if they were produced by human mathematical agents. Indeed, some initial works have recently been done in this direction and some interesting results have already been obtained (see, e.g., Lample and Charton, 2020; Welleck et al., 2021, 2022; Wei et al., 2022). For instance, Welleck et al. (2022) have developed a system called NaturalProver which they describe as "a language model that generates proofs by conditioning on background references (e.g., theorems and definitions that are either retrieved or human-provided), and optionally enforces their presence with a constrained decoding algorithm that leverages the multi-step structure of proofs" (Welleck et al., 2022, p. 2). NaturalProver is designed to either make suggestions for the next step in a proof or to generate a full proof. In their experiments, NaturalProver was able to generate short proofs such as the following one of the triangle inequality for complex numbers (Welleck et al., 2022, p. 23):

> By induction:
>
> For all $n \in \mathbb{N}_{>0}$, let $\mathbf{P}(n)$ be the proposition: $|z_1 + z_2 + \cdots + z_n| \leq |z_1|+|z_2|+\cdots+|z_n|$

$\mathbf{P}(1)$ is true, as this just says $|z_1| \leq |z_1|$.

Basis for the Induction

$\mathbf{P}(2)$ is the case: $|z_1 + z_2| \leq |z_1| + |z_2|$

which has been proved in Triangle Inequality for Complex Numbers.

This is our basis for the induction.

Induction Hypothesis

Now we need to show that, if $\mathbf{P}(k)$ is true, where $k \geq 2$, then it logically follows that $\mathbf{P}(k+1)$ is true.

So this is our induction hypothesis: $|z_1 + z_2 + \cdots + z_k| \leq |z_1| + |z_2| + \cdots + |z_k|$

Then we need to show: $|z_1 + z_2 + \cdots + z_{k+1}| \leq |z_1| + |z_2| + \cdots + |z_{k+1}|$

Induction Step

This is our induction step:

$$\begin{aligned}
|z_1 + z_2 + \cdots + z_{k+1}| &= |z_1 + z_2 + \cdots + z_k + z_{k+1}| \\
&\leq |z_1 + z_2 + \cdots + z_k| + |z_{k+1}| \\
&= |z_1 + z_2 + \cdots + z_k| + |z_{k+1}| \\
&\leq |z_1| + |z_2| + \cdots + |z_k| + |z_{k+1}| \\
&= |z_1| + |z_2| + \cdots + |z_k| + |z_{k+1}|
\end{aligned}$$

So $\mathbf{P}(k) \Rightarrow \mathbf{P}(k+1)$ and the results follows by the Principle of Mathematical Induction.

Therefore: $\forall n \in \mathbb{N} : |z_1 + z_2 + \cdots + z_n| \leq |z_1| + |z_2| + \cdots + |z_n|$

□

In this particular case, it seems perfectly possible for a human mathematical agent to understand this proof, in the same way as the texts produced by ChatGPT are most of the time perfectly understandable by humans.

Can the plan-based account explain what it means for a human mathematical agent to understand such a LLM-generated proof even though the proof was presumably not generated through a planning process? Yes, the approach would here be the same as the one described in the clarificatory remarks of section 3 regarding the possibility of understanding proofs that were not generated by rational planning agents. The main idea is that, in such cases, the agent who aims to understand proceeds *as if* the proof was produced by a rational planning agent. This is a perfectly plausible hypothesis given that what is so striking with the texts generated by ChatGPT or the above small proof generated by NaturalProver is how human-like they are. For this reason, it is perfectly plausible that the understanding processes humans engage in when trying to understand the outputs of these artificial intelligence systems are very similar to the understanding processes they would engage in when trying to understand the same outputs if they were produced by human agents. Interestingly, when Welleck et al. (2022) are commenting on the above proof they say that "the model has learned the format of proof by induction and can apply it in [a] new context" even though "[t]he model did not see a similar proof during training" (Welleck et al., 2022, p. 23). From the perspective of the plan-based account, proof by induction is a paradigmatic case of practical reasoning in which a proving intention is turned into more specific proving intentions. Thus, when commenting on the above proof, one may say that Welleck et al. (2022) are proceeding *as if* the proof was generated by a human mathematical agent. If this hypothesis is correct, that is, if humans proceed this way when trying to understand proofs generated by artificial intelligence systems, then one can apply all the resources of the plan-based account to analyze what it means for humans to understand such proofs.

