Logique & Analyse **179–180** (2002), 399–419

PROOF, COGNITION, AND RATIONALITY

RON ROOD

0. *Introduction: formalized science*

The idea of a formalized science goes back to at least Frege. In fact, he considered what we have called formalization as "the ideal of a strictly scientific method". He thought to have accomplished this ideal (with respect to arithmetic) in *Grundgesetze der Arithmetik* [12]. He describes this ideal as follows:

> The ideal of a strictly scientific method in mathematics, which I have here [i.e., in the *Grundgesetze*] attempted to realize, and which might indeed be named after Euclid, I should like to describe as follows. It cannot be demanded that everything be proved, because that is impossible; but we can require that all propositions used without proof be expressly declared as such, so that we can see distinctly what the whole structure rests upon. After that we must try to diminish the number of these primitive laws as far as possible, by proving everything that can be proved. Furthermore, I demand — and in this I go beyond Euclid — that all methods of inference employed be specified in advance; otherwise we cannot be certain of satisfying the first requirement ([12], vi).

For Frege, it will turn out, the idea of formalized science is wholly decisive for his theory of proof. And Frege's theory of proof remained the standard ever since. It is no underestimation to say that Frege's ideas constitute the earliest predecessor of our modern conception of proof.

For Frege, the idea of formalized science basically rests on one specific rationality assumption. Namely, the assumption that mathematics (or, slightly better, all mathematical knowledge) forms what we shall come to call a *scientific theory*. (We explain this in § 1.)

At bottom, it is the aforementioned rationality assumption that gives shape to Frege's theory of proof. This situation is more or less analogous with, for example, applications of the theory of games in study of the social behavior of agents. For example, as in economics [32]. Here, the rationality assumption is that an economic agent maximizes expected utility. It is this assumption, then, that turns out to be decisive for the theory of economic behavior. Another parallel comes from the theory of belief revision. This theory is, roughly, about the dynamics of an agent's epistemic states (such as belief) under the influence of what Gärdenfors has called *epistemic inputs* ([18], 14). In [1], for example, a theory of belief change is presented. This theory is based on certain rationality assumptions which are nowadays known as the *AGM postulates* of belief revision.[1]

Our goal is to urge the reader to think that only Frege's aforementioned rationality assumption, if we should adopt it at all, is in fact too little for a theory of proof.

Our method is to offer a systematic presentation of Frege's theory of proof in mathematics, and to evaluate it critically.[2] We shall raise some difficulties with Frege's theory, either "directly" or in the light of current developments in logic.

If Frege's rationality assumption is not wholly decisive for the theory of proof, what points the way for another and better collection of rationality assumptions? Suggestions will be provided to the effect that we should turn ourselves towards a (or any) theorem prover, or, more generally, towards a mathematical problem solver.

Insofar as Frege did not do so (which is, as it turns out, not entirely clear), we shall still find the need for a theory of proof based on a different collection of rationality assumptions than Frege's. We shall offer some partial suggestions as to what such a collection of rationality assumptions should look like. The upshot will be that these rationality assumptions create a less austere environment for the study of the cognitive-psychological dimensions of proof than Frege's.

This paper is organized as follows. In the first section, we shall offer a systematic presentation of Frege's theory of proof. In the second section, we will briefly explore some relevant and general issues raised by current views on logic. In the third section, we shall put forward a number of difficulties we have with Frege's conception of proof. In the fourth and final section we shall suggest to think of proof, or indeed mathematical practice in general, in a different way than is suggested by Frege.

[1] An example of such a postulate is the following: expanding a belief set with any belief yields a belief set. (Within the AGM framework any belief set is represented as a consistent set of sentences, closed under logical consequence.)

[2] We shall not be concerned with any of the particularities related to Frege's logicist program.

1. *Frege's conception of proof*

In § 1.1 we shall explore some features of Frege's philosophical views on logic. We highlight that logic is fundamentally occupied with truth, or thought (though in a specific sense of *thought*). We also consider the normative status of logic. In § 1.2 we shall take a closer look at how Frege saw the relation between logic and psychology, and "naturalistic" (or empirical) psychology in particular. Finally, in § 1.3, we explore some features of Frege's philosophy of mathematics. We are particularly interested in the fact that for Frege, the truths of mathematics form what we shall come to call a scientific theory (which we shall introduce as a technical term).

1.1. *Logic*

In the *Logische Untersuchungen* Frege says:

> Just as 'beautiful' points the ways for aesthetics and 'good' for ethics, so do words like 'true' for logic. All sciences have truth as their goal; but logic is also concerned with it in a quite different way: logic has much the same relation to truth as physics has to weight or heat. To discover truths is the task of all sciences; it falls to logic to discern the laws of truth ([14], 350).

In a sense, truth is logic's object of study, and logic attempts to discern the laws of truth accordingly.

But what is a law of truth? To begin with, for Frege, a law of truth is a law prescribing how a subject *ought* to think when she wants to attain truth.[3] In this respect, laws of truth are to be taken as *normative*.

There is more to laws of truth, however. But before we proceed, let us make a terminological remark. Instead of *law of truth*, Frege alternatively uses the expressions *law of thought* and *law of logic*. He takes them all to be mutually synonymous. For convenience, we shall henceforth stick to speaking in terms of laws of truth.

As yet, we've only traced in broad outlines what laws of truth are. However, a more deeper understanding of them is needed. To this end, we first need to take a closer look at Frege's notions of thinking and, related to that, thought.

