Introduction

<u>INTRODUCTION</u>. In a broad sense, Hilbert's Proof Theory was to MEDIATE between the foundational positions that had emerged at the beginning of the 20^{th} century. Modern REDUCTIVE PROOF THEORY aims for "reductions" of possibly problematic theories to theories that concern "constructed" mathematical objects, in particular, "accessible domains". The latter are inductively generated in an equally broad sense that was articulated by Peter Aczel in 1977. Their rule-based and deterministic generation justifies INDUCTION and RECURSION PRINCIPLES; those give rise to CANONICAL ISOMORPHISMS between accessible domains that are models of the same theories.

<u>EXAMPLES</u>. $\mathbb{N}$; formulae and proofs of formal theories; arithmetic, hyperarithmetic, and analytic hierarchies; constructive

Introduction

number classes; type theories; segments of the cumulative hierarchy of sets.

The foundational issues is this: which generating rules are "permissible" from a particular philosophical perspective? And, relatedly, which inferential, logical principles can be used? — There are then deeply connected epistemological and ontological issues that can be investigated sine ira et studio! I published a long paper on these matters, "Methodological frames: Paul Bernays, Mathematical structuralism, and proof theory." It was published already more than a year ago in an OUP volume edited by Erich Reck and Georg Schiemer, "The Pre-history of Mathematical Structuralism". I attempted to address such issues. As to the methodological

Introduction

frames and their conceptual grasp, I mention that Patrick Walsh gave in his dissertation a "category theoretic characterization of accessible domains".

The GENERAL REDUCTIVE INVESTIGATIONS are certainly a very important, central direction of proof theoretic work; they can be viewed as a dramatic generalization of Hilbert's finitist consistency program. However, there is a second direction of proof theoretic work, namely, seeing proof theory as a "theory of mathematical proof". That understanding was already articulated by Hilbert in his 24th problem requesting "a theory of the method of proof in mathematics in general" and raising issues of "proof simplicity" and "identity". In 1927, he formulated the "fundamental

Introduction

idea of my proof theory" as being "none other than to describe the activity of our understanding, to make a protocol of the rules according to which our thinking actually proceeds".

Proofs & Frames

Wilfried Sieg
Carnegie Mellon University

Introduction.

It is this second direction of proof theoretic work I want to describe. It is concerned with *the discovery of the structure* and *the efficient construction* of mathematical proofs.

The issue raises the question, “What ***logical methods*** can be used in such work?”

This has become a central issue through work on computer-based formal verification and proof search.

Introduction.

Using a distinction that goes back to the 1950s and 60s, one can use two different kinds of methods, *machine-oriented* and *human-oriented* ones.

The *machine-oriented* techniques have been mostly refinements of resolution. The *human-oriented* ones range from pure “heuristics” (N&S) to incorporating aspects of natural deduction (WB) and (LR).

The concerns go actually back to the 1920s and 30s!

Introduction.

In 1933, a dissertation was completed promising the discovery of a “new symbolic logic for mathematical proofs”; that is then elaborated as follows:

It applies, as far as I can see, to all proofs in all branches of mathematics .... It makes it possible to write down the proof of a theorem in a very much shorter space than by the usual method and at the same time it makes the proof of the theorem very much clearer. … It is only necessary to give in sequence the leading ideas of the proof.

Introduction.

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. It is possible to define once and for all how the job is to be carried out.

Some History? From 1913 to 1918? From the calculus of *PM* to that introduced by Hilbert & Bernays in 1922? From there to Gentzen's natural deduction?

Introduction.

At this moment, let me just say that Gentzen built on H&B’s 1922-work when formulating the intuitionist ND calculus in his “Urdissertation” of 1932. (I will come back to the H&B calculus a little later.)

In his official dissertation of 1933, he asserted that the calculus is as close to real reasoning as possible. Thus,

… **the calculus is particularly suited for the formalization of mathematical proofs**. (p.166)

Introduction.

In his 1936-paper in which he proved the consistency of classical arithmetic, Gentzen went a step further and considered formal derivations in the ND calculus as ***images*** *of mathematical proofs*. He indicated also a direct way of obtaining these “proof images”.

Then he asserted,

The objects of proof theory shall be the proofs carried out in mathematics proper. (p. 499)

Introduction.

