Journal of Automated Reasoning (2021) 65:1153–1190
https://doi.org/10.1007/s10817-021-09594-z

Human-Centered Automated Proof Search

Wilfried Sieg[1] · **Farzaneh Derakhshan**[1]

Received: 15 June 2020 / Accepted: 28 April 2021 / Published online: 30 July 2021

Abstract
Human-centered automated proof search aims to capture structures of ordinary mathematical proofs and discover human strategies that are used (implicitly) in their construction. We analyze the ways of two theorem provers for approaching that goal. One, the G&G-prover, is presented in Ganesalingam and Gowers (J Autom Reason 58(2):253–291, 2017); the other, Sieg's AProS system, is described in Sieg and Walsh (Rev Symb Logic 1-35, 2019). Both systems make explicit, via their underlying logical calculi, the goal-directedness and bi-directionality of proof construction. However, the calculus for the G&G-prover is a weak fragment of minimal first-order logic, whereas AProS uses complete calculi for intuitionist and classical first-order logic. The strategies for the construction of proofs are dramatically different as well. The G&G-prover uses a waterfall strategy and is thus restricted to problems that can be solved without backtracking. The AProS strategies, by contrast, support a complete search procedure with backtracking. These divergences are rooted in the fact that the concrete goals of the systems are different: The G&G-prover is to yield write-ups indistinguishable from good mathematical writing; AProS is to yield humanly intelligible formal proofs by logically and mathematically motivated strategies. In our final *Programmatic remarks*, we sketch a plausible, but difficult project for achieving more fully G&G's broad goals by radically separating proof search from proof translation: one could use AProS for the proof search and then exploit the strategic structure of the completed proof as the deterministic underpinning for its translation into a natural language.

Keywords Proof search · Automated theorem proving · Natural deduction · Natural intercalation calculus · Fitch diagrams · Normal proofs · Machine-oriented proof methods

Introduction: Context and Overview

Human-centered theorem proving aims to capture structures of ordinary mathematical arguments and to discover efficient strategies that are being (implicitly) used in their construction. To reach these goals, an evolving search procedure should be implemented and tested on

✉ Wilfried Sieg
sieg@cmu.edu

Farzaneh Derakhshan
fderakhs@andrew.cmu.edu

1 Department of Philosophy, Carnegie Mellon University, Pittsburgh, USA

computers. That is one focus of Ganesalingam and Gowers' project in their [15]; see also [14]. Their second, and perhaps central, focus is concerned with the output of the computer's proof search: the proof should be presented in a form that is not distinguishable from good mathematical writing. They view this human-centered approach as incompatible with a machine-oriented approach, because "it does not seem possible to reconstruct genuinely human-like writeups from the proofs produced by automated provers from the machine-oriented tradition" (Ganesalingam and Gowers [15], p. 254).

Indeed, there have been tensions between human- and machine-oriented approaches to theorem proving ever since the beginning of the subject in the middle of the twentieth century. They are vividly described for the period up to 1995 in Mac Kenzie [35] and that description is extended to the present time in Ganesalingam and Gowers [15], Sects. 1 and 2.[1] In both cases, proof search that models human processes of proof construction is opposed to proof search that explicitly views as irrelevant any resemblance between human and machine search processes. A very wide spectrum of approaches can be seen, ranging from Newell and Simon's "purely heuristic" theorem proving in the mid-1950s and early 1960s [36,54] to the latest "brute force" machine proving [22]. The tensions do not indicate, in our view, a deep conflict of principle, but rather a striking difference in programmatic goals. The broad goals can be pursued fruitfully and independently: Why shouldn't one try to discover a computational model of mathematical or logical reasoning and thus better understand a distinctively human capacity? Why shouldn't one try to exploit the "brute force" of SAT machines to solve challenging problems by proofs that are fully automatically obtained in ingenious ways, but are, because of their size, inaccessible to humans?

The very roots of the issue go back, however, to the early decades of the twentieth century, when it was realized that large parts, if not all, of mathematical practice could be developed in strictly formal frames. The question arose, whether there might be a mechanical process that decides, in finitely many steps, the provability of a statement from given axioms. This is the *Entscheidungsproblem* that was resolved negatively for first-order logic by Church and Turing in 1936 [9,57]. Leaving aside the problem of full mechanization (whose positive solution would have had dramatic consequences for our understanding of the nature of mathematics), the bare formalizability of mathematics was of fundamental importance for Hilbert's Program in the 1920s: to ensure the coherence of mathematical practice, the consistency of its formal representation was to be proved by elementary, so-called finitist means.

To be assured of formalizability, actual formalization was required, at least in principle: Hilbert and Bernays [1917–1918] developed analysis in the second-order part of the system of Whitehead and Russell's *Principia Mathematica*. Four years later, they introduced a new logical calculus that was to allow a more direct formalization and that was methodologically structured on the model of Hilbert's *Grundlagen der Geometrie*. There, a group of axioms had been formulated for each basic geometric concept (e.g., incidence and betweenness); here, axioms were introduced for each logical connective. Let us mention the axioms for conjunction and disjunction:

`&I(ntroduction)`	$\phi \to (\psi \to (\phi \,\&\, \psi))$
`&E(limination)`	$(\phi \,\&\, \psi) \to \phi$ and $(\phi \,\&\, \psi) \to \psi$
`∨I(ntroduction)`	$\phi \to (\phi \vee \psi)$ and $\psi \to (\phi \vee \psi)$
`∨E(limination)`	$(\phi \vee \psi) \to ((\phi \to \chi) \to ((\psi \to \chi) \to \chi))$

This calculus was used throughout the 1920s in the proof theoretic investigations of Hilbert and Bernays; it was basic for the two volumes of their *Grundlagen der Mathematik* [1934]

[1] See Harrison et al. [20] for a more comprehensive view on the history of interactive theorem proving.

and [1939]. At the center of those investigations was the consistency question for theories in which the mathematical practice of, say, number theory, analysis or set theory, could be formalized. However, in his 1917 Zürich talk *Axiomatisches Denken*, Hilbert had formulated a perspective for such investigations that is quite different in its emphasis. He articulated the idea of a theory concerning "the concept of the specifically mathematical proof" and suggested:

> …we must—that is my conviction—turn the concept of the specifically mathematical proof into an object of investigation, just as the astronomer considers the movement of his position, the physicist studies the theory of his apparatus, and the philosopher criticizes reason itself. (Hilbert [24], p. 1115)

Hilbert recognized in the next sentence that "the execution of this program is at present …still an unsolved task".

It was clear that the frame in which mathematics could be formalized was to serve as a tool to develop such a theory and address, for example, questions concerning *simplicity* and *uniqueness* of proofs. Hilbert had formulated such questions tentatively in 1900 as his 24th Paris problem.[2] *Proof theory*, as a part of mathematical logic, was initiated only in 1922 with an immediate focus on the consistency issue. However, the broader issues were not pushed aside. Hilbert articulated five years later the "fundamental idea" for the new subject as follows:

> The formula game … has, besides its mathematical value, an important general philosophical significance. For this formula game is carried out according to certain definite rules, in which the *technique of our thinking* is expressed. These rules form a closed system that can be discovered and definitively stated. The fundamental idea of my proof theory is none other than to describe the activity of our understanding, to make a protocol of the rules according to which our thinking actually proceeds. Thinking, it so happens, parallels speaking and writing …(Hilbert [25], p. 475)

The rules "according to which our thinking actually proceeds" are evidently not just the basic rules of a specific logical calculus. There are "conservative" ways in which a calculus can be extended to an inference mechanism that reflects *further techniques of our thinking*; they will be indicated below in Sect. 2.

Hilbert's programmatic remarks describe not only the ultimate purpose of his proof theory, but also one essential goal of G&G's work when the latter is examined more closely. G&G look, in the first place, for an automatic system that outputs proofs in a form "that is hard to distinguish from solutions that human mathematicians might write" (l.c., 253). They realize, in the second place, that such proofs must be based on the rules of a "logical calculus" that makes it possible "to convert each step of its [the system's] reasoning process into a piece of human-style prose, which will form the basis for the human-style output". Consequently, the system must be such that it "closely mirrors the way human mathematicians operate" (l.c., 255). That conclusion is in harmony with Hilbert's ideas when transposed from general proof theory to automated proof search.[3]

[2] On the discovery of this problem, which Hilbert did not include in the official Paris list, see [55].

[3] For Gowers, there was also a deep educational motivation that was forcefully expressed in an interview with the Notices of the American Mathematical Society reported in Diaz Lopez [11]. Sieg's AProS project originally had a more restricted educational goal; for reaching even that goal, a human-centered proof search procedure was needed; see Sieg [46]. These matters are briefly discussed in our concluding *Programmatic remarks*.

G&G then claim that for the design of a program that "imitates human thought" one has to place "severe restrictions on how it operates". That is seen in contrast to standard automated theorem provers; indeed, they assert, as we quoted already earlier, "... it does not seem possible to reconstruct genuinely human-like write-ups from the proofs produced by automated provers from the machine-oriented tradition". Thus, they argue, "In order to be able to produce such write-ups we had to use much more restricted proof methods than those available to modern provers". We agree that it seems to be close to impossible to obtain a genuinely human-like write-up from a machine-generated proof, say, in the sequent calculus or when the system uses tableaux or resolution methods. However, the step to then asserting that the proof methods must be restricted is not logically warranted. One might proceed in a quite different way, namely, use the basic logical tools that were invented to reflect deep aspects of human argumentation, expand the tool chest appropriately and take steps toward a *theory of proofs* in Hilbert's sense.[4] That is the direction that has been pursued by the AProS Project since the late 1980s.

In any case, the goal of constructing a human-centered theorem prover, that goal is shared between the work on the G&G-prover and on AProS. The means of reaching it are, on the surface, strikingly different. Section 1, *The second face of proof theory*, presents historical and systematic roots of a theory for mathematical proofs and developments towards *bi-directional* NI (natural intercalation) calculi for minimal, intuitionist and classical logic.[5] In Sect. 2, titled *Logical steps and mathematical extensions*, we use the features of these calculi to capture faithfully mathematical proofs: we formulate introduction and elimination rules for defined notions and operations; lemmas are shaped into rules, thus making it possible to take larger steps in proofs. G&G pursue the very same goal but, as they see it, from a direct mathematical perspective without explicitly using standard tools of the logicians.

How much G&G's approach is using tools from the logicians' workbench will be apparent when we describe in Sect. 3 the complex proof steps they are taking. There we analyze, under the heading *Mathematical steps and logical foundations* the central aspects of the G&G-prover, namely, its tactics for constructing proofs in human style. We show how the logical rules for just minimal logic suffice to mimic those tactics. The proofs of two elementary theorems concerning metric spaces that are presented as paradigmatic successes of G&G's approach have been obtained directly by AProS. In Sect. 4, *Two kinds of automated proof search*, we formulate the "waterfall strategy" G&G use, but also the strategic proof procedure of AProS. That is used then to juxtapose G&G's proof of one of the two theorems with its AProS proof.[6]

In our final *Programmatic remarks* we discuss how limitations of the G&G-prover can be addressed without giving up the goal of obtaining a well-articulated write-up. In part that is done by using AProS' broader minimal, intuitionist or classical frame. The crucial part

---

4 That point was made vivid to WS through [40] and Andrews' efforts in [1] to associate natural deduction proofs with mating proofs. (WS served on Pfenning's thesis committee and conducted a seminar on automated proof search with Andrews in the early 1990s.) According to (Andrews [2], p. 285), it was Larry Wos in [60] who "drew attention to the importance of investigating mappings between (a) clause representation and natural deduction representation of proofs and (b) corresponding search strategies".

5 These NI calculi are situated properly between sequent and natural deduction calculi. Their rules are formulated in Appendix 1. The general idea of "bi-directional" is that of "forward" and "backward chaining". The precise meaning of the term depends on the graphical representation of proofs. In case of the tree-representation of NIC proofs, the directions are "left to right" and "right to left", whereas in the Fitch-representation they are "top down" and "bottom up". See Sect. 2, but also Sect. 4, where the (direction of) distinct strategic steps are discussed in detail.

6 The proof of the other theorem, obtained by AProS, is presented in the second appendix of Sieg and Walsh [53]. That "other" proof was given in Ganesalingam and Gowers [14].

is played, however, by the radical separation of proof search and translation. The search of the G&G-prover is structured for obtaining a "simultaneous translation": every individual inference step is to be immediately paraphrased in English. The search is consequently forced to be "deterministic", not allowing any backtracking whatsoever. Their prover is thus restricted, as G&G put it, to "routine problems" whose solution reflects the "particularly simple rhetorical structure" of mathematical proofs. (These considerations are analyzed in Sect. 4 below.) AProS is, by contrast, a complete proof search mechanism with extensive non-deterministic steps; it is strategically guided, always goal-directed and bi-directional. Because of these features, the final "sequential" presentation of a proof does not reflect its stepwise construction. It is the sequence of construction steps in the final proof that can be taken as the deterministic underpinning for a write-up from the strategically constructed formal proof. That obviously can be done only, if the search procedure itself is "human-centered".

1 The Second Face of Proof Theory

It is standardly taken for granted that proof search is performed using a "formal representation" of mathematics. That involves a formal language and an appropriate inferential apparatus, i.e., a kind of logical calculus. Such formal frames are the object of proof theoretic investigations. Their role in proof theory can be seen from two different perspectives. The *first face* of proof theory seen from one perspective is most familiar to contemporary logicians, computer scientists and mathematicians. It is directed towards *formal deductions* in such frames and is concerned, or perhaps pre-occupied, with their meta-mathematical investigation. It aims, e.g., to establish the equivalence of different calculi, prove cut-elimination theorems for sequent calculi, investigate the reach of the epsilon-substitution method, prove normalization theorems for natural deduction calculi. We should emphasize that—in the context of proof search—both "formal representation" and "formal proof" have very open-ended meanings that cover also the complex mathematical, logical structures obtained by resolution methods.

There is a *second face* of proof theory. It is less familiar and is directed towards the *proofs of mathematical practice*. In the *Introduction*, we already mentioned Hilbert and Bernays' shift from the logical calculus of *Principia Mathematica* to a calculus with groups of axioms for each logical connective. The reasons for that shift were methodological but importantly also pragmatic, as the axioms justify derived rules reflecting steps in ordinary mathematical arguments and, thus, make it easier to formalize them.[7] Indeed, in his (Gentzen [18], p. 513), Gentzen goes beyond this local conception and views formal deductions as "images" (Abbilder) of mathematical proofs. He analyzes, in detail, a proof of Euclid's theorem on the infinity of primes and argues that his formalism is as close as possible to real mathematical reasoning. It is this second face of proof theory we are going to explore more deeply.