[3] "[...] the laws of logic ought to be guiding principles for thought in the attainment of truth" (Frege [12], xv).

For Frege, to think is to grasp a thought.[4] Note that, as such, to think is described as an action rather than a state.

One way of explicating Frege’s notion of a thought, in turn, is as follows: it is that what is expressed by an indicative sentence on a particular occasion of use ([13], 138). Such a sentence needs to be combined with a nonlinguistic context in case it contains indexical expressions. Alternatively, one may explicate a thought as being the object of a propositional attitude such as belief or knowledge. Such attitudes are typically described by sentences of the form

- *a believes that p*;
- *a knows that p.*

Here, a is an expression denoting a subject; p (or rather, *that p*) denotes a Fregean thought — the object of a propositional attitude. Accordingly, sentences of the above form are interpreted in terms of a relation between a subject and a thought.

Accordingly, a thought is what some, but not all philosophers, tend to call a *proposition*. As such, a thought is a nonlinguistic, non-mental, eternal and stable entity. Moreover, thoughts are the bearers of truth values. Specifically, thoughts are either of two different types: the ones that are true and the ones that are false ([13], 149).

Henceforth, we shall continue to speak of propositions rather than of thoughts.

The laws of truth have what we may call epistemic import, for they bear on matters of justification. As follows: when a subject thinks in compliance with the laws of truth, then a subject comes to believe a *true* proposition; moreover, the subject comes to believe that proposition in such a way that the subject’s belief becomes justified ([12], xvii; [14], 352). In short, thinking in compliance with the laws of truth forms a sufficient condition for truth and justification. Hence, when a subject thinks in accordance with the laws of truth, he comes to *know* a proposition.

A belief can be either justified immediately, or otherwise mediately, on the basis of truths other than the one that is the object of belief. In the latter case, Frege speaks of *inference*. Somewhat more precisely, an inference is a belief, such that one is at the same time cognizant of other truths providing justification for that belief.

Any belief that is not an inference is justified immediately and is accordingly not grounded on any other proposition except for the proposition that is the object of that belief. Such propositions are called *axioms*. For Frege, axioms are true propositions that are self-evident and known ([16]).

[4] “Thoughts are not psychic structures, and thinking is not an inner producing and forming, but an apprehension of thoughts which are already objectively given” [15], 113).

In general, a proposition p (say) can be traced backwards to other truths, namely, to those which p is inferentially based on. The inferred truths among the latter can likewise be traced back to other truths, and so on. At some point, one will always arrive at truths that are not inferred, i.e., at axioms — or so it is assumed. The series of truths found in the way indicated in the previous paragraph is called a *proof* of p.

> If we start from a theorem and trace the chains of inference backwards until we arrive at other theorems or at axioms, postulates or definitions, we discover chains of inference starting from known theorems, axioms, postulates or definitions and terminating with the theorem in question. the totality of these inference-chains constitutes the *proof* of the theorem ([13], 204).[5]

Reversing the order in which the proof is obtained, we may say that a proof of p starts from beliefs that are justified immediately, and proceeds, via chains of inferences to p.

> We may say that a proof starts from propositions that are accepted as true and leads *via* chains of inferences to the theorem ([13], 204).

Let us make the following two important observations, both of which are suggested by the above. First, proofs are discovered in a way which is essentially *regressive*: given a proposition, in order to find its proof, "backtrack" that proposition to other propositions, in a way as indicated above, until no further such backtracking is possible.[6] Second, the justification of beliefs proceeds in a way which is essentially *progressive*. More specifically, a proof starts from (known) axioms, and proceeds via intermediate steps to the theorem. The way of justification of one's beliefs that a proposition p holds true is basically the reverse of the process by means of which a proof of p is discovered.[7]

The regressive way of discovering proofs on the one hand, and the progressive way of justification of beliefs on the other, intimately relate to the fact that for Frege all mathematical truths stand in specific type of logical relationships to one another. We shall come back to this point in § 1.3.

[5] "A postulate is a truth as is an axiom, its only peculiarity being that it asserts the existence of something with certain properties" ([13], 207); "Once this [i.e., setting up a new definition] has happened, one can make out of this definition a self-evident proposition which is then to be used like an axiom" ([16], 7).

[6] Cf. also [11], 4.

[7] See Kowalski [27], preface, for a strikingly parallel view.

1.2. *Logic and psychology*

Frege’s philosophy of logic (and mathematics) is deeply influenced by his so-called *anti-psychologism*.

Roughly, psychologism is an umbrella term for theories holding that logic is based on psychological descriptions or explanations as to how subjects actually reason. Note that psychologism, thus conceived, concerns the foundations of logic in the first place, rather than logical theory itself.[8]

Considering Frege, psychology should be understood as a form of natural science. Comparing psychology to what we are nowadays inclined to call either experimental psychology, neuroscience, or some possible combination of these, will presumably not be entirely adequate if not anachronistic. At any rate, however, we think that such comparisons, whatever their adequacy, put Frege’s conception of psychology into some perspective. Deepening this issue any further would certainly lead us too far afield. Let us conclude, then, that Frege leveled his criticisms against the *naturalistic* variant of psychologism.

For Frege, to think — i.e., to grasp a proposition — is not a purely psychological type of state or event. Thinking always involves a non-mental element, namely a proposition. As Frege sees it, a psychological state or event is always bound to an individual psychological subject; as such, no two psychological states or events belonging to two respective such subjects are ever the same. If thinking were a psychological state or event, then this would invariably lead to either subjectivism (or idealism) or to relativism. Both Frege found unacceptable ([13], 143–4; [12], xvi).