Gentzen’s interest was shared by one of his Göttingen friends who completed in 1933 a thesis under Weyl; that friend wrote in 1979 about this “early logical work”:

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.

Introduction.

In 1933, the friend saw himself as quite a radical and expressed his views about the developing dissertation in a letter to his mother. He saw in it the discovery of "*a new symbolic logic for mathematical proofs*".

You may remember what I read to you earlier and you may recognize him on this photo from 1929:


AT-
TAINS
HIGHEST

Introduction.

Until about 20 years ago, I was not really aware of this direction of work in the proof theoretic tradition. I had worked, however, from 1975 to 1977 under Pat Suppes on CAI in proof theory and became around 1986, for purely pedagogical reasons, interested in:

(1) *The fine structure* of mathematical proofs reflected in appropriate logical calculi, and

(2) *Strategic ways* of constructing proofs interactively and heuristics for their automated discovery.

Overview

That interest is in the background of the work on which I am reporting:

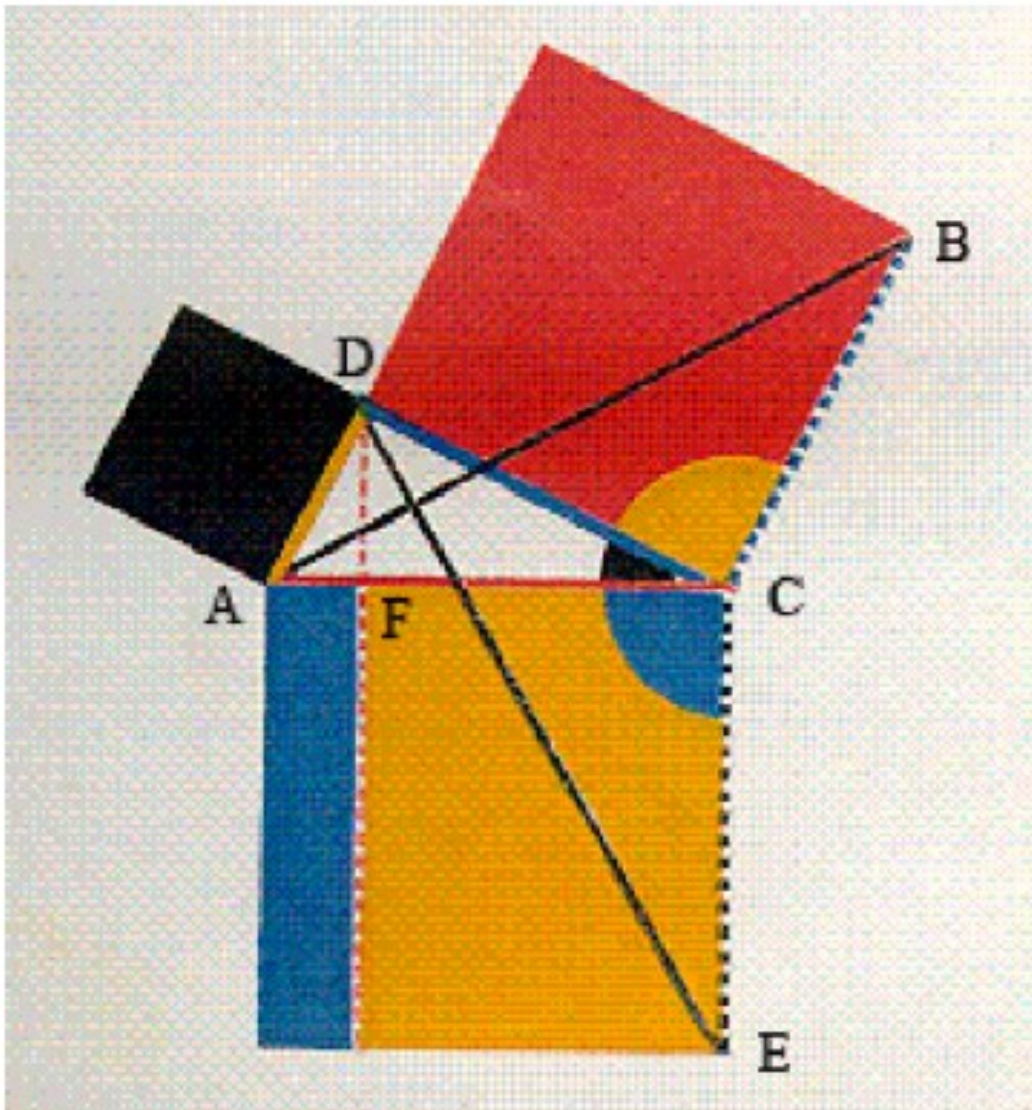
Part A. Natural formalization: Dedekind’s lemma.