Of course, we can only hint here at the complex issue of the potential interaction(s) between humans and proofs generated by artificial intelligence systems—this question will certainly attract a lot of attention in the years to come, in mathematics, computer science and philosophy. The *as if* approach just described already offers one possible way in which the plan-based account could explain some aspects of what happens when a human mathematical agent is trying to understand a proof generated by an artificial intelligence system.

8 Conclusion

In this work, we have tried to articulate a precise and systematic account of what it means to understand a mathematical proof. Our focus has been on the active process of understanding, that is, on what it takes on the part of the agent to understand a proof. The epistemic state of understanding a proof was then conceived as the state that results from having successfully completed the associated understanding process. We have shown that our account was able to diagnose failures of understanding and to propose ways to overcome them. We have also argued that the account can explain a significant range of distinctive phenomena commonly associated with proof understanding.

The capacity to understand mathematical proofs is essential to any agent engaged in mathematical practice, but there are also many other things that mathematical practitioners need to understand: theories, methods, concepts, problems, solutions, theorems, definitions, to mention some of the main ones. Could some of the ideas developed in the plan-based account of proof understanding be adapted to account for the understanding of other mathematical entities? One central idea guiding the plan-based account is that understanding a mathematical proof is intimately connected with how it has been *designed* or *engineered* by one or more rational human agents—as we have seen, this idea is at the core of Poincaré's and Bourbaki's descriptions of proof understanding. In this respect, we can notice that all the other mathematical entities just mentioned are themselves produced by rational human agents.[32] It is thus a perfectly plausible hypothesis that the understanding of these mathematical entities is itself connected to the way they have been designed or engineered. Pursuing an approach to mathematical understanding along this line appears then intimately connected with the project of developing a theory of *mathematical design* as promoted, for instance, by Avigad (2020) and Marquis (2015)—the aim of such a theory is to do for mathematics what others have done for mechanical engineering, software development, architecture, etc. Now, there is also an intimate connection between design and agency as pointed out, for instance, by philosophers of technology (see, e.g., Franssen et al., 2018, sections 2.3 and 2.4). This is due to the fact that "design is a form of action, a structured series of decisions to proceed in one way rather than another", from which it follows that "the form of rationality that is relevant to it is practical rationality, the rationality incorporating the criteria on how to act, given particular circumstances" (Franssen et al., 2018, section 2.4). Thus, from this perspective, an account of mathematical understanding should build on an account of mathematical design, which itself should build on an account of mathematical agency. This is exactly the approach implemented in the plan-based account of proof understanding.

[32]Putting aside mathematical entities that may have been produced by computers.

Interestingly, the idea to build a theory of mathematical understanding on a theory of mathematical agency has a direct counterpart in the philosophy of science where it has been pointed out that the development of a theory of scientific understanding will itself require to attend to scientific agency. For instance, the editors of the collected volume entitled *Scientific Understanding: Philosophical Perspectives* (de Regt et al., 2009) remarked that:

> All chapters in this volume firmly agree on linking understanding to cognition. [...] This emphasis on the cognizant individual involves a reevaluation of the epistemic role of human agency in producing, disseminating, and using scientific knowledge. To understand scientific understanding, philosophers must find ways to study and analyze scientific agency. This means taking scientific practices seriously, for arguably a study of agency in science needs to be based on knowledge of how scientists across various fields actually act. (de Regt et al., 2009, p. 14)

There is then an interesting prospect for a general methodology common to the study of mathematical and scientific understanding focused on agency. To this end, resources from the philosophy of action—e.g., on planning agency, intentional agency, and practical rationality—can appear to be directly useful, as we witnessed in the present work. Furthermore, and as pointed out by de Regt et al. (2009), such a methodology would imply paying particular attention to mathematical and scientific practices. This would provide, in turn, suitable grounds for fruitful interactions with other disciplines concerned with the study of mathematical and scientific practices, such as the history, sociology, and pedagogy of mathematics and the sciences. Thus, it seems that there is much to be gained for philosophers and researchers from these different horizons to join forces in order to build a general and systematic methodology to the study of human understanding focused on human agency and directly informed by empirical studies of the relevant intellectual practices.

References

Anonymized. Details omitted for blind reviewing. forthcoming.