As will be expected, the actual psychological processing that leads a person to believe something forms the subject matter of psychology. Psychologists, at least those of the naturalistic variety, are, among other things, interested in describing or explaining general features of such processes. However, such descriptions or explanations should never be taken as justification for a belief that a psychological process ends up in. As Frege puts it: “[w]ith the psychological conception of logic we lose the distinction between the grounds that justify a conviction and the causes that actually produce it” ([13], 147).

1.3. *Mathematics*

The set consisting of all truths of mathematics (that is, all the truths of arithmetic, or geometry) has a quite specific property. Before introducing it, let us first introduce the following terminology. A *logical system*, or simply

[8] This is also confirmed by what Frege says in the introduction of *Grundgesetze*.

a *system*, is given by specifying (1) a language, (2) a set of sentences, any member of which is regarded as an "axiom", and (3) a set of sound rules of inference. For convenience, we think of a language as a set of sentences. Thus, axioms are simply members of the language.

Note that this matches the so-called Hilbert-style presentation of a logical system. This observation will be of relevance for things we are going to say in § 3.

A *scientific theory* or simply a *theory* is a pair (S, T) consisting of a system S and a set of sentences T satisfying the following two conditions: (i) any member of T is in the language of S, and (ii) any sentence in T is proved from the axioms of S by means of successive applications of the rules of inference of S.[9]

We say that a theory (S, T) satisfies the *truth property* if all the sentences in T are true; we say that (S, T) satisfies the *evidence property* if all the axioms of S are self-evident.[10]

Earlier we said that axioms are always known propositions. Given that any sentence in a scientific theory is proved and that the rules of inference are sound, it follows that any truth in a scientific theory is known. A scientific theory, then, represents a completed body of scientific knowledge in a systematically ordered manner.

In *Grundgesetze* [12], Frege presented a system — call it G — such that G, together with the set consisting of all the truths of elementary arithmetic form a theory. (At least, this is a reasonable way of reading *Grundgesetze*.) Arithmetic is a scientific theory in exactly this sense.[11] Moreover, this theory satisfies both the truth and the evidence property. In fact, we should say that this theory satisfies something stronger than the truth property: all sentences in this theory are not merely true but *logically true*. However, it seems that Frege never made a clear distinction between mere truth and logical truth.

We mention that Frege did not present an exact specification of the language of his system in *Grundgesetze*, nor did he anywhere else. Neither did he give a precise demarcation of the sentences in this system, except,

[9] Our notion of a scientific theory comes close to what Beth has called an *Aristotelian science* ([7], 31).

[10] Beth calls this respectively the *truth postulate* and the *evidence postulate* of a scientific theory (i.e., what Beth calls an Aristotelian science). To be more precise, for Beth, a scientific theory satisfies the evidence postulate if (1) all axioms of that science are self-evident, and (2) if all the primitive terms from which the sentences in the language are composed are self-evident. (See [7], 32.) Since it is not relevant for our purposes we have ignored (2). Beth also introduces a deduction postulate, which relates to (ii) in our definition of a scientific theory.

[11] To be more precise: G, together with the truths of arithmetic as represented in the language of G, forms a scientific theory.

perhaps, that the laws of truth are "the most general laws, which prescribe universally the way in which one ought to think if one is to think at all" ([12], xv). Nonetheless, we do think that what we've said in the previous paragraph forms a reasonable and useful reconstruction of Frege's views on the matter.

Besides the truths of arithmetic, it is likely that Frege also thought that the truths of geometry form a scientific theory. And presumably, he would have said the same about the truths of (theoretical) physics. Although in the case of both physics and geometry, there would be non-logical axioms among the axioms of the underlying system. There are also obvious links with the various "sources of knowledge" mentioned earlier. Due to limitations of time and space we cannot pursue this point any further.

Earlier we noticed that there are immediately and mediately justified beliefs. Given a theory, the axioms correspond to the immediately justified beliefs; all the other truths correspond to mediately justified beliefs. This suggests that justification essentially bears on the logical interrelationships between truths as they are laid down in terms of a scientific theory.

Moreover, the above nicely reveals the double nature of proofs. On the one hand, the function of proof is to *prove*, i.e., to justify a belief; on the other hand, the function of proof is to reveal the interrelationships between truths (as they appear within a scientific theory):

> A proof does not only serve to convince us of the truth of what is proved: it also serves to reveal logical relations between truths.[12]

Again, this suggests that Frege's notion of proof is intimately connected to the supposed fact the truths of mathematics form a scientific theory.

Let us make a brief summary. In § 1.1, we found that for Frege, logic is oriented towards truth and it aims to discern the laws of truth. Moreover, logic is a normative discipline. When a subject wants to attain knowledge, he ought to think in accordance with the laws of truth.

From § 1.2 it became clear that, for Frege, logic and empirical psychology are strictly separated in the sense that logic does not rest on psychology. The actual psychological processing that lead a subject to believe something belongs to the subject matter of psychology while proof is strictly a matter of logic.

In § 1.3, we observed that Frege endorsed a quite specific conception of scientific rationality: all the truths of mathematics form what we have called a scientific theory. In the light of this, all (scientific) knowledge is knowledge "from axioms". The axioms which form the ground of our knowledge are like wise known, and hence true propositions.