Part B. Bi-directional reasoning: NIC calculi.

Part C. Automated search: Gödel’s theorems.

Part A. Natural formalization.

Part A. Natural formalization.



Part A. Natural formalization.

The *natural formalization* of a mathematical proof is achieved by using a rule-based, strategic approach that includes defined notions and operations.

In addition, one applies *Lemmas-as-rules* both "forward" and "backward".

The latter requires a careful conceptual organization, what Hilbert called a "*Fachwerk von Begriffen*".

Part A. Natural formalization.

The *natural formalization* of a mathematical proof is achieved by extending a rule-based, strategic approach to include defined notions and operations.

In addition, one applies “*Lemmas-as-rules*” both forward and backward.

The latter requires a careful conceptual organization.

Patrick Walsh and I formalized in this “natural” way the Cantor-Bernstein Theorem in ZF. (RSL 2019)

Part A. Natural formalization.

Cantor–Bernstein Theorem. Let f be an injection from a to b and g an injection from b to a; then there is a bijection h from a to b, i.e., a ≈ b.

Part A. Natural formalization.

Cantor–Bernstein Theorem. Let f be an injection from a to b and g an injection from b to a; then there is a bijection h from a to b, i.e., $a \approx b$.

The CBT is easily proved from (and easily proves) Dedekind's *Fundamental Lemma* formulated in WZ.

Part A. Natural formalization.

Cantor–Bernstein Theorem. Let f be an injection from a to b and g an injection from b to a; then there is a bijection h from a to b, i.e., $a \approx b$.

The CBT is easily proved from (and easily proves) Dedekind's *Fundamental Lemma* formulated in WZ.

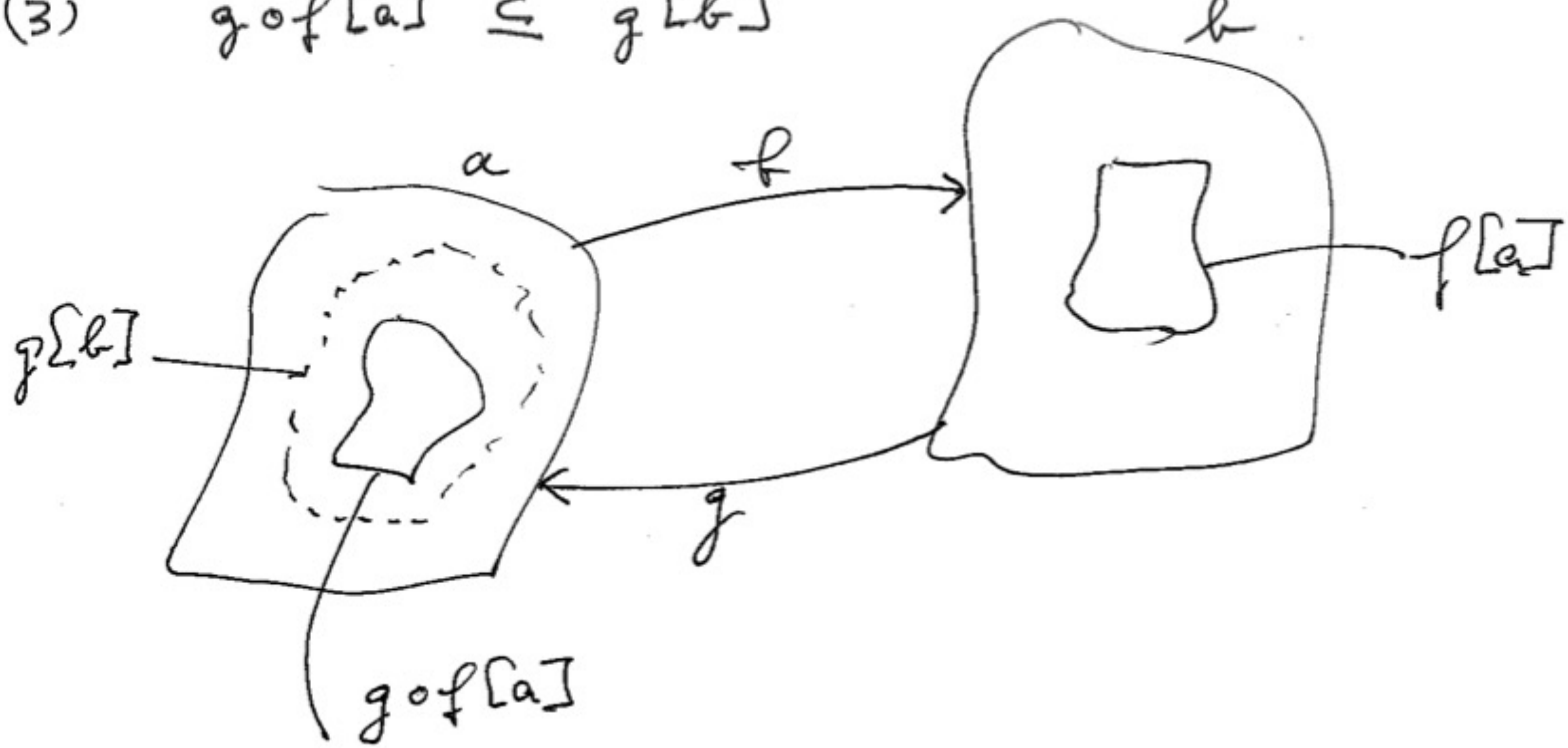
Fundamental Lemma. Let e, d, and a be sets such that $e \subseteq d \subseteq a$ and $a \approx e$; then $a \approx d$.