Jeremy Avigad. Understanding proofs. In Paolo Mancosu, editor, *The Philosophy of Mathematical Practice*, pages 317–353. Oxford University Press, Oxford, 2008.

Jeremy Avigad. Understanding, formal verification, and the philosophy of mathematics. *Journal of the Indian Council of Philosophical Research*, 27:161–197, 2010.

Jeremy Avigad. Modularity in mathematics. *The Review of Symbolic Logic*, 13(1):47–79, 2020.

Christoph Baumberger, Claus Beisbart, and Georg Brun. What is understanding? An overview of recent debates in epistemology and philosophy of science. In Stephen R. Grimm, Christoph Baumberger, and Sabine Ammon, editors, *Explaining Understanding: New Essays in Epistemology and the Philosophy of Science*, pages 1–34. Oxford University Press, New York, 2017.

Nicholas Bourbaki. The architecture of mathematics. *The American Mathematical Monthly*, 57(4): 221–232, 1950.

Michael E. Bratman. *Intention, Plans, and Practical Reason.* Harvard University Press, Cambridge, MA, 1987. Reissued by CSLI Publications, Stanford CA, 1999. (Citations are to the latter edition).

Michael E. Bratman. Agency, time, and sociality. *Proceedings and Addresses of the American Philosophical Association*, 84(2):7–26, 2010.

Alan Bundy. The use of explicit plans to guide inductive proofs. In Ewing Lusk and Ross Overbeek, editors, *9th International Conference on Automated Deduction*, volume 310 of *Lecture Notes in Computer Science*, pages 111–120, Berlin, 1988. Springer-Verlag.

Alan Bundy, David Basin, Dieter Hutter, and Andrew Ireland. *Rippling: Meta-Level Guidance for Mathematical Reasoning*, volume 56 of *Cambridge Tracts in Theoretical Computer Science.* Cambridge University Press, Cambridge, 2005.

J. Adam Carter and Emma C. Gordon. Objectual understanding and the value problem. *American Philosophical Quarterly*, 51(1):1–13, 2014.

Carlo Cellucci. Mathematical beauty, understanding, and discovery. *Foundations of Science*, 20(4): 339–355, 2015.

Henk W. de Regt. *Understanding Scientific Understanding.* Oxford University Press, New York, 2017.

Henk W. de Regt and Dennis Dieks. A contextual approach to scientific understanding. *Synthese*, 144(1):137–170, 2005.

Henk W. de Regt, Sabina Leonelli, and Kai Eigner, editors. *Scientific Understanding: Philosophical Perspectives*. University of Pittsburgh Press, Pittsburgh, 2009.

Sven Delarivière and Bart Van Kerkhove. The mark of understanding: In defense of an ability account. *Axiomathes*, 31(5):619–648, 2021.

Finnur Dellsén. Beyond explanation: Understanding as dependency modelling. *The British Journal for the Philosophy of Science*, 71(4):1261–1286, 2020.

Catherine Z. Elgin. *Considered Judgment*. Princeton University Press, Princeton, NJ, 1996.

Catherine Z. Elgin. Is understanding factive? In Adrian Haddock, Alan Millar, and Duncan Pritchard, editors, *Epistemic Value*, pages 322–330. Oxford University Press, New York, 2009.

Solomon Feferman. And so on...: reasoning with infinite diagrams. *Synthese*, 186(1):371–386, 2012.

Janet Folina. Towards a better understanding of mathematical understanding. In Mario Piazza and Gabriele Pulcini, editors, *Truth, Existence and Explanation*, pages 121–146. Springer, Cham, 2018.

John B. Fraleigh. *A First Course in Abstract Algebra (Seventh Edition)*. Pearson Education, London, 2003.

Joachim Frans. Unificatory understanding and explanatory proofs. *Foundations of Science*, 26(4): 1105–1127, 2021.

Maarten Franssen, Gert-Jan Lokhorst, and Ibo van de Poel. Philosophy of Technology. In Edward N. Zalta, editor, *The Stanford Encyclopedia of Philosophy*. Metaphysics Research Lab, Stanford University, Fall 2018 edition, 2018.

Mohan Ganesalingam and William Timothy Gowers. A fully automatic theorem prover with human-style output. *Journal of Automated Reasoning*, 58(2):253–291, 2017.