[12] See also [11].

Against this background, the way of discovering proofs is essentially regressive: discovering proofs of theorems takes place by tracing those theorems back to other truths and ultimately back to axioms. We also found that the way of justifying beliefs is essentially progressive; the way of justifying the belief that a proposition p holds true, is essentially the reverse of the way in which a proof of p is discovered. Both the way of discovering proofs as well as way of justification of beliefs belong to the province of logic.

Let us state our conclusion. It seems clear that the concept of proof Frege endorses is fundamentally shaped by a certain *rationality assumption*, namely that the truths of mathematics form a scientific theory as explained above. However, suppose, if only for the sake of argument, that Frege is right in holding that the truths of mathematics form a scientific theory in exactly this sense. Then still we cannot resist the impression that the tight connection with Frege's specific ideas on proof and justification goes by and large unmotivated.

2. *The shape of logic today: some contours*

Despite Frege's widely acknowledged and indeed fundamental contributions to logic, the subject has evolved considerably ever since. It is beyond doubt that on a more technical side logic has progressed enormously. And it is claimed that these developments have had their repercussions on a more conceptual level too. Even so much so that we are inclined to say that because of these developments, many features of Frege's philosophy of logic and mathematics cannot be reasonably maintained (see § 3).

It is the purpose of this section to highlight some of these developments. We shall first and foremost focus on the notion of a logical system and how it is broadly conceived by logicians.

For Frege, there is at bottom only one logical system; logic is "universal" (Van Heijenoort [23]). Nowadays, however, logicians do not tend to identify "logic" with a single logical system. There currently exist a great many types of logical systems. Examples are those known under the names: classical logic, intuitionistic logic, deontic logic, temporal logic, epistemic logic, relevance logic, linear logic, many-valued logic, non-monotonic logic, paraconsistent logic, and dynamic logic. Many more could be added to this list.

Broadly conceived, logical systems are nowadays basically defined along either of the following tow lines

- as consequence relations;
- as relations between models and sentences.

The first is characteristic for more proof-theoretic lines of research, the second for more model-theoretic lines of research. Let is briefly discuss these two ways of defining logical systems. Our prime interest is in logical systems defined as consequence relations.

From a proof-theoretic perspective, one can define a logical system in general as, for example, a set-theoretical consequence relation $\Gamma \vdash \Delta$ between sets of sentences Γ and Δ satisfying certain structural rules (for instance, Reflexivity, Monotonicity and Cut). This approach is epitomized in a paper by Gabbay [17]. (We can intuitively interpret $\Gamma \vdash \Delta$ ("Δ is a consequence of Γ") as follows: any $\varphi \in \Delta$ is a consequence of Γ.)

Note that an approach along these lines always presupposes a language from which the sentences in both Γ and Δ are taken. Accordingly, it would be more exact to say that a logical system is given by specifying a language and a relation between sets of sentences from that language. That is, a logical system can be represented in terms of a pair $(L, \vdash)$. Here, L is a language (for convenience defined as a set of sentences) and $\vdash$ is a relation on $\wp(L)$.

Observe also that a definition along the lines indicated above is purely extensional. In the following sense: given logical system, then this system basically demarcates which pairs (Γ, Δ) are such that Δ is a consequence of Γ from all other such pairs. In particular, the system under consideration does not tell us what a proof is; it doesn't even tell us what a proof ought to be (see § 3.2. below). Rather, the definition of $\vdash$ *presupposes* a certain "analysis" of the concept of proof.

Accordingly, a logical system does not really define what a proof is, it merely tells us that certain proofs exist. Some systems are presented in such a way that they lead to a notion of formal derivation. It is not clear, however, whether derivation is the same as proof.

Of course, there is more to a logical system than being merely a consequence relation on a given language. For example, logical systems often lead to methods to *prove* whether or not a given set of sentences is a consequence of another set of sentences. These methods may be either heuristic (e.g., constructing derivations) or effective (e.g., constructing a semantic tableau). Sometimes there even is a decision procedure available (e.g. as in the case of constructing semantic tableaux or truth tables for classical propositional logic).

The second way of defining a logical system plays its role in more model-theoretic lines of research. Here, one can in general define a logical system as a relation $M \vDash \varphi$ between models M and sentences φ. Sometimes, logical systems defined accordingly are called *model-theoretic logics*. See Barwise and Feferman [1].

By defining a logical system accordingly, both a language and a class of models are always presupposed. Thus, to be more precise, we represent a model-theoretic logic as a triple $(L, \mathsf{M}, \vDash)$. Here, L is a language, M a class of models, and $\vDash$ is a subset of $\mathsf{M} \times L$.

The model-theoretic perspective places (model-theoretic) truth at the center of inquiry. Models are mathematical tools for describing the possible semantic interpretations of the sentences of a language. For Frege, however, it is out of the question to think of a single sentence as possibly having different semantic interpretations.

Model-theoretic logics will not play any role in what follows. We mentioned them merely for the sake of completeness.

Given a logical system, defined in whatever way, an important part of a logician's business consists of establishing various of its properties. For example, logicians are interested in answering such questions as: is the system consistent? Is the system decidable? In short, an important part of a logician's task is metasystematic: he tries to establish results which are *about* logical systems.

Logical systems defined in terms of a consequence relation can be presented in various different ways. To illustrate this, let us briefly mention several ways of presenting systems for first-order logic. Let us note beforehand that the various different ways of presenting logical systems determine exactly the same set-theoretical consequence relation.