[7] The methodological reason is a parallelism to Hilbert [23], where groups of axioms were formulated for each basic notion. The significance of derived rules for capturing argumentation is emphasized in Hilbert and Bernays ([26], p. 139) and is explicit in Bernays [4], where rule-based formulations of the sentential calculus of Principia Mathematica are given. Gentzen ([18], p. 514) calls inference rules the "formal counterpart" of an informal reasoning step.

The goal of capturing the structure of mathematical arguments is underlined by Gentzen already in his *Urdissertation* when formulating a system of natural deduction [16].[8] He wrote under the heading *Reasons for setting up the calculus N1J*:

> The formalization of logical reasoning in Russell, Hilbert, Heyting (for intuitionistic reasoning) and others is rather far removed from the kind of reasoning as it is actually carried out (for example, in number theoretic proofs). In this way, one gains, however, considerable formal advantages. I will first set up a formalism ("calculus N1J") that is as close as possible to actual reasoning. After all, one can assume such [formalisms] have also certain advantages, and I believe that I can claim—on the basis of my further results—that this is indeed the case.[9,10]

Gentzen articulates the axioms of Hilbert and Bernays as *Introduction* and *Elimination* rules for the connectives. He adds a novel feature to the system that allows making and discharging assumptions. Gentzen views this feature as reflecting a crucial aspect of mathematical practice, as distinctive for natural reasoning (natürliches Schließen), and as "the most essential difference" to the axiomatic presentation of logic (Gentzen [17], p. 184). Here are the familiar rules that correspond to the above axioms.

$$\&\mathtt{I} \qquad \dfrac{\overset{\downarrow}{\phi} \quad \overset{\downarrow}{\psi}}{\phi \,\&\, \psi}$$

$$\&\mathtt{E} \qquad \dfrac{\overset{\downarrow}{(\phi \,\&\, \psi)}}{\phi} \quad \text{and} \quad \dfrac{\overset{\downarrow}{(\phi \,\&\, \psi)}}{\psi}$$

$$\vee\mathtt{I} \qquad \dfrac{\overset{\downarrow}{\phi}}{(\phi \vee \psi)} \quad \text{and} \quad \dfrac{\overset{\downarrow}{\psi}}{(\phi \vee \psi)}$$

$$\vee\mathtt{E} \qquad \dfrac{\overset{\downarrow}{(\phi \vee \psi)} \quad \overset{\overset{[\phi]}{\downarrow}}{\chi} \quad \overset{\overset{[\psi]}{\downarrow}}{\chi}}{\chi}$$

Note that proofs are no longer finite sequences of formulae satisfying the constraint that every element is either an axiom or has been inferred by a rule from one or more (earlier) formulae; they are now structured as trees. Already in their proof theoretic investigations of 1922, Hilbert and Bernays transformed linear proofs into trees—a crucial technical step that

[8] Cf. [58] and [47]. The *Urdissertation* was probably written in the fall or winter of 1932. The real dissertation [17] was submitted for publication on 21. July 1933.

[9] Gentzen repeats these remarks almost literally in (Gentzen [17], p. 176) and re-asserts on p.183, "Wir wollen einen Formalismus aufstellen, der möglichst genau das wirkliche logische Schließen bei mathematischen Beweisen wiedergibt."

[10] Here is the German text of the quotation: Die Formalisierung des logischen Schliessens bei Russell, Hilbert, Heyting (für das intuit. Schliessen) u.a. entfernt sich ziemlich weit von der Art des Schliessens, wie sie in Wirklichkeit (etwa bei zahlentheoretischen Beweisen) geübt wird. Dafür werden beträchtliche formale Vorteile erzielt. Ich will zunächst einmal einen Formalismus ("Kalkül N1J") aufstellen, der dem wirklichen Schliessen möglichst nahe kommt. Es ist doch anzunehmen, dass auch ein solcher gewisse Vorteile hat, und ich glaube, auf Grund meiner weiteren Ergebnisse behaupten zu können, dass dies der Fall ist.

was preserved in their later work; see [28,29]. That was known to Gentzen and is clearly reflected in Gentzen ([17], pp. 417–431), when the equivalence of the different classical calculi, namely, NK (natural deduction calculus), LK (sequent calculus), and LHK (Hilbert's axiomatic calculus), is established.

Gentzen mentions in the above long quotation from his *Urdissertation* that a calculus close to real reasoning might have "certain advantages" and asserts that his meta-mathematical results confirm this. He alludes to the normalization theorem for proofs in intuitionist first-order logic and the subformula property of normal proofs. In his real thesis, published as [17], Gentzen expands on this remark and asserts that the investigation of natural deduction calculi led him to a "very general theorem", he dubs *Hauptsatz* and formulates as follows:

> The Hauptsatz asserts that every purely logical proof can be brought into a certain normal form, which is, incidentally, by no means unique. The most essential properties of such a normal proof can be roughly expressed as follows: it does not make any detours. One does not introduce into it [the normal proof] any concept that is not contained in its final result and thus has to be necessarily used for obtaining it [the final result].[11,12]

For the proof of the *Hauptsatz* in his *Urdissertation*, Gentzen crucially uses the "formula tree structure" of proofs. It has become the common shape of natural deductions since their re-introduction into the proof theoretic mainstream through Prawitz's 1965-study *Natural Deduction*: $\Gamma \vdash \phi$ stands for "there is a proof-tree of formulae, such that the formula ϕ is at its root and all its open assumptions are contained in Γ".[13]

In contrast, the proofs in the natural deduction calculus of Gentzen [18] are formulated as "sequent tree structures": at every node of a proof tree we have now a sequent of the form $\Delta \supset \psi$ and $\vdash \Gamma \supset \phi$ expresses that "there is a proof-tree of sequents such that the sequent $\Gamma \supset \phi$ is at its root and indicates that the proof of ϕ uses assumptions in Γ". Let us formulate the rules of the modified natural deduction calculus just for disjunction:

$$\vee\mathrm{I} \qquad \frac{\overset{\downarrow}{\Gamma \supset \phi}}{\Gamma \supset (\phi \vee \psi)} \quad \text{and} \quad \frac{\overset{\downarrow}{\Gamma \supset \psi}}{\Gamma \supset (\phi \vee \psi)}$$

$$\vee\mathrm{E} \qquad \frac{\overset{\downarrow}{\Gamma \supset (\phi \vee \psi)} \quad \overset{\downarrow}{\Gamma, \phi \supset \chi} \quad \overset{\downarrow}{\Gamma, \psi \supset \chi}}{\Gamma \supset \chi}$$

[11] (Gentzen [17], p. 177). Gentzen continues by claiming that the Hauptsatz is valid for both intuitionist and classical logic, but for stating and proving it in a "convenient form" he had to introduce a special calculus, the sequent calculus. "Hierzu erwies sich der natürliche Kalkül nicht als geeignet. Zwar weist er schon die für die Gültigkeit des Hauptsatzes wesentlichen Eigenschaften auf, doch nur in seiner intuitionistischen Form, während der Satz vom ausgeschlossenen Dritten, …, im Hinblick auf diese Eigenschaften eine Sonderstellung einnimmt."

[12] Here is the German text of the quotation: Der Hauptsatz besagt, daß sich jeder rein logische Beweis auf eine bestimmte, übrigens keineswegs eindeutige, Normalform bringen läßt. Die wesentlichsten Eigenschaften eines solchen Normalbeweises lassen sich etwa so ausdrücken: Er macht keine Umwege. Es werden in ihm keine Begriffe eingeführt, welche nicht in seinem Endergebnis enthalten sind und daher zu dessen Gewinnung notwendig verwendet werden müssen.

[13] Prawitz [43] establishes the Hauptsatz or the normalization theorem also for (a subtly restricted, but equivalent form of) classical logic. There are further remarkable results in that work. Let us mention two: (1) "branches" in normal proofs can be uniquely divided into E- and I-parts, where only elimination rules are used in the E-part and only introduction rules in the I-part; (2) "contradictory pairs" of formulae are restricted to pairs with the negated component being a positive subformula of an assumption. Those features inform the AProS strategies for proof search in a crucial way.

Note that the genuinely logical actions are all taking place on the right-hand side of the sequent symbol $\supset$; Γ only records the assumptions that are used in the proof (and have not been discharged). These rules thus allow, as the standard rules, the construction of *non-normal* proofs, in which a formula occurrence is simultaneously the consequence of an introduction rule and the major premise of an elimination rule.

Gentzen used this sequent form of natural deduction in his consistency proof of Gentzen [18]. He switched, however, to the full sequent calculus in his [19]. Generations of proof theorists have been investigating sequent calculi in all kinds of forms and contexts. Some began using cut-free sequent calculi for automated proof search in classical logic; it seems that Hao Wang initiated such use in his [59]. Though Prawitz in 1965 rekindled an interest in natural deduction, the calculi were not viewed as suitable for automated proof search. The reason is simple: backward search was viewed as essential for automated proof search and seemed to be incompatible with natural deduction.[14] There was work on non-resolution methods in the computer science community, for example, by Bledsoe and his students; see [5,6]. Their work was partially guided by natural deduction considerations, but it did not have a fully grounded search space for natural deduction proofs.

Let us come back to the path that leads to the calculus underlying AProS. We just saw that "backward search" was viewed as essential for automated search and, simultaneously, as incompatible with natural deduction. In a more restricted sense, backward moves reflect a significant aspect of mathematical and ordinary informal argumentation. Taking a "backward move" from the conclusion that is to be reached, indeed, is pervasive; here are some concrete examples: (1) when seeking a proof of a conjunction, simplify the problem by seeking proofs of the conjuncts; (2) when attempting to prove a conditional, consider the antecedent of the conditional as an assumption and establish its succedent using also this newly made assumption; (3) when looking for a proof of a negation, assume the unnegated goal and prove a contradiction. Such backward moves cannot be seamlessly incorporated into the two standard representations of natural deduction proofs we just formulated. That led Sieg in the late 1980s to a cumbersome formulation of his *intercalation calculi*. We will focus here on a restricted version of the calculi, *normal intercalation calculi*, (NI calculus).

The key idea of the NI calculi was to use elimination rules only in the "forward" direction and (inverted) introduction rules in the "backward" direction, modeled in a unified syntactic configuration of "partial proofs".[15] That can be articulated when one takes as the starting-point not the Gentzen–Prawitz formulation of natural deduction, but rather Gentzen's sequent formulation. The proof trees have now "structured sequents" at their nodes. The latter are of the form $\Gamma; \Delta \supset \phi$ expressing that Γ contains assumptions and insisting that the elements of Δ have been obtained by successive elimination rules applied to an element γ of Γ that contains the goal ϕ as a (strictly) positive subformula. This sequence σ of formulae begins then with γ and the first element of Δ, if non-empty, is the result of applying an E-rule to γ. If σ ends with ϕ, then it is called a *full extraction sequence for* ϕ from γ. In general, even a full extraction sequence does not constitute a proof (in natural deduction) from γ to ϕ, as disjunctions, existential quantifiers and conditionals require special considerations. (They

[14] For example, Melvin Fitting wrote in his (Fitting [13], p. 95): "Hilbert systems are inappropriate for automated theorem proving. The same applies to natural deduction, since Modus Ponens is a rule in both." Only a few intrepid philosophical and psychological souls tried to use natural deduction calculi for automated proof search; among them were Pelletier, Pollock, and Rips; see [37,42] and [44]. The special issue of Studia Logica [41] was devoted to natural deduction; at the time, that was the best we could do to reflect work relating natural deduction and theorem proving.

[15] See [8,10,45,49,52], and [50].

raise “open questions” and are the source of “partiality”.) The latter will be described, when we treat AProS strategies; below we will indicate the treatment of disjunctions.

All the NIC rules are formulated as bottom-up rules that lead from partial proofs to partial proofs. A search tree is *completed* just in case all its top nodes are of the form

$$\Gamma; \Delta, \chi \supset \chi$$

or

$$\Gamma; \supset \chi.$$

In the first case, the sequence γ, Δ, χ is an extraction sequence for χ from γ; in the second case, χ is an element of Γ. The introduction rules are formulated as above, but they are used only for backward steps; for the full rule set see Appendix 1. The elimination rules are formulated relative to an extraction sequence σ; all the uniquely prescribed elimination steps are carried out for constructing the expanded partial proof tree.

$$\&\mathtt{E}\,(\sigma) \qquad \begin{array}{c} \Gamma; \Delta, (\phi \,\&\, \psi), \phi \supset \chi \\ \hline \Gamma; \Delta, (\phi \,\&\, \psi) \supset \chi \\ \uparrow \end{array}$$

Here, the successor of $(\phi \,\&\, \psi)$ in σ is ϕ. If ψ were the successor of the conjunction in σ, then ψ would be inferred. For disjunctions we consider the case that $(\phi \vee \psi)$ is a positive subformula of an element γ of Γ. The extraction sequence leading from γ to $(\phi \vee \psi)$ is indicated by $\sigma[\vee]$ and does not contain another disjunction. The elimination rule for disjunction is now formulated as follows, where $\sigma[\vee]$ is the sequence of formulae Δ, $(\phi \vee \psi)$:

$$\vee\mathtt{E}\,(\sigma[\vee]) \qquad \begin{array}{c} \Gamma, \phi; \supset \chi \quad \Gamma, \psi; \supset \chi \\ \hline \Gamma; \Delta, (\phi \vee \psi) \supset \chi \\ \uparrow \end{array}$$

In this calculus, the logical rules are applied both on the left and right-hand side of the sequent symbol: elimination rules are only applied on the left-hand side, whereas introduction rules are applied only on the right-hand side (and by “inversion”, i.e., by a backward application of the rule). In either case the application is goal-directed. A completed search tree constitutes a *NIC proof*.

The full calculus with rules for the other logical connectives, both in its classical and intuitionist form, is complete. Indeed, we have a *strengthened completeness theorem for natural deduction*: if χ is a logical consequence of Γ, then we can find a *normal* proof of χ from Γ. The NIC proof can be found strategically and can be effectively transformed into a normal natural deduction proof; these procedures are implemented in AProS. We will discuss further details in Sects. 2 and 4. In this section, we wanted to point out that our final logical calculus has important meta-mathematical properties that are crucial for proof search.[16] However, we want to emphasize the fact that it incorporates very broad structures of ordinary arguments, namely, making and discharging assumptions as well as constructing proofs bi-directionally. To make this point, we considered the rules for conjunctions and

[16] Troelstra and Schwichtenberg remark in their (Troelstra and Schwichtenberg [56], pp. 32, 60) that normality of natural deduction derivations is a “global property” as it involves “the order in which the rules are applied.” They took this fact as a reason for considering and investigating sequent calculi. We take it as a reason for locally exploiting the constraints on rule applications in (automated) proof search.

disjunctions and described how they evolved from the axioms of Hilbert and Bernays [27]. It is this historical development that illuminates the second face of proof theory.