Emma C. Gordon. Understanding in epistemology. In Jon Matheson, editor, *The Internet Encyclopedia of Philosophy*. 2017.

Timothy W. Gowers. Mathematics, memory, and mental arithmetic. In Mary Leng, Alexander Paseau, and Michael Potter, editors, *Mathematical Knowledge*, pages 33–58. Oxford University Press, Oxford, 2007.

Stephen R. Grimm. Understanding. In Sven Bernecker and Duncan Pritchard, editors, *The Routledge Companion to Epistemology*, pages 84–94. Routledge, New York, 2011.

Stephen R. Grimm. The value of understanding. *Philosophy Compass*, 7(2):103–117, 2012.

Stephen R. Grimm. Understanding and transparency. In Stephen R. Grimm, Christoph Baumberger, and Sabine Ammon, editors, *Explaining Understanding: New Essays in Epistemology and the Philosophy of Science*, pages 212–229. Oxford University Press, New York, 2017.

Stephen R. Grimm. Understanding. In Edward N. Zalta, editor, *The Stanford Encyclopedia of Philosophy*. Metaphysics Research Lab, Stanford University, Summer 2021 edition, 2021.

Derek Haylock and Anne Cockburn. *Understanding Mathematics for Young Children: A Guide for Foundation Stage and Lower Primary Teachers*. SAGE Publications Ltd, London, 2008.

Gerhard Heinzmann. Poincaré on understanding mathematics. *Philosophia Scientiæ*, 3(2):43–60, 1999.

Matthew Inglis and Juan Pablo Mejía-Ramos. Functional explanation in mathematics. *Synthese*, 198(S26):6369–6392, 2021.

Christoph Kelp. Understanding phenomena. *Synthese*, 192(12):3799–3816, 2015.

Kareem Khalifa. *Understanding, Explanation, and Scientific Knowledge*. Cambridge University Press, Cambridge, 2017.

Hans Kurzweil and Bernd Stellmacher. *The Theory of Finite Groups: An Introduction*. Springer, New York, 2004.

Jonathan L. Kvanvig. *The value of Knowledge and the Pursuit of Understanding*. Cambridge University Press, Cambridge, 2003.

Jonathan L. Kvanvig. Understanding. In *The Oxford Handbook of the Epistemology of Theology*, pages 175–189. Oxford University Press, Oxford, 2017.

Guillaume Lample and François Charton. Deep learning for symbolic mathematics. In *Proceedings of the International Conference on Learning Representations*, 2020. URL https://openreview.net/forum?id=S1eZYeHFDS.

Leslie Lamport. How to write a proof. *The American Mathematical Monthly*, 102(7):600–608, 1995.

Leslie Lamport. How to write a 21st century proof. *Journal of Fixed Point Theory and Applications*, 11(1):43–63, 2012.

Ellen Lehet. Mathematical explanation in practice. *Axiomathes*, 31(5):553–574, 2021.

Uri Leron. Structuring mathematical proofs. *The American Mathematical Monthly*, 90(3):174–185, 1983.

Peter Lipton. Understanding without explanation. In Henk W. de Regt, Sabina Leonelli, and Kai Eigner, editors, *Scientific Understanding: Philosophical Perspectives*, pages 43–63. University of Pittsburgh Press, Pittsburgh, 2009.

Danielle Macbeth. Proof and understanding in mathematical practice. *Philosophia Scientiæ*, 16 (1):29–54, 2012.

Kenneth Manders. Expressive means and mathematical understanding. Unpublished manuscript, 2012.

Jean-Pierre Marquis. Axiomatization as conceptual design. Talk at the IHPST, 2015.

James H. McKay. Another proof of Cauchy's group theorem. *The American Mathematical Monthly*, 66(2):119, 1959.

Edwina Rissland Michener. Understanding understanding mathematics. *Cognitive Science*, 2(4): 361–383, 1978.

Rebecca Lea Morris. Motivated proofs: What they are, why they matter and how to write them. *The Review of Symbolic Logic*, 13(1):23–46, 2020.

Henri Poincaré. Intuition and logic in mathematics. In William Ewald, editor, *From Kant to Hilbert: A Source Book in the Foundations of Mathematics (Volumes I and II)*, volume 19, pages 1012–1020. Clarendon Press, Oxford, 1900/1996.