Hilbert-style systems, for example, are a fairly simple and elegant way of presenting logical systems in an axiomatic manner. As suggested earlier, with hindsight we can reasonably say that Frege's notion of a logical system is that of an Hilbert-style system.

However, as is commonly acknowledged by logicians, Hilbert-style systems fail to reflect accurately the steps taken in actual proofs. To some extent, this lack is remedied by natural deduction-style systems ([37]). Systems presented accordingly are intended to offer a more accurate reflection of the steps that have led to an actual proof.

However, we think that, in order to forestall overwrought expectations, the latter point should be put somewhat more carefully. Along the following lines: natural deduction systems accurately reflect the steps that have been taken in the course of a proof, but only insofar as this proof is represented in a specific language (such as, e.g., a first-order language). This point is stressed in Ebbinghaus, Flum, Thomas [10].

There are two other ways of presenting logical systems which are not of real import for our purposes, but which seem nonetheless worth mentioning.

First, sequent-style systems (or Gentzen-style systems, as they are sometimes called) are most suited for the meta-systematic study of logical systems themselves. Second, resolution-style systems are of interest primarily because they are relatively easy to implement on a computer. This makes them fairly suitable for various applications in computer science such as e.g. automated reasoning.

3. *Difficulties with Frege's theory of proof*

In this section, we shall point out a number of difficulties arising from Frege's views on logic and mathematics. It is not our purpose to treat any of these difficulties decisively. Rather, we want to urge the reader to think that Frege's very conception of logic and mathematics is problematic in various different ways.

Some of the difficulties mentioned arise more or less directly from the things said in § 1, others arise more or less in view of current developments as set out in § 2.

1. The first difficulty we find concerns the supposed normativity of logic. It is a fundamental one, in the sense that it gives force to many others that follow.

As we noted earlier, for Frege logic is universal. However we observed that logicians nowadays study a multiplicity of logical systems. This multiplicity of available logical systems, together with the metasystematic perspective from which any of these systems is studied, are likely to have radically altered our view on the normativity of logic. In fact, it seems not unreasonable to claim that many logicians hold that logic is much less normative in nature than Frege thought, if at all.

Indeed, quite a few logicians nowadays think that one of the central problems of logic consists of the systematic classification of valid and invalid arguments and the study of the methods used in classifying them accordingly. Copi, for example, says that

> [t]he study of logic, then, is the study of the methods and principles used in distinguishing correct (good) from incorrect (bad) arguments ([9], 1).[13]

Logic, roughly, consists of certain methods and principles which are used to distinguish arguments in terms of "good" and "bad". The methods and principles are delivered by a given presentation of a logical system. And it are these principles that are studied by logicians. Thus, the task of logic is much more descriptive in nature rather than normative.

Accordingly, logic itself presupposes certain norms that apply to arguments. This contrasts with Frege, who held that logic itself sets the norms as to what "good" and "bad" arguments are.

[13] See also Lemmon [9]; Prior [38]; Mates [31]. Remarkably, in 1925, C.S. Peirce already said the following:

> It will, however, be generally conceded that its [i.e., logic's] central problem is the classification of arguments, so that all bad are thrown into one division, and those which are good are thrown into another ([34]).

2. The second difficulty we find is Frege's suggestion that proofs are discovered in a way which we broadly characterized as "regressive". Frege may be right when using a Hilbert-style presentation, but the situation changes when we turn to a natural deduction-style presentation. For in the latter case, proofs are often found in exactly the reverse way.

We add that Polya mentions several other ways of finding proofs. For example, by making analogies with other but similar problems, or by breaking up the result in a couple of more specific cases and proving each of them separately [35]. It is far from clear why these should be ruled out as legitimate ways of finding proofs.

3. Related to this, we observed earlier that for Frege the way of justifying beliefs is essentially progressive. Justification basically proceeds "from axioms to theorem". Moreover, the axioms that form the starting point are always known, and hence true.

Let us make three observations on behalf of this. First, the fact that proofs are always to start from known axioms seems to rule out proofs by *reductio*, or indirect proofs as they are sometimes called. In fact, reductio proofs start from assumptions that are (or turn out to be) false.

Now whatever opinion one might have about reductio proofs, it is far from clear why they *should* be ruled out, if at all. Even in case when they can be reconstructed in terms of direct proofs. In our view, such a claim cannot be made on the basis of logic alone. At any rate, we think that some substantial philosophical argument is required on this point (as, for example, attempted by some of the intuitionist philosophers of mathematics). As far as we can see, however, Frege offers no such argument.

The second remark more or less elaborates on the first. It concerns Frege's insistence that axioms are always self-evident. In her two interesting studies "Believing [*sic*] the axioms" [29], [30], Maddy argued, among other things, that the axioms of what is nowadays called Zermelo-Fraenkel set theory (including the axiom of choice), including some of the potential axioms of higher set theory, where not generally taken as self-evident. It is Maddy's claim that some of the axioms of set theory where initially accepted solely for extrinsic reasons.

In order to understand what this means, let us note the following. Maddy distinguishes between "intrinsic reasons" and "extrinsic reasons" for believing an axiom. The former relate to matters of self-evidence or obviousness. The latter relate to various pragmatic considerations. For example, axioms may be accepted just because of the consequences they have, or because of the inter-theoretic connections they establish, their explanatory power, or avoidance of paradoxes ([29], 482–3).