In the next section we start out by focusing on a different format for natural deduction proofs that is more suitable for a presentation on a computer screen! To see very quickly the usefulness of that new presentation, we give a NIC proof of a very simple logical claim: $A\&(B \vee C); \supset A\&(C \vee B)$ and construct then the corresponding natural deduction proof.

$$\dfrac{\dfrac{A\,\&\,(B \vee C);\, A \supset A}{A\,\&\,(B \vee C);\, \supset A} \qquad \dfrac{\dfrac{\dfrac{A\,\&\,(B \vee C), B;\, \supset B}{A\,\&\,(B \vee C), B;\, \supset (C \vee B)} \quad \dfrac{A\,\&\,(B \vee C), C;\, \supset C}{A\,\&\,(B \vee C), C;\, \supset (C \vee B)}}{A\,\&\,(B \vee C);\, (B \vee C) \supset (C \vee B)}}{A\,\&\,(B \vee C);\, \supset (C \vee B)}}{A\,\&\,(B \vee C);\, \supset A\,\&\,(C \vee B)}$$

From this NIC proof one can obtain, in a completely canonical way, a normal natural deduction proof in the Gentzen–Prawitz style.

$$\dfrac{\dfrac{A\,\&\,(B \vee C)}{A} \qquad \dfrac{\dfrac{A\,\&\,(B \vee C)}{(B \vee C)} \quad \dfrac{[B]}{(C \vee B)} \quad \dfrac{[C]}{(C \vee B)}}{(C \vee B)}}{A\,\&\,(C \vee B)}$$

In the new proof presentation, these proofs will be depicted as "Fitch-diagrams"; however, they will be constructed in the NIC way, i.e., via E-rules from above and via inverted I-rules from below.

$$\boxed{\begin{array}{l} A\,\&\,(B \vee C) \\ (B \vee C) \\ \boxed{\begin{array}{l} B \\ (C \vee B) \end{array}} \\ \boxed{\begin{array}{l} C \\ (C \vee B) \end{array}} \\ (C \vee B) \\ A \\ A\,\&\,(C \vee B) \end{array}}$$

Let's look at this new format more systematically in the next section.[17]

2 Logical Steps and Mathematical Extensions

It is remarkable how right Gentzen was when asserting in his *Urdissertation* (1) that a calculus close to real argumentation might have "certain advantages" and (2) that his meta-mathematical results confirm this claim. It seems that this assertion can be extended beyond

[17] Prawitz ([43], Appendix C, section 1) points to Jaśkowski [30] as a second root for natural deduction calculi. Jaskowski, inspired by a question of Lukasiewicz in 1926, reports in his [30] that he answered the question by putting the "method of an arbitrary supposition" under the "form of structural rules" and settling the relation to the standard "theory of deduction", i.e., proving the equivalence of his calculus with a standard axiomatic one. For another, complementary perspective, see [38,39] and [21].

Gentzen's formulation of the natural deduction calculus for intuitionist first-order logic to the formulation of the NI calculus, both in its intuitionist and classical form. Indeed, the claim can be strengthened, as any NIC proof has the subformula property and is translatable into a normal natural deduction proof.

We are going to use for pragmatic reasons a quite different format for the presentation of proofs, as we indicated at the end of Sect. 1. That presentation is derived from Jaskowski and Fitch's diagrammatic indication of sub-proof relations. Jaskowski calls it "the method of supposition" (Jaśkowski [30], p. 6), whereas Fitch named it "the method of subordinate proofs" (Fitch [12], p. iv). Fitch asserts that it "vastly simplifies the carrying out of complicated proofs". As to the origins of his method, Fitch remarks later on, "The method of subordinate proofs was suggested by techniques due to Gentzen and Jaskowski." Fitch refers to their respective 1934 papers. We will see that the construction and presentation of proofs in our paper differ from Fitch's in three dramatic ways. The construction is (1) goal-directed, (2) bi-directional, and (3) strategic. That means, in particular, there is no general assumption or hypothesis rule and no rule of re-iteration.

A careful modern description of Jaskowski's system and variants thereof is found in Appendix C of [43]. The format of "Fitch diagrams" is well known; we adopted it already in the example from Sect. 1, indicating the subordination of proofs by the "boxes" that had been initially used by Jaskowski.[18] Here is the construction of the proof in the Fitch-Jaskowski calculus of the simple logical theorem $(\neg A \to \neg B) \to (B \to A)$—that was given in Jaskowski and then reviewed in Prawitz. The rule Hyp(othesis) allows considering the antecedent of this conditional as a "proof".

$$\boxed{(\neg A \to \neg B)}$$

The same rule leads to

$$\boxed{\begin{array}{l}(\neg A \to \neg B)\\ \qquad \boxed{B}\end{array}}$$

Now we reiterate the hypothesis $(\neg A \to \neg B)$ into the sub-proof we opened with B.

$$\boxed{\begin{array}{l}(\neg A \to \neg B)\\ \qquad \boxed{\begin{array}{l}B\\ (\neg A \to \neg B)\end{array}}\end{array}}$$

We use the rule Hyp again to extend the proof:

$$\boxed{\begin{array}{l}(\neg A \to \neg B)\\ \qquad \boxed{\begin{array}{l}B\\ (\neg A \to \neg B)\\ \qquad \boxed{\neg A}\end{array}}\end{array}}$$

Continuing in this *purely forward march* one constructs the following completed proof.

[18] Jaskowski claims in a footnote of (Jaśkowski [30], p. 7) that he had used boxes in 1926 to articulate his method of supposition. He gives in that very footnote, as an example, the proof we will construct next.

1.	$(\neg A \to \neg B)$	Hyp
2.	$\quad B$	Hyp
3.	$\quad (\neg A \to \neg B)$	Reit : 1
4.	$\quad\quad \neg A$	Hyp
5.	$\quad\quad (\neg A \to \neg B)$	Reit : 3
6.	$\quad\quad \neg B$	$\to$E : 5, 4
7.	$\quad\quad B$	Reit : 2
8.	$\quad A$	$\neg$E : 6, 7
9.	$(B \to A)$	$\to$I : 2, 8
10.	$(\neg A \to \neg B) \to (B \to A)$	$\to$I : 1, 9

Notice that every boxed sequence of items is a "proof", where each item either (1) is a hypothesis or (the consequence of) a subordinate "proof", (2) has been obtained by reiteration or (3) has been inferred by a logical rule applied to formulae in the box. This way of inductively generating "proofs" is meta-mathematically motivated: it allows proving by induction one direction of Jaskowski's central result, namely, that his calculus is equivalent to an axiomatic formulation in the style of Hilbert [25].

Fitch's name for his way of proceeding, "method of subordinate proof", is most appropriate and points to its characteristic feature. Our modified Fitch-diagrammatic system would best be called the "method of partial proofs" or the "method of reasoning with gaps". The above logical claim can be proved via a sequence of *partial proofs*, when the very articulation of the goal is counted as a partial proof. Coming back to Jaskowski's example, we have this goal articulation:

$(\neg A \to \neg B) \to (B \to A)$ Goal

Inversion with $\to$ I leads to the partial proof whose gap is indicated by an ellipsis:

1.	$(\neg A \to \neg B)$	Hyp
	$\quad\vdots$	
2.	$(B \to A)$	Goal
3.	$(\neg A \to \neg B) \to (B \to A)$	$\to$I : 1, 2

The gap can be narrowed by the same rule yielding:

1.	$(\neg A \to \neg B)$	Hyp
2.	$\quad B$	Hyp
	$\quad\vdots$	
3.	$\quad A$	Goal
4.	$(B \to A)$	$\to$I : 2, 3
5.	$(\neg A \to \neg B) \to (B \to A)$	$\to$I : 1, 4

There is no I-rule or E-rule that is applicable to this partial proof. However, we can refute classically and obtain:

1.	$(\neg A \to \neg B)$	Hyp
2.	B	Hyp
3.	$\neg A$	Hyp
	$\vdots$	
4.	$\bot$	Goal
5.	A	$\neg$E : 3, 4
6.	$(B \to A)$	$\to$ I : 2, 5
7.	$(\neg A \to \neg B) \to (B \to A)$	$\to$ I : 1, 6

So, we must find an appropriate "contradictory pair" of formulae.[19] The pair B and $\neg B$ is suitable. We have B already accessible and $\neg B$ can be obtained by $\to E$. So, we have this final proof:

1.	$(\neg A \to \neg B)$	Hyp
2.	B	Hyp
3.	$\neg A$	Hyp
4.	$\neg B$	$\to$ E : 1, 3
5.	$\bot$	$\bot$ I : 4, 2
6.	A	$\neg$E : 3, 5
7.	$(B \to A)$	$\to$ I : 2, 6
8.	$(\neg A \to \neg B) \to (B \to A)$	$\to$ I : 1, 7

Every step in this sequence of partial proofs is strategically motivated by the question, "Which inference rule allows bridging (narrowing) the gap between hypotheses (assumptions) and goals?" This question is answered by the strategic moves *extraction*, *inversion* and *refutation*: extraction consists of elimination steps that are uniquely determined by an extraction sequence; inversion consists of a backward application of an introduction rule to the goal; refutation applies not-elimination to the goal in the backward direction. In the NI calculus, hypotheses are consequently introduced only in the context of their elimination, reiteration is replaced by back-references to "accessible" formulae, and the strategic moves force bi-directionality in the general case. We pointed out that Jaskowski-Fitch proofs are generated by forward moves only. Thus, the sequence of construction steps is reflected by the sequence of formulae in the final proof. That is most decidedly not the case in our calculus. For example, when trying to prove a negation $\neg\phi$ we assume ϕ and then plan to prove a contradiction.

[19] The use of the $\bot$ symbol is discussed in Appendix 1. Let us mention here that $\bot$ is used only in proofs to note that a pair of contradictory formulae is to be proved. As we are searching for normal proofs, the choice of contradictory pairs can be restricted to those that have as their "negative" part a formula that has a positive occurrence in an assumption or premise (on which the line with $\bot$ depends).

It is precisely this planning aspect and the resulting higher-level structure that is crucial for the smooth translation of formal proofs into human-style proofs; that will be discussed briefly in the concluding *Programmatic remarks*.

Up to now, our considerations have been exclusively concerned with logic, and the rules we have been appealing to are the absolutely basic steps of argumentation. Indeed, the rules have been *reduced* to those that allow us to use (via eliminations) and *infer* (via introductions) logically complex statement. Gentzen emphasized this reductive viewpoint and contrasted it with "real reasoning" that extends beyond the basic logical moves:

> In actual reasoning, one frequently skips a sequence of inferences that have become obvious by force of habit. This is not to be done in the formalization that follows. On the contrary, every step of a "proof" is to be individually represented, because that is crucial for uncovering the fundamental components of logical reasoning.[20]

Using the analysis of the basic components of logical reasoning by Gentzen (following Hilbert and Bernays), we can join sequences of steps into a single, directly intelligible one, i.e., formulate a derived rule. Consider, as an example, the application of de Morgan's Law to infer $(\neg A \,\&\, \neg B)$ from $\neg(A \vee B)$. It is very useful in an argument, as the rule's consequence, $(\neg A \,\&\, \neg B)$, allows using via $\&\,\mathtt{E}$ the component conjuncts.

We are going to consider now the pertinent and important example of *bounded quantifiers*. In the context of set theory, we have formulae of the form $(\forall x)(x \in b \rightarrow \gamma(x))$ and $(\exists x)(x \in b \,\&\, \gamma(x))$ that are abbreviated by $(\forall x \in b)\ \gamma(x)$ and $(\exists x \in b)\ \gamma(x)$. For these bounded quantifiers, we can formulate appropriate introduction and elimination rules. Let us do that just for the *bounded universal quantifier*; on the left, the $\mathtt{I}$-rule is represented, whereas the $\mathtt{E}$-rule is on the right:

$$\begin{array}{c} \boxed{\begin{array}{l} x \in b \\ \vdots \\ \gamma(x) \end{array}} \\ (\forall x \in b)\ \gamma(x) \end{array} \qquad\qquad \begin{array}{l} t \in b \\ \vdots \\ (\forall x \in b)\ \gamma(x) \\ \vdots \\ \gamma(t) \end{array}$$

In the case of the introduction rule, the standard variable condition must be satisfied. These introduction and elimination rules can be generalized to any "bounded formula" of the form $(\forall x)(\phi(x) \rightarrow \gamma(x))$ or $(\exists x)(\phi(x) \,\&\, \gamma(x))$. One might even introduce appropriate abbreviations, say, $(\forall x : \phi(x))\, \gamma(x)$ and $(\exists x : \phi(x))\, \gamma(x)$. We call the appropriate rules simply bounded quantifier rules. The use of these *bounded quantifier rules* is one modest way of *abbreviating* proofs that both preserves their formality and makes them more intelligible. The conclusion of the I-rule for the bounded universal quantifier can be obtained in the basic calculus in the following way:

[20] Here is the German text: Beim wirklichen Schliessen überspringt man oft eine Reihe von Schlüssen, die durch lange Gewohnheit selbstverständlich geworden sind. Dies soll bei der folgenden Formalisierung nicht nachgemacht werden, sondern jeder Schritt eines "Beweises" soll einzeln wiedergegeben werden, da es darauf ankommt, die Grundbestandteile des logischen Schliessens aufzudecken.

$$\begin{array}{l} \boxed{\begin{array}{l}\phi(x)\\ \vdots\\ \gamma(x)\end{array}} \\ (\phi(x) \rightarrow \gamma(x)) \\ (\forall x)(\phi(x) \rightarrow \gamma(x)) \end{array}$$

As emphasized, the rule for bounded universal quantification is directly intelligible and avoids, what is justly viewed as a purely "bureaucratic" logical step.