Part A. Natural formalization.

Assume $f \in inj(a,b)$ and $g \in inj(b,a)$; then

(1) $a \equiv f[a]$, (2) $f[a] \equiv g \circ f[a]$, and

(3) $g \circ f[a] \subseteq g[b]$



Part A. Natural formalization.

Our final ZF proof, built on about 200 lemmas is short … eight lines:

Part A. Natural formalization.

Our final ZF proof, built on about 200 lemmas is short … eight lines:

1.	$f \in \mathrm{inj}(a, b)$	*Premise*
2.	$g \in \mathrm{inj}(b, a)$	*Premise*
3.	$1(g \circ f) \in \mathrm{bij}(a, g \circ f[a])$	*Theorem (Core12)*: 1, 2
4.	$g[b] \subseteq a$	*Theorem (Func17)*: 2
5.	$g \circ f[a] \subseteq g[b]$	*Theorem (Comp11)*: 1, 2
6.	$a \approx g[b]$	*Theorem (Fundamental Lemma)*: 3, 4, 5
7.	$b \approx g[b]$	*Theorem (Equi4)*: 2
8.	$a \approx b$	*Theorem (Equi8)*: 6, 7

Part A. Natural formalization.

Our final ZF proof, built on about 200 lemmas is short … eight lines:

1.	$f \in \mathrm{inj}(a, b)$	*Premise*
2.	$g \in \mathrm{inj}(b, a)$	*Premise*
3.	$1(g \circ f) \in \mathrm{bij}(a, g \circ f[a])$	*Theorem (Core12)*: 1, 2
4.	$g[b] \subseteq a$	*Theorem (Func17)*: 2
5.	$g \circ f[a] \subseteq g[b]$	*Theorem (Comp11)*: 1, 2
6.	$a \approx g[b]$	*Theorem (Fundamental Lemma)*: 3, 4, 5
7.	$b \approx g[b]$	*Theorem (Equi4)*: 2
8.	$a \approx b$	*Theorem (Equi8)*: 6, 7

The lemmas in the formal proof articulate the direct facts for the concepts of set and function I used in the informal “diagrammatic” argument.

Part A. Natural formalization.

Remark. There are presumably many different proofs of CBT. Our analysis suggests that there are exactly two different proofs. One is Dedekind's 1887 proof and the other Zermelo's 1908 proof. They differ in the way in which they make explicit inductive definitions.