Third, it seems questionable whether genuine scientific knowledge in mathematics is only knowledge "from axioms", as Frege seems to suggest. In this respect Brown has observed that quite a few branches of mathematics are not axiomatized at all. He mentions the theory of matrices as an illustrative example ([8]). And he adds that it would be somewhat odd to deny that we know a lot about matrices, indeed, that we have genuine proofs showing interesting properties that matrices have.

4. Second, we have observed that there are various ways of presenting logical systems. With hindsight, however, we may say that Frege has what seems to be an unmotivated predilection with Hilbert-style systems We suggested this earlier.

Taken on itself, this need not be a real problem. However, it does become problematic in the light of our conclusion that Frege's predilection for Hilbert-style systems is decisive for his theory of proof, and in fact for his philosophy of mathematics in general. With hindsight, however, we can say that there are other ways of presenting logical systems. Any of these leads to different notions of proof. For example, in case of natural deduction systems, axioms hardly play any role, if at all. In fact, natural deduction proofs are proofs from mere assumptions which need not be true.

5. Frege himself acknowledged that the proofs in his system do not accurately reflect the steps that have been taken in actual proofs ([12], vii–viii). These proofs are, among other things, considerably greater in length than those normally produced by mathematicians. Given that Frege in *Begriffsschrift* defined only one rule of deduction (i.e., *modus ponens*), he made a modest attempt to remedy this point in *Grundgesetze* by increasing the number of such rules. However, adding more rules of inference only decreased the length of proofs to a certain extent; it did not make the steps taken substantially more natural.

As will be expected, Frege would have explained (away) the divergence of the proofs in his system from actual mathematical proofs by pointing out that the former set a norm. It will come as no surprise that in our view such a way out becomes highly problematic.

Moreover, there is another reason why one should not pass over this point too lightly. For there are purely practical matters of computational complexity lurking at the background here. Those who carry out proofs are without exception in the possession of limited computational capacities such as time and memory space and attention span. In short, to prove bound to limited resources (cf. Harman [22], 12). And this may well turn out to be problematic. For example, the ideal proofs that Frege sets as the norm may be too complex to carry out in practice.

6. The sixth difficulty is related to the previous one and concerns the relationship between logic and psychology.

Let us begin by noting that it is very likely that Frege's strict separation of the domain of logic from the domain of psychology is nowadays known under a specific guise. Namely, as the distinction between "context of justification" and "context of discovery".

This distinction is known from the philosophy of science.[14] Usually, it is roughly explained as one between psychological belief forming processes on the one hand and logical considerations bearing on justification of beliefs on the other. In this respect, from the logical literature we mention Kahane [26], 2–3; Salmon [41], 10–11. For relevant pointers to the literature from the philosophy of science we advise the reader to consult the paper by Hoyningen-Heune [24].[15]

More specifically, however, one way of understanding the distinction between context of justification and context of discovery is as a distinction between two kinds of processes.[16] That is, as a distinction between processes that lead to states of belief on the one hand and processes of justification on the other. Thus understood, the distinction matches Frege.

The distinction between discovery and justification may be illustrated by pointing out a few examples from the folklore of the history of science. Hardy, Seshu Aiyar and Wilson report Ramanujan saying of himself that he was often prompted mathematical theorems by the Goddess of Namakkal during his sleep ([21], xii). Apparently, proving those theorems was something Ramanujan did afterwards, while he was awake.

Poincaré reports another well-known case about himself. While stepping on the footboard of a bus in order to leave for a geological excursion, he suddenly saw the light on a problem he was pondering over for days. He wrote down the proofs on the matter a few days after this remarkable event ([36]).

These examples are dramatic, however. They raise the impression that the process of belief formation and the process of justifying one's beliefs are temporally disjoint processes, which is what Frege seems to have thought. However, one should be very reserved with respect to drawing one's general conclusions from such examples. In fact, it seems not unreasonable to hold that quite often these two processes go virtually hand in hand. Salmon, for example, mentions the type of cases in which a subject forms a belief by going through a sound algorithmic procedure [40]. To this, we may add that a similar conclusion may hold when one forms a belief by going through an heuristic procedure.

[14] The distinction, at least in these terms, is commonly said to originate with Reichenbach [39].

[15] Thanks to Henk de Regt for making me aware of the existence of this illustrative paper, and for borrowing me his copy.

[16] There are other ways of understanding the distinction which need not bother us here. Again, see [24].

7. The next difficulty that arises more or less directly from Frege's theory of proof is his seemingly one-sided view that in mathematics one encounters only what Polya has called *problems to prove* ([35]). That is, cases were one is given the task of finding a proof of a given theorem.

However, Polya has distinguished a second type of problems a mathematician may be confronted with. He called them *problems to find*. These are problems such that one is not so much given the task of finding a proof of a theorem but of finding a mathematical object of some sort.[17] Examples of such problems are: solving equations of various sorts (e.g. polynomial equations, differential equations), finding a greatest common divisor of several numbers, construction problems in geometry, and finding algorithms. To this, we may add that finding a definition is often also a (substantial) "problem to find".

4. *From scientific theory to theorem prover*

The difficulties that we found in the previous section often suggest that certain matters coming from mathematical practice bear on our concept of proof. They are matters of the following type: those proving theorems often appear to do such and such, or cannot but do within such and such limits. And Frege's notion of proof is such that it cannot account for these issues, or it simply ignores these them.