Now, we will extend the joining of atomic steps into a single step of "real reasoning" from logic to mathematics. Recall that the meaning of the logical connectives is captured by their E- and I-rules. The same can be done for mathematical definitions through elimination and introduction rules obtained from their defining bi-conditionals. Here are two examples. The first concerns the powerset operation in set theory. The defining axiom for that operation is: $(\forall x)(x \in \mathcal{P}(a) \leftrightarrow x \subseteq a)$. Its associated rules are:

$$\mathcal{P}\text{–I} \quad \begin{array}{l} x \subseteq a \\ \vdots \\ x \in \mathcal{P}(a) \end{array} \qquad\qquad \mathcal{P}\text{–E} \quad \begin{array}{l} x \in \mathcal{P}(a) \\ \vdots \\ x \subseteq a \end{array}$$

Here we took for granted that we have defined already the subset relation; the latter has of course the following defining axiom: $a \subseteq b \leftrightarrow (\forall x \in a)x \in b$. Its associated rules are:

$$\subseteq\text{-I} \quad \begin{array}{l} (\forall x \in a)\, x \in b \\ \vdots \\ a \subseteq b \end{array} \qquad\qquad \subseteq\text{- E} \quad \begin{array}{l} a \subseteq b \\ \vdots \\ (\forall x \in a)\, x \in b \end{array}$$

These definitional rules have the same abbreviatory power as the rules for the logical connectives when compared to their respective axiomatic formulations.[21]

One of the last logic students in Hilbert's Göttingen and a friend of Gentzen, Saunders Mac Lane, took steps toward a faithful representation of mathematical proofs. In his 1934-thesis [31] *Abgekürzte Beweise im Logikkalkul* [sic], he pursued the goal of using "complex" rules to exhibit not only the broader, surveyable structure of individual proofs, but also "the structure of any body of mathematical doctrine" (Mac Lane [32], p. 118). His work is a remarkable attempt to uncover the rational structure of proofs and strategies for finding them.[22] To do so

[21] In AProS there are generic introduction and elimination rules for definitions, `Def-I` and `Def-E`. These generic rules use the correct definition automatically.

[22] At the end of his [33], Mac Lane emphasized: "There remains the real question of the actual structure of mathematical proofs and their strategy. It is a topic long given up by mathematical logicians, but one which still—properly handled—might give us some real insight." (p. 66)—Mac Lane expressed the very radical aspirations of the thesis in letters to his mother, written when working on the thesis; for example, on April 20, 1933, he asserted: "Perhaps I have time to tell you about my new discovery. It is a new symbolic logic for mathematical proof. It applies, as far as I can see, to all proofs in all branches of mathematics (a rather large order). It makes it possible to write down the proof of a theorem in a very much shorter space than the usual method and at the same time it makes the proof of the theorem much clearer. In essence, it eliminates practically all the long mechanical operations necessary to prove a theorem. It is only necessary to give in

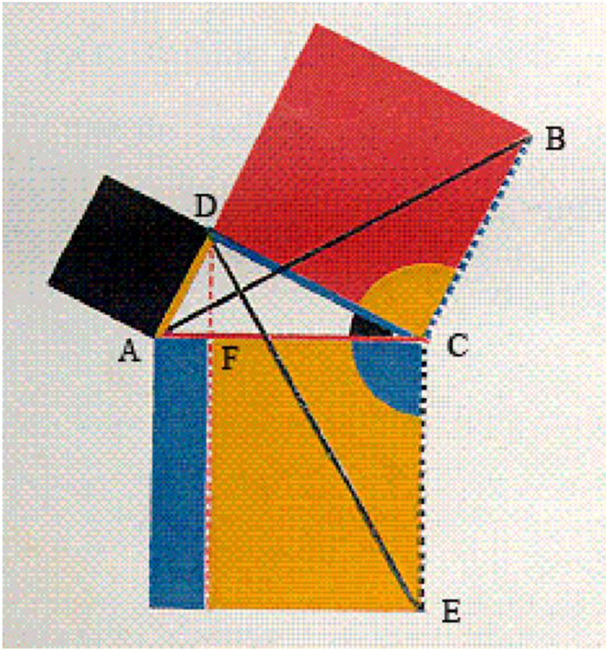


Fig. 1 Euclid's diagram for the Pythagorean Theorem

efficiently and intelligibly, atomic logical steps must be synthesized into larger mathematical steps and given direction by an overarching plan. That has to be done when moving from the *local* search for logical proofs in the basic system to the *global* search for mathematical proofs relative to a particular theory or a foundational frame like ZF, Higher-Order Logic or Martin–Löf-type theory. We are not going to explore that in any detail, as the comparison to G&G's work does not require it. However, we want to point to critical elements and refer to the discussion of *natural formalization* in [53]. In that paper, a complete and computer verified formalization of the proof of the Cantor–Bernstein Theorem is carried through; that is done in a definitional extension of Zermelo–Fraenkel set theory.

The critical elements of a natural formalization are not a modern invention. The classical and paradigmatic case exhibits them clearly and in a very beautiful way. Not surprisingly, we are thinking of the development in Book I of Euclid's *Elements* that culminates in the proof of the Pythagorean theorem and its converse. The proof of the Pythagorean theorem is short, as it uses earlier propositions or lemmas as "local axioms". Those earlier lemmas make mathematically meaningful claims concerning complex defined notions (triangles and rectangles, for example); they have been established, in turn, from other lemmas with the aim of "deepening the foundations". Ultimately, the "lemmas" used in proofs must be Euclid's axioms and common notions. We examine one step in Euclid's proof. Figure 1 shows the colorful diagram from Byrne's edition of the *Elements* [7]. The labels to special points have been added by us. In the central step (of relating the area of the square over the right leg to

Footnote 22 continued
sequence the leading ideas of the proof. In fact, once these leading ideas are given—together with a few directions—then it becomes possible to compute from the leading ideas just what the proof of the theorem will be. In other words, once the leading ideas are given, all the rest is a purely mechanical sort of job." (Mac Lane [34], pp. 60–61)

part of the area of the square over the hypotenuse) one wants to show that the triangles DCE and ABC are equal (in area, as we would say). This is obtained by first observing that (i) the sides CE and AC are equal, (ii) the sides DC and BC are equal and (iii) the angles enclosed by DC and CE, respectively AC and CB, are equal as well. By Euclid's Proposition I.4, we can infer that the triangles are equal. Proposition I.4 just states the familiar SAS principle, namely, if two triangles have two equal sides and the angles enclosed by those sides are equal, then the triangles are equal.

This example points to two important features of the conceptual organization of Book I, (i) the systematic introduction of notions and the proving of lemmas for these concepts and (ii) the explicit hierarchical structure that allows us to use lemmas as rules in later arguments. The application of the general inference mechanism *lemmas-as-rules* for a lemma of the form (H_1 & $\cdots$ & $H_n \rightarrow C$) requires only proofs of the H_i (on lines that are accessible from the line on which C is to be concluded). The mathematical frame in combination with this inference mechanism is clearly a tool for shortening proofs. In addition, they divide the overall argument into intelligible chunks with the aim of aiding the (limited human) understanding.[23] The next ambitious step for such proof theoretic investigations will move the human-centered interactive verification to fully automated proof search. We mention an early success for this broad methodology: AProS found the proofs of Gödel's Incompleteness Theorems and results like Löb's Theorem *RELATIVE* to the local axiomatic assumption of representability and derivability conditions; that was presented in Sieg and Field [51].

We opened the discussion of our Fitch-diagrammatic presentation of proofs in this section by claiming "practical reasons" for its use; it is also logically natural. Though it is not suitable for writing proofs in a linear fashion on paper, it is ideal for working in a computer-supported interface joining logical and presentational perspicuity. In particular, some of the operations on partial proofs induce a higher-level structure of argumentation that is also guided by meta-mathematical facts. We discussed that after presenting our proof for the Jaskowski theorem as a typical example of "an argument by contradiction". The other purely logical theorem we proved at the end of the Sect. 1 was paradigmatic for "an argument by cases" reflected by the use of the $\vee$-elimination rule. The "discrepancy" between the orderly way of listing the lines of a (completed) proof and the actual steps of the proof construction play a major role in distinguishing the G&G-prover from AProS.

3 Mathematical Steps and Logical Foundations

In the previous section, we built the proof search machinery from the bottom up, using `E`- and `I`-rules for the logical connectives of first-order logic; we expanded this rule-based approach to mathematical definitions and integrated it—via the mechanism of lemmas-as-rules—with a conceptually and hierarchically organized mathematical frame. The latter could be, for example, Book I of Euclid's *Elements* or the basic development of ZF set theory. Our explicit intention was to represent formally what Gentzen called the "real reasoning" in mathematics or, expressing it differently, to ground central aspects of mathematical proof in a powerful logical calculus. In the current section, we will follow G&G's top-down approach, isolating from mathematical practice central inferential steps or *tactics*. G&G focus on *applying tactics* and view a tactic as an *applying* one if it "corresponds to what human mathematicians would call applying an assumption, result or definition". (Ganesalingam and Gowers [15], p. 274)

[23] The analysis of the mathematical frame for the proof of the Pythagorean theorem in Book I and the analogy to the proof of the Cantor–Bernstein Theorem are both found in [48], Sect. 4 that is entitled *Local Axiomatics*.

Not surprisingly, there are genuine similarities between top-level elements of the search procedures carried out by the G&G-prover and AProS. There are also striking dissimilarities that arise mostly from G&G's concern with the *deterministic generation* of proof write-ups and the resulting restrictive focus on, what they call, *routine problems*.

The underlying strategic decisions will be discussed in Sect. 4. Here we are going to focus on the question, "What are the tactical *mathematical* steps of the G&G-prover?" The subsidiary issue concerns the logical foundation of these tactics. The syntactic structure of assumptions and goals, which G&G call *targets*, guide the tactical steps. The underlying language is that of first-order logic. However, no proof steps are associated with negation (though it is officially part of the language) and "pure" conditionals are not even part of the language; see (l.c., pp. 267–268). The formulae are consequently built up inductively from atomic formulae by closing under the formation of negations, disjunctions, conjunctions, universally and existentially quantified statements but also by using "bounded universal quantification" to construct formulae of the form

$$(\forall x_1, \cdots, x_m)(F_1 \& \cdots \& F_n \rightarrow F_{n+1}).$$

The basic syntactic configurations involved in the proof search are called *boxes*.[24] They are inductively defined together with the notion of a *target* as follows:
1. The special symbol $\mathbb{T}$ and any formula is a target;
2. If $A_1 \cdots A_m$ are formulae and $T_1 \cdots T_n$ are targets, then the configuration

$$\begin{array}{|c|} \hline A_1 \\ \cdots \\ A_m \\ \hline T_1 \\ \cdots \\ T_n \\ \hline \end{array}$$

is a box;
3. If $B_1 \cdots B_k$ are either boxes or $\mathbb{T}$, then the list $B_1 \cdots B_k$ is a target.
Note that an individual box can thus be a target. *Tactics* are functions that transform boxes into boxes; the box displayed above expresses the goal of "proving the formula" $A_1 \& \cdots \& A_m \rightarrow T_1 \& \cdots \& T_n$, prefixed by universal quantifiers over all the free variables in $A_1 \& \cdots \& A_m \rightarrow T_1 \& \cdots \& T_n$.[25] Tactics applied to formulae A_i lead to "forward" moves, whereas tactics applied to formulae T_j lead to "backward" moves. G&G describe their prover as "goal-oriented" and "tactic-based". Let us see what tactics are being used, what "forward" and "backward" moves amount to, and whether the prover is genuinely goal-oriented.

G&G provide a list of tactics grouped into five broad categories: (1) Deletion, (2) Tidying, (3) Applying, (4) Suspension, and (5) Equality substitution. They stress that their efforts "have been concentrated less on the tactics themselves and more on how the program chooses which tactic to apply". We reverse matters and focus on the individual tactics to see what G&G take

24 We follow G&G and do not include sequences of variables in the definition of boxes. They assert, "we suppress the variables as they tend to clutter the exposition" (l.c., p. 268). For us there is also a second reason. These variables are involved in choosing instantiations of existentially quantified statements; that problem is addressed by AProS in a quite different way. It is discussed below.

25 G&G introduce in this context a more general concept of "formula corresponding to a box"; that can be defined easily by recursion on the very concept of box and target: if a target is a list of boxes then the formulae associated with them are joined into a disjunction.

to be the crucial components for modelling "human thought processes". In the next section, we will discuss the "justification for the order of priority" of the application of tactics and see them in action for the proof of one theorem from metric space theory. We return now to the individual tactics.

"..., the aim of the program is to reach a goal with no target." That is asserted at the very beginning of section 5.5, where the tactics of category (1) are discussed. The crucial logical point of the main deletion tactic is the removal of a target as soon as it has been established from the given assumptions. Tidying and applying tactics are the central logical and mathematical steps the G&G prover takes; they are complemented by "suspension" tactics concerning existential quantifiers. After fixing the role of the special symbol $\top$, we will divide the G&G tactics in two big groups. The first group consists of all the forward moves, whereas the second group contains then, of course, all the backward moves. The individual moves are categorized in terms of the AProS steps induced by the elimination and introduction rules for the logical connectives, for definitions and lemmas, i.e., use of the mathematical frame or—as G&G put it—use of the library. As we will see, that differs from G&G's "forwards" and "backwards" steps, but only in inessential ways.

Remark
Readers may wish to jump ahead and look at the reformulation of the G&G approach in terms of the NI calculus; see Appendix 2. We think that clarifies and illuminates the purely logical aspects of boxes and tactics.[26]

In our listing of all the tactics, please keep in mind the informal understanding of a box in this basic form:

$$\begin{array}{|c|}\hline A_1 \\ \cdots \\ A_m \\ \hline T_1 \\ \cdots \\ T_n \\ \hline \end{array}$$

Here, all the $A_1, \cdots A_m$ and $T_1, \cdots, T_n$ are formulae. The goal is to prove T_1 and $T_2 \cdots$ and T_n from the assumptions $A_1, \cdots, A_m$. This goal has been reached as soon as the initial box of the above form has been transformed into a box of the form:

$$\begin{array}{|c|}\hline A_1 \\ \cdots \\ A_m \\ \hline \\ \hline \end{array}$$

[26] Our remark conflicts, of course, with G&G's rather dismissive observation on page 262, where they distinguish sharply between their "sequents" and the ordinary sequents of, well, of Gentzen's sequent calculi: only one formula as a succedent or consequent. They write: "Consequents [interpreted disjunctively] may be more attractive from a logical point of view for symmetry reasons, but the convention we adopt …is in our judgment closer to how humans would think of 'the current state of play' when they are in the middle of solving a problem, and that is more important to us than logical neatness." This remark is only applicable to the sequent calculus for classical logic; in the case of intuitionist logic the consequent is either empty or consists of one formula. As G&G treat only intuitionist logic, the difference in the basic form does not matter at all. It is the completely different "meaning" of the sequents that is significant here; see Appendix 2.

I.e., the box does not contain any targets and the list of "assumptions" may have been expanded by forward steps.