Henri Poincaré. *Science et Méthode.* Ernest Flammarion, Paris, 1908. Translated to English by Francis Maitland as *Science and Method*, Thomas Nelson & Sons, London, 1914. (Citations are to translation).

George Pólya. With, or without, motivation. *The American Mathematical Monthly*, 56(10):684–691, 1949.

Michael D. Resnik. On understanding mathematical proofs. In J.L. Greffe, G. Heinzmann, and K. Lorenz, editors, *Henri Poincaré. Wissenschaft und Philosophie*, pages 459–466. Akademie Verlag/Blanchard, Berlin/Paris, 1996.

Wayne D. Riggs. Understanding 'virtue' and the virtue of understanding. In Michael DePaul and Linda Zagzebski, editors, *Intellectual Virtue: Perspectives from Ethics and Epistemology*, pages 203–226. Oxford University Press, New York, 2003.

John Alan Robinson. Proof = guarantee + explanation. In Steffen Hölldobler, editor, *Intellectics and Computational Logic*, volume 19 of *Applied Logic Series*, pages 277–294. Kluwer Academic Publishers, Dordrecht, 2000.

Joseph J. Rotman. *An Introduction to the Theory of Groups (Fourth Edition).* Springer-Verlag, New York, 1995.

Simon Singh. *Fermat's Last Theorem: The Story of a Riddle that Confounded the World's Greatest Minds for 358 Years.* Fourth Estate, London, 1997.

Michael Strevens. No understanding without explanation. *Studies in History and Philosophy of Science Part A*, 44(3):510–515, 2013.

Fenner Tanswell. Go forth and multiply! On actions, instructions and imperatives in mathematical proofs. In Joshua Brown and Otávio Bueno, editors, *Essays on the Philosophy of Jody Azzouni.* Springer, Cham, forthcoming.

Jamie Tappenden. Proof style and understanding in mathematics I: Visualization, unification and axiom choice. In Paolo Mancosu, Klaus Frovin Jørgensen, and Stig Andur Pedersen, editors, *Visualization, Explanation and Reasoning Styles in Mathematics*, pages 147–214. Springer, Dordrecht, 2005.

William P. Thurston. On proof and progress in mathematics. *Bulletin of the American Mathematical Society*, 30:161–177, 1994.

Jason Wei, Xuezhi Wang, Dale Schuurmans, Maarten Bosma, Brian Ichter, Fei Xia, Ed Chi, Quoc V. Le, and Denny Zhou. Chain-of-thought prompting elicits reasoning in large language models. In S. Koyejo, S. Mohamed, A. Agarwal, D. Belgrave, K. Cho, and A. Oh, editors, *Advances in Neural Information Processing Systems*, volume 35, pages 24824–24837. Curran Associates, Inc., 2022.

Sean Welleck, Jiacheng Liu, Ronan Le Bras, Hanna Hajishirzi, Yejin Choi, Kyunghyun Cho, and Kyunghyun Cho. Naturalproofs: Mathematical theorem proving in natural language. In J. Vanschoren and S. Yeung, editors, *Proceedings of the Neural Information Processing Systems Track on Datasets and Benchmarks*, volume 1, pages 1–14. Curran Associates, Inc., 2021.

Sean Welleck, Jiacheng Liu, Ximing Lu, Hannaneh Hajishirzi, and Yejin Choi. Naturalprover: Grounded mathematical proof generation with language models. In S. Koyejo, S. Mohamed, A. Agarwal, D. Belgrave, K. Cho, and A. Oh, editors, *Advances in Neural Information Processing Systems*, volume 35, pages 4913–4927. Curran Associates, Inc., 2022.

Hans Wussing. *The Genesis of the Abstract Group Concept: A Contribution to the History of the Origin of Abstract Group Theory.* MIT Press, Cambridge, MA, 1984.

Linda Zagzebski. Recovering understanding. In Matthias Steup, editor, *Knowledge, Truth, and Duty: Essays on Epistemic Justification, Responsibility, and Virtue*, pages 235–252. Oxford University Press, New York, 2001.

Linda Zagzebski. Toward a theory of understanding. In Stephen R. Grimm, editor, *Varieties of Understanding: New Perspectives from Philosophy, Psychology, and Theology*, pages 123–135. Oxford University Press, New York, 2019.