A terminological issue is in order here. The conception of mathematical practice presupposed here is, in a way, fairly limited. Somewhat more specifically, we understand mathematical practice in a broadly cognitive sense. Even more specifically, we understand mathematical practice in terms of relatively short time tasks such as proving a theorem.[18] Mathematical practice, in this sense, basically comes down to proving theorems (or, more generally, solving mathematical problems) by individuals. It are considerations arising from mathematical practice thus understood that will be relevant for what we have to say.

Accordingly, then, the issues that we have brought up against Frege were often, but not always, of a cognitive nature. This may seem odd, or even question-begging. For Frege held that what a proof is merely bears on matters of logic, and that it must never be based on any consideration coming from psychology. However, if one is prepared to take at least some of these

[17] This suggest that proofs are not mathematical objects. The point is harmless however.

[18] Other ways of understanding mathematical practice are, for example, in terms of fairly large scale historical developments (cf. the work of Lakatos), the sociology of mathematics, and issues bearing on education.

difficulties seriously — and we think one should —, the gap between logic and psychology turns out less wider than Frege claimed.

Upon closer inspection, however, we can interpret the difficulties we brought up in a different manner. Namely, as pointing the way for a different set of rationality assumptions than Frege's. Our proposal is not to have a theory of proof based on the rationality assumption Frege assumed, namely that the truths of mathematics form a scientific theory. Such an assumption creates only a very austere environment for the study of the cognitive dimensions of proof. Rather, we propose to find another collection of rationality assumptions. These rationality assumptions, then, should offer a more realistic account of proof.

In broad terms, our proposal is not to focus wholly on the idea of formalized science and to orient ourselves towards on a (or any) theorem prover. Or, more generally, towards a mathematical problem solver.

Newell and Simon [33] and their followers in cognitive science and AI, explicitly state that their theory of problem solving is a theory of the individual. They assume that someone has a problem when he wants something and does not know immediately what actions to perform to get it (Newell and Simon [33], 72). To have a problem implies that one be given certain initial information concerning what is given, the desired goal, under what conditions this goal is to be reached, the admissible actions, and so on. It seems likely that one also be given some information about the accessible resources during the problem solving process. For example, information concerning the available amount of time, the available amount of memory space, whether or not one is allowed to use external resources, and so on.

Solving a problem comes down to finding a solution path from an initial state to a goal state in a problem space. A problem space basically consists of a set of states together with a collection of actions that can be executed on these states in order to obtain new states. Thus, problem solving in a problem space is always a problem of search.

What rationality assumptions to adopt is far from trivial. And we do not claim to offer a satisfactory treatment of the matter here. We suffice to offer a couple of suggestions which pave the way for future work.

In the light of a problem solving framework, rationality assumptions will fundamentally bear on the possible solutions paths in a given problem space. More specifically, matters related to relevance of goals become prominent. Indeed, solving a problem is generally a goal directed activity. Further, questions as to what actions are admissible in order to reach a goal become urgent. For within a given domain of science (e.g. mathematics), not every action is admissible. This leads us to speculate that there are possibly domain-specific principles related to a given field.

Related to this, attention should be given to matters of heuristics and plans. And as we said earlier, the fact that problem solvers only have limited cognitive resources at their disposal needs to be accounted for.

Of course, assumptions along these lines can be hardly be counted as serious and realistic alternatives for Frege's until they are worked out in considerable detail. We do think, however, that this can be done and, indeed, that serious attempt have been undertaken in the theory of reasoning and general epistemology.

Harman suggests that, for example, conservatism, simplicity, and coherence are important aspects of what he calls theoretical rationality ([22], chapter 1). Other examples he mentions bear on such matters the relevance of goals and interests, and available cognitive resources. (See also Goldman [19]).

Moreover, our proposals do not serve to deliver an alternative for logic, but rather to supplement it. In this respect, we should note that matters of complexity of tasks have been put on the logical agenda for quite some time. And similarly for issues related to the cognitive structure of proofs (see Van Benthem [6]). To this, we can add that the growing interest in nonlinguistic forms of reasoning among logicians can also be interpreted as an attempt to obtain a more realistic theory of proof (see [4], [5], [25]).

However, still considerable work needs to be done. Interesting and important questions we have in mind are: what are plausible rationality assumptions? Should we pose rationality assumptions which apply exclusively to mathematics and not to other fields? And related to this: is mathematical reasoning a specific kind of reasoning, and if so, what does make it a distinguished type of reasoning? Are there specific mathematical cognitive actions?

Department of Philosophy
Vrije Universiteit, Amsterdam
E-mail: `p.t.m.rood@ph.vu.nl`

REFERENCES

[1] C.E. Alchourrón, P. Gärdenfors, D. Mackinson. On the logic of theory change: partial meet contraction and revision functions. *Journal of symbolic logic* 50, 510–30 (1985).
[2] Aristotle. *Analytica posteriora.* In: *The works of Aristotle*, vol. 1, W.D. Ross (ed.), Oxford: Oxford University Press (1963).
[3] J. Barwise, S. Feferman (eds.). *Model-theoretic logics.* New York: Springer-Verlag (1985).