The elimination of targets is achieved using the special symbol $\mathbb{T}$. Under specific conditions a target is replaced by $\mathbb{T}$ and that occurrence of $\mathbb{T}$ is then immediately removed. That is reflected by the `I`- and `E`-rules for $\mathbb{T}$.

$$\begin{array}{|c|}\hline A_1 \\ \cdots \\ A_j \\ \cdots \\ A_m \\ \hline T_1 \\ \cdots \\ T_k \\ \cdots \\ T_n \\ \hline \end{array} \quad \xRightarrow{\mathbb{T}\text{-}\mathtt{I}} \quad \begin{array}{|c|}\hline A_1 \\ \cdots \\ A_j \\ \cdots \\ A_m \\ \hline T_1 \\ \cdots \\ \mathbb{T} \\ \cdots \\ T_n \\ \hline \end{array}$$

Corresponding to clause 3 of the inductive generation of boxes and targets, there is a second $\mathbb{T}$-`I` rule.

$$\begin{array}{|c|}\hline A_1 \\ \cdots \\ A_j \\ \cdots \\ A_m \\ \hline T_1 \\ \cdots \\ B_1, \cdots \begin{array}{|c|}\hline C_1 \\ \cdots \\ C_\ell \\ \hline \\ \hline \end{array} \cdots, B_k \\ \cdots \\ T_n \\ \hline \end{array} \quad \xRightarrow{\mathbb{T}\text{-}\mathtt{I}} \quad \begin{array}{|c|}\hline A_1 \\ \cdots \\ A_j \\ \cdots \\ A_m \\ \hline T_1 \\ \cdots \\ B_1, \cdots, \mathbb{T}, \cdots, B_k \\ \cdots \\ T_n \\ \hline \end{array}$$

The primary logical condition for applying the first tactic, their *removeTarget*, is the syntactic identity of A_j and T_k. G&G indicate, however, additional conditions [l.c., p. 247]. They give as an example that T_k is a bounded existential statement $(\exists u : P(u))Q(u)$ and that the assumptions contain $P(x)$ and also $Q(x)$. They state then, "The other circumstances are similar.". The example suggests that the target can be obtained by a single "single" step, here, by the introduction rule for a bounded existential statement, $\exists$: `I`; that rule is discussed below. The point of the second $\mathbb{T}$-`I` is to indicate that one of the boxes of the disjunctive box-target has been completed.

The *elimination* of $\mathbb{T}$ is divided into two corresponding cases. Here are our formulations:

$$\begin{array}{|c|}\hline A_1\\ \cdots\\ A_m\\ \hline T_1\\ \cdots\\ \mathbb{T}\\ \cdots\\ T_n\\ \hline\end{array} \quad \stackrel{\mathbb{T}\text{-}\mathtt{E}}{\Longrightarrow} \quad \begin{array}{|c|}\hline A_1\\ \cdots\\ A_m\\ \hline T_1\\ \cdots\\ T_n\\ \hline\end{array}$$

$$\begin{array}{|c|}\hline A_1\\ \cdots\\ A_m\\ \hline T_1\\ \cdots\\ B_1, \cdots, \mathbb{T}, \cdots, B_k\\ \cdots\\ T_n\\ \hline\end{array} \quad \stackrel{\mathbb{T}\text{-}\mathtt{E}}{\Longrightarrow} \quad \begin{array}{|c|}\hline A_1\\ \cdots\\ A_m\\ \hline T_1\\ \cdots\\ T_n\\ \hline\end{array}$$

The first $\mathbb{T}$-$\mathtt{E}$ rule is G & G's *deleteDone* tactic, whereas the second is their *deleteDoneDisjunct*.[27]

We now come to the genuinely logical steps that are associated with the connectives of first-order logic as used by G&G: disjunction, conjunction, (bounded) universal and (bounded) existential quantifiers. In addition, we discuss G&G's rules for definitions of special forms and for theorems available from a library. We are first considering forward steps reflecting appropriate elimination rules applied to assumptions, i.e., formulae in the upper part of a box.

[27] There are two additional tactics that allow the deletion of assumptions, namely, *deleteDangling* and *deleteUnmatchable*; the assumptions to be deleted have been previously used and " have no obvious use", (l.c., p. 272). The general purpose of these tactics is described in G&G's section 3.3 on pp. 264–265 and will be discussed below.

Forward steps

$$\begin{array}{|c|}\hline A_1\\ \dots\\ P\vee Q\\ \dots\\ A_m\\ \hline T_1\\ \dots\\ T_n\\ \hline\end{array} \quad\xRightarrow{\vee\text{-}\mathtt{E}}\quad \begin{array}{|c|}\hline A_1\\ \dots\\ A_m\\ \hline \begin{array}{|c|}\hline P\\ \hline T_1\\ \dots\\ T_n\\ \hline\end{array}\\ \\ \begin{array}{|c|}\hline Q\\ \hline T_1\\ \dots\\ T_n\\ \hline\end{array}\\ \hline\end{array}$$

This is G&G's *splitDisjunctiveHypothesis* tactic that corresponds to the standard form of the mathematical "argument by cases". However, together with the corresponding $\vee$-`Introduction` rules (their *splitDisjunctiveTarget*), G&G describe the proper treatment of these rules as "work in progress". They give the following reason: "they [these two disjunction tactics] are important for generating certain human-like proofs, but (because they split the proof into cases) they are not compatible with the incremental, concatenative tactic-by-tactic generation of write-ups. We intend to extend the write-up generation mechanism to accommodate these tactics in future versions. Note that neither tactic is used in the problems we evaluate on." (l.c., p. 273)

The elimination rule for conjunction, &-`E`, can be formulated as follows:

$$\begin{array}{|c|}\hline A_1\\ \dots\\ (P\&Q)\\ \dots\\ A_m\\ \hline T_1\\ \dots\\ T_n\\ \hline\end{array} \quad\xRightarrow{\&\text{-}\mathtt{E}}\quad \begin{array}{|c|}\hline A_1\\ \dots\\ (P\&Q)\\ \dots\\ A_m\\ P\\ \hline T_1\\ \dots\\ T_n\\ \hline\end{array} \quad or \quad \begin{array}{|c|}\hline A_1\\ \dots\\ (P\&Q)\\ \dots\\ A_m\\ Q\\ \hline T_1\\ \dots\\ T_n\\ \hline\end{array}$$

The rule is used only implicitly in the context of the tactic *elementaryExpansionOfHypothesis*, i.e., for the elimination of definitions (l.c., p. 278). Similarly, the standard elimination rule for universal quantifiers, $\forall$-`E`, is only used implicitly in the context of the elimination of bounded universal quantifiers that, in turn, is employed for the elimination of definitions. The latter is formulated as follows:

$$\begin{array}{|c|}\hline A_1\\ \dots\\ P(x)\\ \dots\\ (\forall y : P(y))\,Q(y)\\ \dots\\ A_m\\ \hline T_1\\ \dots\\ T_n\\ \hline\end{array} \quad \overset{\forall\text{:-E}}{\Longrightarrow} \quad \begin{array}{|c|}\hline A_1\\ \dots\\ P(x)\\ \dots\\ (\forall y : P(y))\,Q(y)\\ \dots\\ A_m\\ Q(x)\\ \hline T_1\\ \dots\\ T_n\\ \hline\end{array}$$

This tactic is clearly a forward step, but as the &-E it is not goal-directed. G&G call this rule *universalmodusponens* and includes, in principle, their *libraryreasoning*. There is a second tactic for bounded universal statements that is considered by G&G as a backward step. (In the AProS categorization it is a forward, goal-directed step.) Here is the explicit formulation of the second tactic:

$$\begin{array}{|c|}\hline A_1\\ \dots\\ (\forall y : P(y))\,Q(y)\\ \dots\\ A_m\\ \hline T_1\\ \dots\\ Q(x)\\ \dots\\ T_n\\ \hline\end{array} \quad \Longrightarrow \quad \begin{array}{|c|}\hline A_1\\ \dots\\ (\forall y : P(y))\,Q(y)\\ \dots\\ A_m\\ \hline T_1\\ \dots\\ P(x)\\ \dots\\ T_n\\ \hline\end{array}$$

The difference between these two uses of a bounded universal statement lies in the fact that in the first case a suitable antecedent is already available, whereas in the second case a different proof obligation is incurred.

Elimination rules for (bounded) existential statements, ∃-E and ∃:-E, can be easily formulated in the standard natural deduction style. However, these E-rules as well as &-E are not discussed in any detail, as they don't contribute to the write-ups. They are considered only implicitly in the context of the elimination of definitions; we discuss that below.

Now we are going to consider the Backward Steps that are associated with introduction rules applied to goals, i.e., to targets in G&G's terminology.

Backward Steps

$$\begin{array}{|c|}\hline A_1\\ \dots\\ A_m\\ \hline T_1\\ \dots\\ (P\&Q)\\ \dots\\ T_n\\ \hline\end{array} \quad \overset{\&\text{-}\mathtt{I}}{\Longrightarrow} \quad \begin{array}{|c|}\hline A_1\\ \dots\\ A_m\\ \hline T_1\\ \dots\\ P\\ \dots\\ Q\\ \dots\\ T_n\\ \hline\end{array}$$

This is G&G's *splitConjuctiveTarget* [273] that does not contribute to the write-up.

$$\begin{array}{|c|}\hline A_1\\ \dots\\ A_m\\ \hline T_1\\ \dots\\ (P \vee Q)\\ \dots\\ T_n\\ \hline\end{array} \quad \overset{\vee\text{-}\mathtt{I}}{\Longrightarrow} \quad \begin{array}{|c|}\hline A_1\\ \dots\\ A_m\\ \hline T_1\\ \dots\\ \begin{array}{|c|}\hline \\ \hline P\\ \hline\end{array} \qquad\qquad \begin{array}{|c|}\hline \\ \hline Q\\ \hline\end{array}\\ \dots\\ T_n\\ \hline\end{array}$$

This is the *splitDisjunctiveTarget* from (l.c., p. 273). G&G point out that this tactic presents the same kind of problems for their program as the tactic *splitDisjunctiveHypothesis* we discussed.

For universal quantifiers, G&G have two tactics, the *peelBareUniversalTarget* that corresponds directly to the standard ∀-I rule, and the *peelAndSplit-UniversalConditionalTarget*. The latter is just ∀:-I formulated in the box-format as follows:

$$\begin{array}{|c|}\hline A_1\\ \dots\\ A_m\\ \hline T_1\\ \dots\\ (\forall y : P(y))Q(y)\\ \dots\\ T_n\\ \hline\end{array} \quad \Longrightarrow \quad \begin{array}{|c|}\hline A_1\\ \dots\\ A_m\\ \hline T_1\\ \dots\\ \begin{array}{|c|}\hline P(x)\\ \hline Q(x)\\ \hline\end{array}\\ \dots\\ T_n\\ \hline\end{array}$$

Clearly, suitable standard conditions must obtain for the variable x.

There are also tactics for (bounded) existential quantifier namely, *unlockExistentialUniversalConditionalTargets* and *unlockExistentialTargets* (l.c., p. 278). The latter tactic corresponds to ∃-I, but the discussion pertains also to ∃:-I.

Remarks

1. We do not discuss their complex way of "creating Meta-variables" involving diamonds and bullets. What is treated here in Full View is pushed into the background by AProS. AProS uses in its search Skolem functions and Herbrand variables to allow the (generalized) unification of a goal formula with a "canonical instance" of a positive subformula of a premise or assumption; that is discussed in great detail by Sieg and Byrnes in their (Sieg and Byrnes [49], pp. 90–100). In the completed proof these function terms and variables are replaced by ordinary variables that have to satisfy, of course, the side conditions for ∃-`E` and ∀-`I`.
2. As we mentioned already, G&G have no tactics for negation. "Bare" conditionals are not even permitted in the language (see (l.c., p. 269)), because they are "barely of interest of [to?] mathematicians".

The logical steps we just discussed are complemented and, to a large extent, joined with tactics for the elimination, resp. introduction of definitions, what is done by `Def-E` and `Def-I` in our framework. G&G consider two such tactics for the elimination of a definition. The first is their *expandPreexistentialHypothesis* that combines `Def-E` and ∃-`E`, but also ∃:-`E`; it is discussed in (l.c., pp. 275–276). The second is the *elementaryExpansionOfHypothesis*: it combines `Def-E` with the appropriate rule for the form of the definiens, mostly a conjunction. One could, of course, add a rule *expandPreUniversalHpothesis* combining `Def-E` with ∀-`E` and ∀:-`E`. That is treated by G&G under forwards-Reasoning (l.c., p. 274). Finally, the introduction of definitions joins `Def-I` with the immediate logical steps *elementaryExpansionOfTarget*, *expandUniversalTarget*, and *expandPreExistentialTarget*; these tactics are found in (l.c., 277 and 279).

We emphasized in our discussion how crucial the Lemma-as-rule mechanism is both for natural formalization and automated search. That issue is treated by G&G under the heading "Library Reasoning". They state quite correctly:

> Logically speaking, one could unify forwards library reasoning with ordinary forwards reasoning by adding the entire (allowable) content of the library to our list of assumptions. (l.c., p. 275)

The parenthetical "(allowable)" points to a subtle, but extremely important issue we address by the hierarchical organization of the mathematical frame. That kind of conceptual structuring is hinted at, when G&G discuss very briefly their future plan of ordering the library results "using a relation 'is more advanced than', so that for each problem we can simply instruct the library not to use results beyond a certain point in the ordering." There are additional remarks on the use of the library and what should be contained in it as "appropriate background results". They assert:

> It may be that for a more sophisticated program one should not expect to be able to judge the appropriateness of a statement in advance, but should store promising statements and then gradually forget about them if they turn out not to be especially useful. (l.c., p. 263)

AProS has, as we mentioned, the lemma-as-rule mechanism, and we will see in the next section how this mechanism is built into the overall proof search.

4 Two Strategic Approaches

G&G had their prover establish two theorems from "metric space theory"; AProS proves both these theorems quite directly. The first theorem and its proof were presented in Ganesalingam and Gowers [14]; the AProS proof is given in Appendix B of [53]. The second theorem,

> (*) if A and B are open subsets of a metric space X, then their intersection $A \cap B$ is open,

was considered in Ganesalingam and Gowers [15]; that is the article we have been examining. Below we discuss first the strategy that allows the G&G-prover to construct a proof of (*). Next, the strategic approach AProS takes is discussed and used to generate its proof of (*).

The goals of the two approaches are, prima facie, quite different. The G&G-prover aims to give a well-written natural language argument (with access to a library), whereas AProS' goal is to find a well-structured formal proof (within a mathematical frame). Well-written arguments are based, however, on formal steps in the box calculus. The steps are taken strategically, imitating human thought processes; they are also paraphrased directly in natural language. The combination of taking steps individually, and paraphrasing them immediately, forces the search of the G&G-prover to be "deterministic" without any backtracking. Thus, the search is restricted to what G&G call "routine problems" and it is forced to follow a *waterfall strategy*.