Part A. Natural formalization.

Remark. There are presumably many different proofs of CBT. Our analysis suggests that there are exactly two different proofs. One is Dedekind's 1887 proof and the other Zermelo's 1908 proof. They differ in the way in which they make explicit inductive definitions.

The difference has a beautiful mathematical core. With Dedekind's inductive definition is associated a monotone operator: Dedekind's explicit definition yields the smallest, Zermelo's the largest fixed-point of that operator.

Part B. Bi-directional frames.

Part B. Bi-directional frames.

Up to now, I have not mentioned which logical principles were used in the “natural formalization” of the Cantor-Bernstein Theorem. In the Introduction I mentioned Gentzen’s calculi of “natural reasoning” (Kalküle des natürlichen Schließens).

In this part, I’ll move from H&B’s calculus and Gentzen’s natural deduction calculi to “natural intercalation calculi” that reflect directly bi-directional reasoning, thus, more of the “actual structure of mathematical proofs”.

Part B. Bi-directional frames.

H&B's axiomatic calculus from 1922 articulates ND rules as axioms:

&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)	$(\phi \vee \psi) \to ((\phi \to \chi) \to ((\psi \to \chi) \to \chi))$

Or rather, Gentzen turned, in a very significant move, this axiomatic formulation into his rule-based system:

Part B. Bi-directional frames.

&I

$$\frac{\overset{\downarrow}{\phi} \quad \overset{\downarrow}{\psi}}{\phi \,\&\, \psi}$$

&E

$$\frac{\overset{\downarrow}{(\phi \,\&\, \psi)}}{\phi} \quad \text{and} \quad \frac{\overset{\downarrow}{(\phi \,\&\, \psi)}}{\psi}$$

∨I

$$\frac{\overset{\downarrow}{\phi}}{(\phi \vee \psi)} \quad \text{and} \quad \frac{\overset{\downarrow}{\psi}}{(\phi \vee \psi)}$$

∨E

$$\frac{\overset{\downarrow}{(\phi \vee \psi)} \quad \overset{[\phi]}{\overset{\downarrow}{\chi}} \quad \overset{[\psi]}{\overset{\downarrow}{\chi}}}{\chi}$$

Part B. Bi-directional frames.

Gentzen viewed as distinctive for his ND calculus not the interplay of I- and E-rules, but rather the *making and discharging of assumptions*. That reflected for him an important structural element of proofs in mathematical practice.

Part B. Bi-directional frames.

Gentzen viewed as distinctive for his ND calculus not the interplay of I- and E-rules, but rather the *making and discharging of assumptions*. That reflected for him an important structural element of proofs in mathematical practice.

That emphasis reveals an *unnatural* feature of ND:

(1) to prove $(A \rightarrow B)$ one assumes A and aims for B;
(2) to prove $\neg A$ one assumes A and aims for "Falsum";
(3) to prove $(\forall x)\ P(x)$ one asserts, it suffices to prove $P(z)$ (under the usual variable conditions).

Part B. Bi-directional frames.

The “proof contexts” are not part of the syntactic configuration one uses to start the sequence of proof steps.

Part B. Bi-directional frames.

The "proof contexts" are not part of the syntactic configuration one uses to start the sequence of proof steps.

Can one reflect such a proof context in a direct way via a "proof configuration"? In such a configuration, can one make "forward" and "backward" moves?

A **first step** can be taken via Gentzen's ***sequent formulation*** of ND in his (1936).

Part B. Bi-directional frames.

That *sequent formulation* indicates, locally at each proof node, on which assumptions the proof of the formula depends. It also presents the proof context.

∨I

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

∨E

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

Part B. Bi-directional frames.

That *sequent formulation* indicates, locally at each proof node, on which assumptions the proof of the formula depends. It also presents the proof context.

$\vee$I

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

$\vee$E

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

The rules operate on the rhs and retain another *unnatural* feature of ND: *detours*!

Part B. Bi-directional frames.

A second step leads essentially to Gentzen's sequent calculus with E-rules applied on the lhs and I-rules on the rhs …

… it leads also to the cut-rule, cut-elimination, and the completeness proof for the cut-free calculus.