[4] J. Barwise, J. Etchemendy. Visual information and valid reasoning. In: *Visualization in mathematics*, Mathematical Association of America, notes no. 19, W. Zimmermann and S. Cunningham, (eds.), Washington: Mathematical Association of America, 9–24 (1991). Reprinted in *Philosophy and the computer*, L. Burkholder (ed.), Boulder: Westview Press, 160–182 (1992).
[5] J. Barwise, J. Etchemendy. Heterogeneous logic. In: *Diagrammatic reasoning: cognitive and computational perspectives*, J. Glasgow, N. Hari Narayanan, B. Chandrasekaran (eds.), Cambridge: MIT Press, and Menlo Park: AAAI Press, American Association for Artificial Intelligence, 211–234 (1995).
[6] J.F.A.K. van Benthem. *Exploring logical dynamics.* Stanford: CSLI Publications, Center for the Study of Language and Information, Stanford University (1996).
[7] E.W. Beth. *The Foundations of Mathematics, A Study in the Philosophy of Science.* Studies in Logic and the Foundations of Mathematics. Amsterdam: North-Holland Publishing Company (1959).
[8] J.R. Brown. *Philosophy of mathematics: an introduction to the world of proofs and pictures.* London: Routledge (1999).
[9] I.M. Copi. *Symbolic logic.* New York: MacMillan (1973).
[10] H.-D. Ebbinghaus, J. Flum, W. Thomas. *Mathematical logic.* New York: Springer-Verlag (1989).
[11] G. Frege. *The foundations of arithmetic: a logico-mathematical enquiry into the concept of number, 2nd ed.* Transl. by J. L. Austin, Oxford: Blackwell (1978).
[12] G. Frege. *The basic laws of arithmetic: exposition of the system.* Transl. and ed. by M. Furth, Berkeley: University of California Press (1982).
[13] G. Frege. *Posthumous writings.* H. Hermes, F. Kambartel, F. Kaulbach (eds.), Oxford: Basil Blackwell (1979).
[14] G. Frege. *Collected papers on mathematics, logic, and philosophy*, B. McGuinness (ed.), Oxford: Blackwell (1984).
[15] G. Frege. Letter to Husserl. In: E. Husserl, *Briefwechsel.* Vol. VI, Dordrecht: Kluwer, 1994.
[16] G. Frege. *Gottlob Frege on the foundations of geometry and formal theories of arithmetic.* Transl. and intr. by E. Henner W. Kluge, New Haven: Yale University Press (1971).
[17] D.M. Gabbay. A general theory of structured consequence relations. In: *Substructural logics*, K. Došen, P. Schroeder-Heister (eds.), Oxford: Clarendon Press (1993).
[18] P. Gärdenfors. *Knowledge in flux. Modeling the dynamics of epistemic states.* Cambridge: The MIT Press (1988).
[19] A.I Goldman. *Epistemology and cognition.* Cambridge: Harvard University Press (1986).

[20] Hacking, I. What is logic? *The journal of philosophy* 76(6), 285–319 (1979).
[21] G.H. Hardy, P.V. Seshu Aiyar, B.M. Wilson (eds.). *Collected papers of Srinivasa Ramanujan.* Cambridge: Cambridge University Press (1927).
[22] Harman, G. *Reasoning, meaning and mind.* Oxford: Oxford University Press (1999).
[23] J. van Heijenoort. Logic as calculus and logic as language. *Synthese* 17, 324–30 (1967).
[24] P. Hoyningen-Heune. Context of discovery and context of justification. *Studies in the history and philosophy of science* 18(4), 501–15 (1987).
[25] M. Jamnik. *Mathematical reasoning with diagrams. From intuition to automation.* Stanford: CSLI Publications, Center for the Study of Language and Information, Stanford University (2001).
[26] H. Kahane. *Logic and Philosophy.* Belmont: Wadsworth (1973).
[27] R. Kowalski. *Logic for problem solving.* New York: North-Holland (1983).
[28] E.J. Lemmon. *Beginning logic.* London: Nelson (1965).
[29] P. Maddy. Believing the axioms I. *The journal of symbolic logic* 53(2), 481–511 (1988).
[30] P. Maddy. Believing the axioms II. *The journal of symbolic logic* 53(3), 736–64.
[31] B. Mates. *Elementary logic*, 2nd ed. New York: Oxford University Press (1972).
[32] J. von Neumann, O. Morgenstern. *Theory of games and economic behavior.* Princeton: Princeton University Press (1953).
[33] Newell, A., H.A. Simon. *Human Problem Solving.* Englewood Cliffs, NJ: Prentice Hall (1972).
[34] C.S. Peirce. Logic. In: *Dictionary of philosophy and psychology*, vol. 1, London: MacMillan (1925).
[35] G. Polya. *Mathematical discovery: on understanding, learning and teaching problem solving.* 2 vol.'s, New York: Wiley (1962–1965).
[36] H. Poincaré. *Science and method.* In: *The foundations of science* transl. and ed. by G. B. Halsted (1982).
[37] D. Prawitz. *Natural deduction. A proof-theoretical study.* Stockholm: Almqvist & Wiksell (1965).
[38] A.N. Prior. *Formal logic*, 2nd ed. London: Oxford University Press (1963).
[39] H. Reichenbach. *Experience and prediction.* Chicago: Chicago University Press (1961).
[40] W.C. Salmon. Bayes's theorem and the history of science. *Minnesota studies in the philosophy of science 5: historical and philosophical*

perspectives of science, R. H. Stuewer (ed.) Minneapolis: University of Minnesota Press (1970).

[41] W. C. Salmon. *Logic*, 3rd edition. Englewood Cliffs: Prentice Hall (1984).