Before delving into the details of G&G's strategic approach, it will be useful to analyze their perspective of human theorem proving. After all, that perspective shapes the structure of the automated search carried out by their prover. A search is viewed as successful if it delivers a "good" proof. What then is a good, humanly understandable proof? For mathematicians, a good proof allows them "to achieve understanding rather than to feel confident that a statement is true." (G&G 2017, p. 255) Automatic theorem provers should, consequently, not just provide a "certification of correctness", but they should find proofs that are "explanatory". G&G argue:

> So, an automatic theorem prover that imitates the way humans do mathematics is more likely to produce proofs that are appealing to human mathematicians.
> For a proof to appeal to human mathematicians, a minimum requirement is that it should be written in a way that is similar to the way humans write.

The latter requirement is much easier to satisfy, if the prover actually constructs proofs in a way that "closely mirrors the way human mathematicians operate" and can convert the reasoning process into human-style output. The goal is consequently to design a program "that imitates human thought" in mathematics. Obviously, this goal is not far-removed from the goal Hilbert articulated for proof theory in 1927. What characteristics should a prover have in order to take real steps towards reaching this goal? G&G articulate one *broad characteristic* in their section 1.3, namely, that it operate on a very high level of argumentation. This characteristic is secured in part by (1) taking "mathematical" steps not just "logical" ones, and by (2) delaying the expansion of definitions. The AProS proof search displays these supportive elements as well. The first element corresponds to the inferential mechanism "lemmas-as-rules" relative to a conceptually organized mathematical frame; the second element will be made explicit when we discuss the strategies for AProS.

G&G give two reasons for another extremely important feature of their current prover. That feature is in conflict, as they admit, with mathematical practice: *the search does not involve backtracking*. Their first reason is a determined focus on "routine problems" [l.c., pp. 257–258], i.e., problems "for which humans do not consciously backtrack". As a matter of research strategy, they seem to think that the "routine problems" should be solved before even trying to tackle the general problem of automated proof search. It might be a reasoned pragmatic way of proceeding by taking on the general problem only *in stages*. It is certainly not the only one. Indeed, one might rightly be skeptical when, at the very first stage, a feature is left out that is as deeply integrated into the proof construction process as backtracking indeed is.

Their second reason is, at first sight, more principled. They assert that "...mathematical proofs have a particularly simple rhetorical structure." [l.c., p. 265] Furthermore, the rhetorical structure is additionally constrained by "a strong convention":

> A proof proceeds by the presentation of a sequence of assertions, each of which follows from the premises of the theorem being proved or from previous assertions. This structure is not accidental; it is a direct reflection of the fact that mathematicians process proofs by reading and verifying one sentence at a time, and would not expect the justification of a fact presented in one sentence to be deferred to a later sentence.

This convention gives "an easy way to produce write-ups of our proofs" and motivates an "obvious strategy": each application of a tactic is associated with a number of sentences; these sentences are then concatenated to form a humanly readable text. G&G emphasize:

> Note that this strategy is viable only because we are absolutely rigorous about requiring our tactics to reflect steps in human reasoning; in effect, the strategy is mimicking a human who is carefully writing down a proof while coming up with it, which is quite straightforward for an experienced mathematician.

The write-up process is taken to be *deterministic*; see [l.c., p. 267]. G&G observe explicitly that their arguments for the determinism of this process and their two reasons for "No Backtracking" are far less convincing, when difficult or long proofs are being considered. Their discussion indicates the intricate way in which proof search and text construction are intertwined. Indeed, the need for a deterministic write-up process imposes special constraints on the proof search; that is reflected by their use of the *waterfall architecture* familiar from the Boyer-Moore provers.

Having seen already in Sect. 3 the central tactics of the G&G-prover, we turn our attention now to the *priority ordering* of these tactics in their waterfall list: a tactic is defined as having a higher priority than a second one if it appears above the second one in the waterfall list. In each iteration of the procedure, the G&G-prover chooses the applicable tactic with the highest priority and uses it to map the current box to the "next" one. The table of Fig. 2 lists the waterfall elements in order of their strictly decreasing priority; the table is just a "beautified" version of the G&G-prover's code as presented in (G&G 2017, pp. 270–271). Below the table, G&G re-emphasize the character and use of the list:

> We stress here that because our system is fully automatic and intended to model human thought processes, our efforts have been concentrated less on the tactics themselves and more on how the program chooses which tactic to apply. For this program, the general form of the answer is as follows: it just chooses a tactic of the first type it can apply from the list above. (l.c., p. 271)

We described the tactics already in Sect. 3 and further explain their function in Appendix 2. There, we also show how they can be mimicked with NI calculus rules. G&G give (l.c.,

- Deletion
  - deleteDone
  - deleteDoneDisjunct
  - deleteDangling
  - deleteUnmatchable
- Tidying
  - peelAndSplitUniversalConditionalTarget
  - SplitDisjunctiveHypothesis
  - SplitConjuctiveTarget
  - SplitDisjunctiveTarget
  - peelBareUniversalTarget
  - removeTarget
  - collapseSubtableauTarget
- Applying
  - Forwards Reasoning
    - forwardsLibraryReasoning
    - expandPreExistentialHypothesis
    - elementaryExpansionOfHypothesis
  - Backwards Reasoning
    - backwardsLibraryReasoning
    - elementaryExpansionOfTarget
    - expandPreUniversalTarget
    - solveBullets
    - automaticRewrite
- Suspension
  - unlockExistentialUniversalConditionalTarget
  - unlockExistentialTarget
  - expandPreExistentialTarget
  - convertDiamondToBullet
- Equality Substitution
  - rewriteVriabelVariableEquality
  - rewriteVariableTermEquality

Fig. 2 Waterfall of tactics

pp. 280–282) a "justification for the order of priority". Their justification has two main components. First, it has a quasi-empirical grounding as they have informally worked through a large number of problems. They describe, how they arrived at the order of priority reflected in Fig. 2:

> After a while we were in a position to make a first guess at a suitable method for choosing tactics. We then tried the method out on more problems, adjusting it when it led to inappropriate choices. After several iterations of this, we arrived at the order of priority of the tactics that we set out in the previous section. (l.c., p. 280)

Secondly, the justification appeals to an "informal guiding principle" stating that " the program prefers 'safe' tactics to 'dangerous' tactics". However, when a tactic is to be viewed as safe is left rather vague:

> Broadly speaking, a tactic is safe if the risk that it will lead to an undesirable result, such as a dead end or a step that is completely irrelevant to the eventual proof, is small.(l.c., p. 280)

They consider tidying tactics as safe, because these tactics just express "the goal in a more convenient form" and provide "new options without closing off any old ones".(l.c., p. 281) That seems to be correct for the tactics dealing with (bounded) universal quantifiers and conjunction. As we saw in Sect. 3, they judge the treatment of disjunctions as "work in progress". It seems that neither the *splitDinjunctiveHypothesis* tactic nor the *splitDisjunctiveTarget* tactic is fully integrable into their (restricted deterministic) strategic approach, as both require backtracking.

The discussion of the applying tactics reveals another problematic aspect of G&G's strategic ordering: forwards reasoning has always a higher priority than backwards reasoning, in spite of the fact that forwards reasoning is judged to be "in general fairly unsafe". That ordering is then "supported" by reference to the highly routine problems the G&G-prover can tackle: for these problems forwards reasoning "tend to be safe". Thus, in general, the G&G-prover does not capture one crucial aspect of human theorem proving, namely, being able to switch directions in argumentation. That lack is described as follows:

> This aspect of our program is, however, unstable, for the reason just given. When humans are faced with several possibilities for forwards reasoning, they will often switch to backwards reasoning in order to lessen the risk of making irrelevant deductions, but our program does not yet have any facility for making this kind of judgment.(l.c., p. 281)

We come back to this issue when describing AProS' strategic approach. Now, let us see how the G&G-prover establishes the theorem that $\mathsf{A} \cap \mathsf{B}$ is open, in case A and B are open subsets of a metric space X. The detailed construction of the sequence of boxes that lead to the proof displayed in Fig. 3 can be obtained from the justification of each line. We note that, surprisingly, the write-up for this proof is not given by G&G. The G&G-prover's library contains the definition of a metric space (X, d) but also the definition of when a subset of X is open as well as the definition of set theoretic intersection.

Definition 1 (Open Set) $\mathsf{Open(A)} \leftrightarrow (\forall x \in \mathsf{A})(\exists \delta > 0)\ (\forall u)\ (\mathsf{d}(x, u) < \delta \rightarrow u \in \mathsf{A})$

Definition 2 (Intersection) $x \in \mathsf{A} \cap \mathsf{B} \leftrightarrow (x \in \mathsf{A} \mathbin{\&} x \in \mathsf{B})$

The formal statement of the theorem can be written as:

Theorem 1 (Intersection of Open Sets is Open)

$$(\mathsf{METS(X,d)} \mathbin{\&} \mathsf{A} \subseteq \mathsf{X} \mathbin{\&} \mathsf{B} \subseteq \mathsf{X}) \rightarrow ((\mathsf{Open(A)} \mathbin{\&} \mathsf{Open\ (B)}) \rightarrow \mathsf{Open(A} \cap \mathsf{B}))$$

The initial box has $\mathsf{Open(A)}$ and $\mathsf{Open(B)}$ as assumptions and $\mathsf{Open(A} \cap \mathsf{B})$ as the target. The complete proof of G&G is given in Fig. 3. In the proof, they use two auxiliary lemmas from the library:

Lemma 1 *If* $\mathsf{a} < \mathsf{b}$ *and* $\mathsf{b} \leq \mathsf{c}$, *then* $\mathsf{a} < \mathsf{c}$.

Lemma 2 $\min\mathsf{(a,b)} \leq \mathsf{a}$ *&* $\min\mathsf{(a,b)} \leq \mathsf{b}$

The considerations underlying this proof are all based on "mathematical experience", but it takes for granted the formulation of statements in the language of first-order logic. The

1.	$\mathsf{Open}(\mathsf{A})$	Premise
2.	$\mathsf{Open}(\mathsf{B})$	Premise
3.	$x \in \mathsf{A} \cap \mathsf{B}$	Hyp
4.	$x \in \mathsf{A}$	$\mathtt{elemExpOfHyp} : 1$
5.	$x \in \mathsf{B}$	$\mathtt{elemExpOfHyp} : 2$
6.	$(\forall u)(\mathsf{d}(x,u) < \eta_1[x] \to u \in \mathsf{A})$	$\mathtt{forwardsReasoning}, \exists\mathtt{E} : 1, 4$
7.	$(\forall u)(\mathsf{d}(x,u) < \eta_2[x] \to u \in \mathsf{B})$	$\mathtt{forwardsReasoning}, \exists\mathtt{E} : 2, 5$
8.	$(\mathsf{d}(x,u) < \delta^{\bullet}$	Hyp
9.	$\delta^{\bullet} \leq \eta_1[x]$	$\mathtt{Construction} - \mathtt{Min}^*$
10.	$\delta^{\bullet} \leq \eta_2[x]$	$\mathtt{Construction} - \mathtt{Min}^*$
11.	$\mathsf{d}(x,u) < \eta_1[x]$	$\mathtt{backwardsLibraryReasoning} : 8, 9$
12.	$\mathsf{d}(x,u) < \eta_2[x]$	$\mathtt{backwardsLibraryReasoning} : 8, 10$
13.	$u \in \mathsf{A}$	$\mathtt{backwardsReasoning} : 6, 11$
14.	$u \in \mathsf{B}$	$\mathtt{backwardsReasoning} : 7, 12$
15.	$u \in \mathsf{A} \cap \mathsf{B}$	$\mathtt{elemExpOfTarget} : 13, 14$
16.	$(\exists \delta)\ (\forall u)\ (\mathsf{d}(x,u) < \delta \to u \in \mathsf{A} \cap \mathsf{B})$	$\mathtt{unlock}\,\exists\forall \to \mathtt{Targ} : 8 - 15$
17.	$(\forall x)(x \in \mathsf{A} \cap \mathsf{B} \to (\exists \delta)\ (\forall u)\ (\mathsf{d}(x,u) < \delta \to u \in \mathsf{A} \cap \mathsf{B})$	$\mathtt{pS}\,\forall \to \mathtt{Targ} : 3 - 16$
18.	$\mathsf{Open}(\mathsf{A} \cap \mathsf{B})$	$\mathtt{expandPreUniversalTarget} : 17$

Fig. 3 G&G proof for Theorem 1

proof can be given a systematic logical foundation by taking into account the "mimicking" of G&G's mathematical steps by the strategic use of NIC rules in a mathematical frame. After all, the tactical steps of the G&G-prover are either basic logical inference steps or "abbreviations" of proofs that use logical and definitional rules, as well as lemmas from the library. These steps can all be taken in the NI calculus, indeed, in a fragment of its minimal version; one has only to take for granted that the same definitions and lemmas are available. Thus, using AProS as a proof assistant, any proof construction carried out by the G&G-prover can be duplicated in the AProS system. However, more importantly and more interestingly, AProS—without modifying its proof search strategies in any way—proves the theorem fully automatically and finds a very natural and direct proof. Let us turn then to the discussion of the AProS strategies and of the construction of the direct proof of the theorem, i.e., the bi-directional strategic approach.

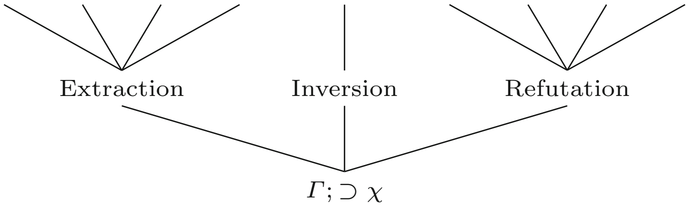
The AProS search procedure—with backtracking and discriminatory evaluations when alternatives have to be considered—reveals the second face of proof theory that is directed towards mathematical arguments: it is *strategically guided* to obtain formal proofs that reflect the structure of informal proofs. We indicated in Sect. 2 that *bi-directionality* is a crucial feature of the construction of proofs; indeed, it is directly incorporated in the NI calculus. The forward moves by elimination rules, and the backward moves by inverted introduction rules are always *goal-directed*. Here are crucial steps that are taken in the indicated order:
EXTRACTION, when the goal has to be a positive subformula of a premise or an assumption;
INVERSION, that includes $\neg\mathtt{I}$ and $\bot\mathtt{I}$;
REFUTATION, that takes two forms, $\bot\mathtt{E}$ and $\neg\mathtt{E}$.

Remarks

1. Disjunctions and existentially quantified formulas are treated in special ways that are reflected in the strategic moves CONVERSION and DIVISION.