In particular the Tait style sequent calculi for classical logic are great tools for metamathematical work.

Part B. Bi-directional frames.

This second step does not at all consider one major drawback, namely, that the applications of E-rules are in general not “goal-directed”.

Let me redescribe, in a different way, the logical problem to be addressed; that will lead then to **a third step**.

Part B. Bi-directional frames.

The task is to *close the gap* between assumptions and a goal

(i) by exploiting assumptions via sequences of E-rule applications that are *directed towards* the goal

Or

(ii) by simplifying a goal via inverted I-rules.

Part B. Bi-directional frames.

So, *closing the gap* between assumptions and a goal in this way is a species of “goal-directed” *forward* ***and*** *backward chaining*.

Part B. Bi-directional frames.

So, *closing the gap* between assumptions and a goal in this way is a species of "goal-directed" *forward* ***and*** *backward chaining*.

I have called this approach, the *intercalation method*. The rules of **N**atural **I**nter**c**alation, the **NIC** rules, have a special formulation, in particular, their E-rules.

Part B. Bi-directional frames.

So, *closing the gap* between assumptions and a goal in this way is a species of “goal-directed” *forward* ***and*** *backward chaining*.

I have called this approach, the *intercalation method*. The rules of **N**atural **I**nter**c**alation, the **NIC** rules, have a special formulation, in particular, their E-rules.

Notions: (strictly) positive subformula; extraction sequence, formula unification; structured sequent.

Part B. Bi-directional frames.

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

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

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

$$\frac{\Gamma; \Delta, (\phi \to \psi), \psi \supset \chi \quad \Gamma; \supset \phi}{\Gamma; \Delta, (\phi \to \psi) \supset \chi} \;\to E$$

Part B. Bi-directional frames.

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

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

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

$$\frac{\Gamma;\Delta,(\phi\to\psi),\psi\supset\chi\quad \Gamma;\supset\phi}{\Gamma;\Delta,(\phi\to\psi)\supset\chi}\ \to E$$

The NIC calculi for classical and intuitionist first-order logic are complete.

Part B. Bi-directional frames.

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

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

The NIC calculi for classical and intuitionist first-order logic are complete.

NIC proofs are “isomorphic” to normal ND proofs and one can strategically search for proofs via “extraction”, “inversion”, and “refutation”.

Part B. Bi-directional frames.

So, we have calculi that support a bi-directional and goal-directed proof construction in both intuitionist and classical first-order logic.

Part B. Bi-directional frames.

So, we have calculi that support a bi-directional and goal-directed proof construction in both intuitionist and classical first-order logic.

In the same way as ordinary ND trees, such NIC proofs are not “easily” displayed on a computer screen. We chose a graphical display that is inspired by the representation of ND-proofs by ***Fitch diagrams***.

Part B. Bi-directional frames.

So, we have calculi that support a bi-directional and goal-directed proof construction in both intuitionist and classical first-order logic.

In the same way as ordinary ND trees, such NIC proofs are not “easily” displayed on a computer screen. We chose a graphical display that is inspired by the representation of ND-proofs by ***Fitch diagrams***.

Differences: no “free” assumption rule; no reiteration!

Part C. Automated search.

Part C. Automated search.

Let me describe briefly the automated search for Gödel's incompleteness theorems and related results, like Löb's Theorem, for ZF.

The proofs do not use an arithmetization of syntax and the representability of the arithmetized syntactic notions. Rather, it uses the direct representation of the inductive and (structurally) recursive definitions of syntactic notions and operations in ZF.

Part C. Automated search.

Let me describe briefly the automated search for Gödel's incompleteness theorems and related results, like Löb's Theorem, for ZF.

The proofs do not use an arithmetization of syntax and the representability of the arithmetized syntactic notions. Rather, it uses the direct representation of the inductive and (structurally) recursive definitions of syntactic notions and operations in ZF.

There is one leading idea: **Prov-I** and **Prov–E**.

Part C. Automated search.

Given the representability and derivability conditions, particular self-referential sentences, and the *crucial leading idea*, AProS very efficiently found proofs of those theorems.

The proofs AProS found are *canonical*; let me show you the proof of the unprovability of the Gödel sentence G.