2. The finitely branching tree that uses *all* intercalating strategic moves can be exploited to prove the completeness of the NI calculus for both intuitionist and classical first-order logic. The proof is similar to the proof that establishes the completeness of cut-free sequent calculi.
3. NIC proofs can easily be translated into **normal** natural deduction proofs. So we have via (2) direct completeness proofs for the system consisting of just **normal proofs**.
4. Conversely, some structural properties of normal derivations inform the proof search in the NI calculus: for example, extraction parallels the division of branches of normal proofs into E- and I-parts; see ([43] p. 41 and p. 53).

The above considerations are statically reflected in this *basic* **search tree**:



This basic search strategy, refined by conversion and division, is systematically expanded through E- and I-rules for definitions; see Sect. 2.[28]

As the logical rules express the meaning of the logical connectives, the meaning of the defined mathematical notions is captured by their E- and I-rules. The crucial step that goes beyond the purely meaning-based organization of the search integrates it with the **mathematical frame**. That frame is constructed to capture genuinely mathematical insights. The integration is achieved through lemmas that are used in rule-like *forward* and *backward* moves, i.e., the mechanism of lemmas-as-rules; see end of Sect. 2 and beginning of Sect. 3. (That can be seen in effective action in Sieg and Walsh [53].)

The issue of Lemmas-as-rules is treated by G&G under the heading of "Library Reasoning" and was discussed by us above.[29] G&G consider this as work in progress; indeed, their extremely preliminary considerations provide no insight of how lemmas can be integrated systematically with the search. For us, the discovery of an "effective" mathematical frame is best pursued in the quasi-empirical testbed of natural formalization: through the computer implementation different frames can be used and their efficacy for the mathematical development can be explored. Let us come back to the theorem of metric space theory and their proofs found by the G&G-prover, respectively by AProS. Although there are striking differences between the two search procedures, they are "similar" for routine problems. We presented the G&G proof of our theorem above; now, we will discuss how AProS finds it. The proof generated by AProS is given in Fig. 4.

The initial step is to use inversion on the goal, using Def I backward and then applying ∀:-I backward. The next step in AProS is to apply Conversion on Line 1. Strategically we can

[28] Much more could be said about the use of these rules through **deep extraction** using an extended **subformula property** when the definiens is viewed as the "subformula" of the definiendum. If a disjunction or an existentially quantified is discovered as a positive subformula, conversion or division might be triggered.

[29] In [3], the role of the Mizar mathematical library in interactive theorem proving is described in great detail and with a good deal of overlap. The general spirit we expressed for the mathematical frame is expressed then for the *Mizar mathematical library* (on p. 23), namely, that it "was considered an experiment of practical formal modeling of mathematics."

1.	Open(A)	Premise
2.	Open(B)	Premise
3.	$(\forall x \in \mathsf{A})(\exists \delta > 0)\ (\forall u)\ (\mathsf{d}(x,u) < \delta \to u \in \mathsf{A})$	def E : 1
4.	$(\forall x \in \mathsf{B})(\exists \delta > 0)\ (\forall u)\ (\mathsf{d}(x,u) < \delta \to u \in \mathsf{B})$	def E : 2
5.	$x \in \mathsf{A} \cap \mathsf{B}$	Hyp
6.	$x \in \mathsf{A}\ \&\ x \in \mathsf{B}$	def E : 3
7.	$x \in \mathsf{A}$	& EL : 6
8.	$x \in \mathsf{B}$	& ER : 6
9.	$(\exists \delta > 0)\ (\forall u)\ (\mathsf{d}(x,u) < \delta \to u \in \mathsf{A})$	$\forall \in$ E : 3, 7
10.	$\eta_1 > 0\ \&\ (\forall u)\ (\mathsf{d}(x,u) < \eta_1 \to u \in \mathsf{A})$	Hyp
11.	$(\exists \delta > 0)\ (\forall u)\ (\mathsf{d}(x,u) < \delta \to u \in \mathsf{B})$	$\forall$E : 3, 8
12.	$\eta_2 > 0\ \&\ (\forall u)\ (\mathsf{d}(x,u) < \eta_2 \to u \in \mathsf{B})$	Hyp
13.	$\min(\eta_1, \eta_2) > 0\ \&\ (\forall u)\ (\mathsf{d}(x,u) < \min(\eta_1, \eta_2) \to u \in \mathsf{A} \cap \mathsf{B})$	MH : 10, 12
14.	$(\exists \delta > 0)\ (\forall u)\ (\mathsf{d}(x,u) < \delta \to u \in \mathsf{A} \cap \mathsf{B})$	$\exists$:I : 13
15.	$(\exists \delta > 0)\ (\forall u)\ (\mathsf{d}(x,u) < \delta \to u \in \mathsf{A} \cap \mathsf{B})$	$\exists$E : 11, 14
16.	$(\exists \delta > 0)\ (\forall u)\ (\mathsf{d}(x,u) < \delta \to u \in \mathsf{A} \cap \mathsf{B})$	$\exists$E : 9, 15
17.	$(\forall x \in \mathsf{A} \cap \mathsf{B})(\exists \delta > 0)\ (\forall u)\ (\mathsf{d}(x,u) < \delta \to u \in \mathsf{A} \cap \mathsf{B})$	$\forall \in$ I : 5, 16
18.	Open($\mathsf{A} \cap \mathsf{B}$)	def I : 17

Fig. 4 AProS Proof for Theorem 1

do this because the definition of Open(A) has an existentially quantified sentence as a positive subformula. So, the rules def E, ∀E, and ∃E are applied to Lines 1, 3, and 9 respectively. For the same reason, these rules are also applied to Lines 2, 4, and 11. While applying ∃E, the variable δ on Lines 9 and 11 are instantiated by η_1 and η_2, respectively. The next step is to apply ∃I, as an inversion step leading to line 13. Here, the variable δ is instantiated by the term $\min(\eta_1, \eta_2)$. By this choice the proof can be completed in just one step using the Minimum Heuristic Lemma (MH).

In the MH Lemma, Fig. 5, we prove $\min(\eta_1, \eta_2) > 0\ \&\ (\forall u)\ (\mathsf{d}(x, u) < \min(\eta_1, \eta_2) \to u \in \mathsf{A} \cap \mathsf{B})$ holds, having the premises $\eta_1 > 0\,\&\,(\forall u)\ (\mathsf{d}(x, u) < \eta_1 \to u \in \mathsf{A})$ and $\eta_2 > 0\,\&\,(\forall u)\ (\mathsf{d}(x, u) < \eta_2 \to u \in \mathsf{B})$. We use the following lemmas to prove the MH Lemma in AProS:

Lemma 3 if a $<$ min(b,c) then a $<$ b.

Lemma 4 if a $<$ min(b,c) then a $<$ c.

Lemma 5 if a $>$ 0 & b $>$ 0, then min (a,b) $>$ 0

The proof, depicted in Fig.5, starts by applying &I to Line 19, as an inversion step since neither extraction nor conversion are applicable. We continue by applying several inversions to Line 18. To prove Lines 10 and 14, we apply extraction to Lines 7 and 11, respectively, since the goals are positive subformulae of (the formulae on) those lines. To prove Lines 9 and 13, the only option is applying Lemmas 3 and 4 to Line 6. The rest of the proof in AProS, i.e., to complete the proof of Line 5, is straightforward; one appeals to Lemma 5.

1.	$\eta_1 > 0 \,\&\, (\forall u)\ (d(x,u) < \eta_1 \rightarrow u \in \mathsf{A})$	Premise
2.	$\eta_2 > 0 \,\&\, (\forall u)\ (d(x,u) < \eta_2 \rightarrow u \in \mathsf{B})$	Premise
3.	$\eta_1 > 0$	$\&\,\mathtt{EL} : 1$
4.	$\eta_2 > 0$	$\&\,\mathtt{EL} : 2$
5.	$\min(\eta_1, \eta_2) > 0$	$\mathtt{Lemma5} : 3, 4$
6.	$\mathsf{d}(x,u) < \min(\eta_1, \eta_2)$	Hyp
7.	$(\forall u)(\mathsf{d}(x,u) < \eta_1 \rightarrow u \in \mathsf{A})$	$\&\,\mathtt{ER} : 1$
8.	$\mathsf{d}(x,u) < \eta_1 \rightarrow u \in \mathsf{A}$	$\forall\mathtt{E} : 7$
9.	$\mathsf{d}(x,u) < \eta_1$	$\mathtt{Lemma3} : 6$
10.	$u \in \mathsf{A}$	$\rightarrow\mathtt{E} : 8, 9$
11.	$(\forall u)(\mathsf{d}(x,u) < \eta_2 \rightarrow u \in \mathsf{B})$	$\&\mathtt{ER}$
12.	$\mathsf{d}(x,u) < \eta_2 \rightarrow u \in \mathsf{B}$	$\forall\mathtt{E} : 11$
13.	$\mathsf{d}(x,u) < \eta_2$	$\mathtt{Lemma4} : 6$
14.	$u \in \mathsf{B}$	$\rightarrow\mathtt{E} : 12, 13$
15.	$(u \in \mathsf{A} \,\&\, u \in \mathsf{B})$	$\&\mathtt{E} : 10, 14$
16.	$u \in \mathsf{A} \cap \mathsf{B}$	$\mathtt{def\,I} : 15$
17.	$\mathsf{d}(x,u) < \min(\eta_1, \eta_2) \rightarrow u \in \mathsf{A} \cap \mathsf{B}$	$\rightarrow\mathtt{I} : 6 - 16$
18.	$(\forall u)\ (\mathsf{d}(x,u) < \min(\eta_1, \eta_2) \rightarrow u \in \mathsf{A} \cap \mathsf{B})$	$\forall\mathtt{I} : 17$
19.	$\min(\eta_1, \eta_2) > 0\ \&\ (\forall u)\ (\mathsf{d}(x,u) < \min(\eta_1, \eta_2) \rightarrow u \in \mathsf{A} \cap \mathsf{B})$	$\&\,\mathtt{I} : 5, 18$

Fig. 5 AProS proof for Minimum Heuristic Lemma

5 Programmatic Remarks

Both the G&G-prover and AProS aim to reflect mathematical practice. There is, as we have seen, a quite dramatic difference in emphasis as well as in strategic direction: the G&G-prover is to output a write-up that is not distinguishable from good mathematical proof writing; AProS is to construct formal proofs that are humanly intelligible. To reach their goal, G&G are not using any backtracking and, thus, are forced to only consider-what they call-*routine problems* whose proofs can be constructed "deterministically". As matters stand, G&G's tactics are extremely restricted: they do not involve (bare) conditionals or negations; the treatment of disjunctions is described as "work in progress"; even conjunctions are treated in a perfunctory way, as backward &`I` does not contribute to the write-up and forward &`E` is used only within the tactic *elementaryExpansionOfHypothesis*. So, we are really left with argumentation that involves ∀:`E`, ∃:`E`, ∀`I`, ∃`I`, ∀:`I`, and ∃:`I`. The standard rules ∀`E` and ∃`E` are not used at all. These facts allow AProS to mimic directly the G&G-prover's proof constructions in a small fragment of minimal logic. It would not be difficult to have AProS follow the waterfall strategy of the G&G-prover and thus obtain formal proofs in our framework that reflect the logical structures underlying the write-ups generated through the G&G-prover.

As we emphasized throughout, the AProS project pursues a different strategic direction, namely, it aims to create an efficient mechanism that constructs humanly intelligible formal proofs. It does so in a natural bi-directional way and with logically grounded strategic guidance. At first, this goal was pursued for full classical and intuitionist first-order logic. More recently, it was extended to elementary set theory within a conceptual frame that incorporates the central feature of any systematic axiomatic development of parts of mathematics. It is a topic of current interest, how far the interactively obtained proof of the Cantor–Bernstein theorem can be obtained automatically.

Another project, stimulated by G&G's goal of obtaining excellent write-ups, can actually be pursued in two stages. In stage one, the standard AProS procedure would be at work to find humanly intelligible formal proofs. The completed proof can be, in stage 2, systematically translated into English (or other natural languages). Though the proof search of AProS uses backtracking, the construction steps involved in the completed proof are deterministically organized and can be used as the fixed underpinning for the translation. Note that the line-numbering of an AProS proof does not reflect the order in which the lines were strategically obtained. Indeed, the strategic guidance allows *chunking* the output of the translation procedure and articulating the strategic significance of the next sequence of steps, for example, that a "proof by contradiction" or an "arguments by cases" is to be given. In the first example, the translation could say: "We are proceeding now indirectly to prove $\neg\phi$, by assuming ϕ and establishing the contradiction $\neg\chi$ and χ". In the second example, when a disjunction $(\phi \vee \psi)$ is used to establish a goal χ, the translation could say: "The disjunction $(\phi \vee \psi)$ is available and we give an argument by cases, that means, we establish χ first under the assumption ϕ (case 1) and then under the assumption ψ (case 2)."

The separation of searching for a proof from obtaining an excellent write-up is absolutely crucial. It is this separation that allows us to use backtracking in the search and the deterministic underpinning of the formal proofs for a translation that is strategically informed. Clearly, this makes sense only when the automated proof search is human-centered and yields humanly intelligible formal proofs. It is equally clear that a translation in the spirit of G&G aiming to be indistinguishable from excellent mathematical writing will require a substantial amount of detailed, mostly linguistic work. It is no accident that G&G do not give a translation of their proof of the elementary theorem from metric space theory.

Finally, we come to the educational issues that work on the G&G-prover, respectively, AProS was to address. In his interview [11], Gowers raised the question: How can we train mathematicians to write better proofs? What better method, he suggested, than to train a computer to find and write proofs? If that could be done, then we might be able to use that method also in mathematics education. This goal underlies the extensive discussion in Sects. 1 and 2 of (G&G, 2017) that leads to the constraints a prover should satisfy and that are formulated in Sect. 3 entitled, *Key Features of the Prover*. G&G emphasize (l.c., p. 261) that a "goal-oriented tactic-based style of proof" should be pursued using "essentially first-order logic with equality" as the logical framework.

Sieg, in collaboration with many colleagues and students, has been pursuing since the late 1980s the goal of teaching (undergraduate) students how to construct proofs in a natural deduction calculus -with the guidance of well-grounded strategies. These strategies were, of course, to be human ones. To accommodate in one syntactic configuration the forward and backward moves of informal reasoning, the NI calculi were formulated in Sieg [45] and then in Sieg and Byrnes [49]. The learning of *strategic thinking* was to be supported by a *dynamic Proof Tutor*. This Tutor was to enter, on student demand, into a dialogue and not only help students to clarify the strategic situation, but also to provide genuine advice. That advice was to be based on the completion of the student's partial proof by AProS. That was the vision for a hopefully more effective education in logic. AProS was built over many years, and the introductory course *Logic&Proofs* was designed and implemented as a fully web-based course; only a couple of years ago did we realize this vision by integrating *Logic&Proofs* and *AProS* via the Proof Tutor.[30]

We mentioned at the end of Sect. 2 AProS's expansion to directly find proofs of Gödel's Incompleteness Theorems relative to the local axiomatic assumptions of representability and derivability conditions. The considerations underlying the automated search were described in Sieg and Field [51]. However, we expanded the interface of *Logic&Proofs* to a genuine *ProofLab* only during the last few years. Now we can work interactively with AProS as a *proof assistant* and construct proofs in meta-mathematics, but also in set theory and modern algebra. That allows us to illuminate the second face of proof theory and bring this aspect of the subject to radiant life by a thorough analysis of mathematical proofs, their presentation in suitable formalisms and expanding human-centered heuristics for finding explanatory proofs fully automatically. In this way we will not only uncover remarkable capacities of the mathematical mind, but more broadly of the human mind. After all, mathematics is common sense made rigorous.

Appendix 1

Here are the rules of the NI calculus. We present first the elimination rules; they are to be read "upward" as one is taking a "gap-closing" or intercalating step. This continues the description of the NI calculus from toward the end of Sect. 1; in particular, we do not repeat the role of extraction sequences for the elimination rules.

$$\frac{\Gamma;\Delta,(\phi\,\&\,\psi),\phi\supset\chi}{\Gamma;\Delta,(\phi\,\&\,\psi)\supset\chi}\;\&E_1(\sigma)\qquad\qquad\frac{\Gamma;\Delta,(\phi\,\&\,\psi),\psi\supset\chi}{\Gamma;\Delta,(\phi\,\&\,\psi)\supset\chi}\;\&E_2(\sigma)$$

$$\frac{\Gamma,\phi;\supset\chi\quad\Gamma,\psi;\supset\chi}{\Gamma;\Delta,(\phi\vee\psi)\supset\chi}\;\vee E(\sigma[\vee])\qquad\qquad\frac{\Gamma;\Delta,(\phi\rightarrow\psi),\psi\supset\chi\quad\Gamma;\supset\phi}{\Gamma;\Delta,(\phi\rightarrow\psi)\supset\chi}\;\rightarrow E(\sigma[\rightarrow])$$

$$\frac{\Gamma;\Delta,(\forall\alpha)\psi,\psi(\tau)\supset\chi}{\Gamma;\Delta,(\forall\alpha)\psi\supset\chi}\;\forall E(\sigma)\qquad\qquad\frac{\Gamma,\psi(\beta);\supset\chi}{\Gamma;\Delta,(\exists\alpha)\psi\supset\chi}\;\exists E(\sigma[\exists])$$

The I-rules for these connectives are the usual ones; see (Gentzen [17], p. 186) or (Prawitz [43], p. 20). The rules for $\bot$ and negation are formulated next.

$$\frac{\Gamma;\supset\phi\quad\Gamma;\supset\neg\phi}{\Gamma;\supset\bot}\;\bot I\quad \textit{where } \neg\phi \textit{ is a positive subformula of an element of } \Gamma$$

$$\frac{\Gamma;\supset\bot}{\Gamma;\supset\phi}\;\bot E\qquad\qquad\frac{\Gamma,\phi;\supset\bot}{\Gamma;\supset\neg\phi}\;\neg I\qquad\qquad\frac{\Gamma,\neg\phi;\supset\bot}{\Gamma;\supset\phi}\;\neg E$$

This is, without $\neg E$, "exactly" the formulation for intuitionist logic given in (Gentzen [17], p. 186). Unfortunately, Gentzen called $\bot I$, "Negationsbeseitigung" and expanded this cal-

[30] *Logic&Proofs* has been offered as a fully web-based course, in various stages of development, since 2007. It has been successfully completed by more than 12000 students at many different institutions. For more information, see AProS Site (http://www.phil.cmu.edu/projects/apros/) and Logic&Proofs Course(https://oli.cmu.edu/courses/logic-proofs-copy/).

culus to its classical form by adding, for example, the law of the excluded middle $\phi \vee \neg\phi$ or the rule of double negation elimination. Dropping $\perp E$ leads to minimal logic.

Appendix 2

We discussed most of the tactics used by the G&G-prover already in Sect. 3. Let us look here at a few more examples on how tactics can be mimicked by rules of the NI calculus, i.e., can be viewed as derived rules.
A Tidying Tactic: peelAndSplitUniversalonditionalTarget

$$\cfrac{\cfrac{\Gamma, \Phi(z); \supset \Psi(z)}{\Gamma; \supset \Phi(z) \to \Psi(z)} \to I}{\Gamma; \supset (\forall x)(\Phi(x) \to \Psi(x))} \forall I$$

An Applying Tactic:

- *expandPreExistentialHypothesis.*
  Let χ be by definition $(\exists x)\phi(x)$;

$$\cfrac{\cfrac{\Gamma, \chi, \phi(z); \supset \psi}{\Gamma, \chi; (\exists x)\phi(x) \supset \psi} \exists E}{\Gamma, \chi; \supset \psi} Def\, E$$

- *backwardsLibraryReasoning.*
  Let $(\forall u)(P(u) \to Q(u))$ be a lemma; this is reflected by backward reasoning in our setting using the lemma as a rule:

$$\frac{\Gamma; \supset P(x)}{\Gamma; \supset Q(x)}$$

References

1. Andrews, P.B.: Transforming matings into natural deduction proofs. In: Bibel, W., Kowalski, R., (eds.) 5-th Conference on Automated Deduction, Lecture Notes in Computer Science, pp. 281–292. Springer (1980)
2. Andrews, P.B.: More on the problem of finding a mapping between clause representation and natural-deduction representation. J. Autom. Reason. **7**(2), 285–286 (1991)
3. Bancerek, G., Byliński, C., Grabowski, A., Korniłowicz, A., Matuszewski, R., Naumowicz, A., Pak, K.: The role of the Mizar mathematical library for interactive proof development in Mizar. J. Autom. Reason. **61**(1–4), 9–32 (2018)
4. Bernays, P.: Beiträge zur axiomatischen Behandlung des Logik-Kalküls. Georg-August-Universität Göttingen, Habilitation (1918)
5. Bledsoe, W.: Non-resolution theorem proving. Artif. Intell. **9**(1), 1–35 (1977)
6. Bledsoe, W.: The UT natural deduction prover. Technical Report, Department of Computer Science, University of Texas at Austin (1983)
7. Byrne, O.: The First Six Books of the Elements of Euclid. Pickering, London (1847)
8. Byrnes, J.: Proof search and normal forms in natural deduction. PhD thesis, Carnegie Mellon University (1999)
9. Church, A.: An unsolvable problem of elementary number theory. Am. J. Math. **58**(2), 345–363 (1936)
10. Cittadini, S.: Intercalation calculus for intuitionistic propositional logic. Technical Report PHIL-29, Philosophy, Methodology, and Logic, Carnegie Mellon University (1992)
11. Diaz Lopez, A.: Sir Timothy Gowers interview. Not. Am. Math. Soc. **63**(9), 1026–1028 (2016)

12. Fitch, F.B.: Symbolic Logic: An Introduction. Ronald Press, New York (1952)
13. Fitting, M.: First-Order Logic and Automated Theorem Proving. Springer, Berlin (1990)
14. Ganesalingam, M., Gowers, W.T.: A fully automatic problem solver with human-style output. arXiv:1309.4501 (2013)
15. Ganesalingam, M., Gowers, W.T.: A fully automatic theorem prover with human-style output. J. Autom. Reason. **58**(2), 253–291 (2017)
16. Gentzen, G.: Urdissertation, Wissenschaftshistorische Sammlung, Eidgenössische Technische Hochschule. Zürich, Bernays Nachlass, Ms ULS (A detailed description of the manuscript is found in (von Plato 2009) (1932-3)
17. Gentzen, G.: Untersuchungen über das logische Schließen I, II. Mathematische Zeitschrift 39(1):176–210, 405–431 (1934-5)
18. Gentzen, G.: Die Widerspruchsfreiheit der reinen Zahlentheorie. Mathematische Annalen **112**(1), 493–565 (1936)
19. Gentzen, G.: Neue Fassung des Widerspruchsfreiheitsbeweises für die reine Zahlentheorie. Forschung zur Logik und zur Grundlegung der exakten Wissenschaften. Neue Folge 4, S Hirzel:19–44 (1938)
20. Harrison, J., Urban, J., Wiedijk, F.: History of interactive theorem proving. Comput. Logic **9**, 133–214 (2014)
21. Hazen, A.P., Pelletier, F.J.: Gentzen and Jaśkowski on natural deduction: Fundamentally similar but importantly different. Stud. Logica. **102**(6), 1103–1142 (2014)
22. Heule, M.J., Kullmann, O.: The science of brute force. Commun. ACM **60**(8), 70–79 (2017)
23. Hilbert, D.: Grundlagen der Geometrie. Teubner, Stuttgart (1899)
24. Hilbert, D.: Axiomatisches Denken. Mathematische Annalen 78:405–415, this paper is translated in: From Kant to Hilbert, W.B. Ewald, editor, Oxford University Press 1996, 1107–1115 (1918)
25. Hilbert, D.: Die Grundlagen der Mathematik. Abhandlungen aus dem mathematischen Seminar der Hamburgischen Universität (6):65–85 (1927)
26. Hilbert, D., Bernays, P.: Prinzipien der Mathematik. In: Ewald, W.B., Sieg, W. (eds.) David Hilbert's Lectures on the Foundations of Arithmetic and Logic 1917–1933, pp. 64–214 . Springer (2013) (1917–1918)
27. Hilbert, D., Bernays, P.: Grundlagen der Mathematik. In: Ewald, W.B., Sieg, W. (eds) David Hilbert's Lectures on the Foundations of Arithmetic and Logic 1917–1933, pp. 431–518. Springer (2013) (1921–1922)
28. Hilbert, D., Bernays, P.: Grundlagen der Mathematik, vol. I. Springer, Berlin (1934)
29. Hilbert, D., Bernays, P.: Grundlagen der Mathematik, vol. II. Springer, Berlin (1939)
30. Jaśkowski, S.: On the rules of suppositions in formal logic. Stud. Logica. **1**, 5–32 (1934)
31. Mac Lane, S.: Abgekürzte Beweise im Logikkalkul [sic]. Dissertation, Georg-August-Universität Göttingen (1934)
32. Mac Lane, S.: A logical analysis of mathematical structure. The Monist **45**(1), 118–130 (1935)
33. Mac Lane, S.: A late return to a thesis in logic. In: Kaplansky, I. (ed) Saunders Mac Lane-Selected Papers, pp. 63–66. Springer (1979)
34. Mac Lane, S.: Saunders Mac Lane: A Mathematical Autobiography. A.K, Peters (2005)
35. Mac Kenzie, D.: The automation of proof: a historical and sociological exploration. IEEE Ann. Hist. Comput. **17**(3), 7–29 (1995)
36. Newell, A., Simon, H.A.: The logic theory machine a complex information processing system. Technical report, (No .P-868) The Rand Corporation (1956)
37. Pelletier, F.J.: Automated natural deduction in THINKER. Stud. Logica **60**(1), 3–43 (1998)
38. Pelletier, F.J.: A brief history of natural deduction. Hist. Philos. Logic **20**(1), 1–31 (1999)
39. Pelletier, F.J., Hazen, A.P.: A history of natural deduction. In: Gabbay, D., Pelletier, F.J., Woods, J. (eds.) Handbook of the History of Logic, vol. 11, pp. 341–414. Elsevier, London (2012)
40. Pfenning, F.: Proof transformations in higher-order logic. PhD thesis, Carnegie Mellon University (1987)
41. Pfenning, F., Sieg, W. (eds): Natural deduction. Special Issue of Studia Logica 60 (1998)
42. Pollock, J.: Skolemization and unification in natural deduction. unpublished paper available at http://www.uarizonaedu/~pollock (1997)
43. Prawitz, D.: Natural Deduction—A Proof-Theoretical Study. Almqvist & Wiksell, Stockholm (1965)
44. Rips, L.J.: The Psychology of Proof—Deductive Reasoning in Human Thinking. MIT Press, London (1994)
45. Sieg, W.: Mechanisms and Search—Aspects of Proof Theory, vol. 14. Associazione Italiana di Logica e sue Applicazioni (1992)
46. Sieg, W.: The AProS project: Strategic thinking and computational logic. Logic J. IGPL **15**(4), 359–368 (2007)
47. Sieg, W.: Review of (von Plato 2008). Mathematical Reviews 2413304 (2009)

48. Sieg, W.: Searching for proofs (and uncovering capacities of the mathematical mind). In: Feferman, S., Sieg, W. (eds.) Proof, Categories and Computations, pp. 189–215. College Publications, New York (2010)
49. Sieg, W., Byrnes, J.: Normal natural deduction proofs (in classical logic). Stud. Logica **60**(1), 67–106 (1998)
50. Sieg, W., Cittadini, S.: Normal natural deduction proofs (in non-classical logics). In: Mechanizing Mathematical Reasoning. Lecture Notes in Computer Science 2605, pp. 169–191. Springer (2005)
51. Sieg, W., Field, C.: Automated search for Gödel's proofs. Ann. Pure Appl. Logic **133**, 319–338 (2005)
52. Sieg, W., Scheines, R.: Searching for proofs (in sentential logic). In: Burkholder, L. (ed.) Philosophy and the Computer, pp. 137–159. Westview Press, Nashville (1992)
53. Sieg, W., Walsh, P.: Natural formalization: Deriving the Cantor–Bernstein theorem in ZF. Rev. Symb. Logic 1–35 (2019)
54. Stefferud, E.: The logic theory machine: A model heuristic program (Memorandum No. RM-3731-CC). The Rand Corporation (1963)
55. Thiele, R.: Hilbert's twenty-fourth problem. Am. Math. Mon. **110**(1), 1–24 (2003)
56. Troelstra, A.S., Schwichtenberg, H.: Basic Proof Theory, vol. 43. Cambridge University Press, Cambridge (2000)
57. Turing, A.M.: On computable numbers, with an application to the Entscheidungsproblem. Proc. Lond. Math Soc. **42**, 230–265 (1936)
58. Von Plato, J.: Gentzen's proof of normalization for natural deduction. Bull. Symb. Logic **14**(2), 240–257 (2008)
59. Wang, H.: Toward mechanical mathematics. IBM J. Res. Dev. **4**(1), 2–22 (1960)
60. Wos, L.: The problem of finding a mapping between clause representation and natural-deduction representation. J. Autom. Reason. **6**(2), 211–212 (1990)