Sieg & Field, Automated search for Gödel's proofs, Annals of Pure and Applied Mathematics, 133, 2005.

Part C. Automated search.

Unprovability of the Gödel sentence G:

1.	$\vdash(\neg G \leftrightarrow \mathrm{thm}(\ulcorner G \urcorner))$	Prem
2.	$(\mathrm{ZFCONS\ IFF\ NOT} \vdash(\bot))$	Prem
3.	ZFCONS	Assum
4.	$\vdash(G)$	Assum
5.	$\mathrm{NOT} \vdash(\bot)$	$\leftrightarrow$ER: 2, 3
6.	$(\neg G \leftrightarrow \mathrm{thm}(\ulcorner G \urcorner))$	ProvE: 1
7.	$\mathrm{thm}(\ulcorner G \urcorner)$	Rep: 4
8.	$\neg G$	$\leftrightarrow$EL: 6, 7
9.	G	ProvE: 4
10.	$\bot$	$\bot$I: 8, 9
11.	$\vdash(\bot)$	ProvI: 10
12.	$\perp\!\!\!\perp$	$\bot$I: 5, 11
13.	$\mathrm{NOT} \vdash(G)$	$\neg$I: 12
14.	$(\mathrm{ZFCONS\ IMPLIES\ NOT} \vdash(G))$	$\rightarrow$I: 13

Part C. Automated search.

Unprovability of the Gödel sentence G:

1.	$\vdash(\neg G \leftrightarrow \text{thm}(\ulcorner G \urcorner))$	Prem
2.	(ZFCONS IFF NOT $\vdash(\bot)$)	Prem
3.	ZFCONS	Assum
4.	$\vdash(G)$	Assum
5.	NOT $\vdash(\bot)$	$\leftrightarrow$ER: 2, 3
6.	$(\neg G \leftrightarrow \text{thm}(\ulcorner G \urcorner))$	ProvE: 1
7.	$\text{thm}(\ulcorner G \urcorner)$	Rep: 4
8.	$\neg G$	$\leftrightarrow$EL: 6, 7
9.	G	ProvE: 4
10.	$\bot$	$\bot$I: 8, 9
11.	$\vdash(\bot)$	ProvI: 10
12.	$\perp\!\!\!\perp$	$\bot$I: 5, 11
13.	NOT $\vdash(G)$	$\neg$I: 12
14.	(ZFCONS IMPLIES NOT $\vdash(G)$)	$\rightarrow$I: 13

Let me show you how this proof is found strategically!

Part C. Automated search.

The particular simplified proofs of the Incompleteness Theorems were developed and implemented when I was working in Pat Suppes' IMSSS (1975-77).

The automated proof search (with Prov-E and Prov-I) was preceded, in 2003-04, by a detailed analysis and interactive construction of the proofs.

That, I assume, will hold for the formulation of any automated proof search procedure that is human-centered.

Concluding remarks.

Concluding remarks.

Let me return to Hilbert’s 1900-call for a “a theory of the method of proof in mathematics in general”.

We have taken steps towards such a theory by introducing bi-directional reasoning (and its representation in NIC calculi) that make it possible to reflect important structures of logical and mathematical proofs.

Concluding remarks.

I described some meta-mathematical results for the NIC calculi and discussed their use for natural formalization and automated search.

Such work allows us to illuminate the second direction of proof theoretic work and brings it to life by the analysis of mathematical proofs, their formal representation in suitable formalisms, and the expansion of strategies for finding intelligible proofs fully automatically.

Thank you!

References.

[1] Sieg, Searching for proofs (and uncovering capacities of the mathematical mind); in: Feferman & Sieg (eds), *Proofs, Categories and Computations* – Essays in honor of Grigori Mints, College Publications, 2010.

[2] Sieg & Walsh, Natural formalization: Deriving the Cantor-Bernstein theorem in ZF; The Review of Symbolic Logic, 2019.

[3] Sieg, The Cantor-Bernstein theorem: how many proofs?; Phil. Transactions of the Royal Society A, 2019.

[4] Sieg & Derakhshan, Human-centered automated proof search; Journal of Automated Reasoning, 2021.

These papers have a full list of references!

URL for the course Logic & Proofs: https://oli.cmu.edu/courses/logic-proofs-copy/