Proof mining

Jeremy Avigad

Department of Philosophy

Carnegie Mellon University

http://www.andrew.cmu.edu/∼avigad

Proof theory

Proof theory: the general study of deductive systems

Structural proof theory: . . . with respect to structure, transformations between proofs, normal forms, etc.

Hilbert's program:

- Formalize abstract, infinitary, nonconstructive mathematics.
- Prove consistency using only finitary methods.

More general versions:

- Prove consistency relative to constructive theories.
- Understand mathematics in constructive terms.
- Study mathematical reasoning in "concrete" terms.

Proof mining

There are

- mathematical
- computational
- foundational
- philosophical

reasons for modeling mathematical proof in formal terms.

Proof mining: using proof theoretic techniques to extract additional mathematical information from inexplicit proofs.

Proof mining

A history of proof mining:

- Kreisel (1950's-): proposes "unwinding" program, with ideas regarding Littlewood's theorem, Hilbert's Nullstellensatz, Artin's solution to Hilbert's 17th problem, L-series, Roth's theorem
- Girard (1981): analyzed proofs of van der Waerden's theorem
- Luckhardt (1989): bounds related to Roth's theorem

More recently:

- Kohlenbach and students: applications to functional analysis
- Berger, Constable, Hayashi, Schwichtenberg, et al.: extraction of algorithms from proofs
- Coquand, Delzell, Lombardi, et al.: effective versions of nonconstructive theorems in algebra

Proof mining

Notes:

- Work situated in a broader tradition.
- Recent applications to functional analysis are quite striking.
- Broader range of application is likely.

Proof mining

General requirements:

- Develop suitable (restricted) theories.
- Formalize mathematics.
- Develop general metamathematical tools.

Specific requirements:

- Find a suitable domain of application.
- Understand methods, information sought.
- Refine general metamathematical tools.

Overview of lectures

1. General frameworks: theories of arithmetic

   primitive recursive, first-order, higher-type, higher-order

2. General methods and results

   cut-elimination, double-negation translations, realizability, the Dialectica interpretation, Herbrand's theorem, conservation results

3. Formalizing analysis

   complete separable metric spaces, Hilbert and Banach spaces, continuity, compactness, distance, representation issues

4. Weak König's lemma and monotone functional interpretation

5. Applications to functional analysis

   extracting rates of strong unicity from approximation theorems, rates of asymptotic regularity in fixed-point theorems for nonexpansive mappings, uniformity results

Mathematics in restricted frameworks

Some historical landmarks:

- Weyl (*Das Kontinuum*, 1918): Developed analysis (informally) in predicative subsystems of second-order arithmetic
- Heyting (1930's): Formalizations of intuitionistic logic
- Hilbert and Bernays (*Grundlagen der Mathematik*, 1934/1939): Analysis in second-order arithmetic

Subsequent efforts by Kreisel, Feferman, Takeuti, Friedman, Simpson, and many others.

Ongoing work in:

- Constructive mathematics
- Recursive mathematics
- Reverse mathematics

Mathematics in restricted frameworks

The words of a budding “Bolshevik”:

The *circulus vitiosis*, which is cloaked by the hazy nature of the usual concept of set and function, but which we reveal here, is surely not an easily dispatched formal defect in the construction of analysis.. . . But the more distinctly the logical analysis is brought to givenness and the more deeply and completely the glance of consciousness penetrates it, the clearer it becomes that, given the current approach to foundational matters, every cell (so to speak) of this mighty organism is permeated by the poison of contradiction and that a thorough revision is necessary to remedy the situation.

Weyl, *Das Kontinuum*, Chapter 1, §6

Mathematics in restricted frameworks

Distinctions:

- Formal vs. informal
- Normative vs. descriptive
- A priori vs. a posteriori

Proof mining:

- With some care (cutting down on generality, paying attention to representations) many mathematical arguments can be represented in restricted theories.
- General metamathematical methods and techniques are available for extracting additional information from proofs in such theories.
- In specific cases, this can be mathematically informative.

Logic

Mathematics = logic + axioms.

Logic comes in two flavors:

- Constructive (intuitionistic or minimal): rules have a clear computational interpretation.
- Classical: add the law of the exclude middle

The same goes for theories:

- Constructive theories have constructive axioms.
- Classical theories describe objects independent of constructions.

A related dichotomy:

- Intensional: math deals with representations
- Extensional: math deals with objects independent of representations

Primitive recursive functions

The set of *primitive recursive functions* is the smallest set

- containing 0, $S(x) = x + 1$, $p_i^n(x_1, \dots, x_n) = x_i$
- closed under composition
- closed under primitive recursion:

$$f(0, \vec{z}) = g(\vec{z}), \qquad f(x + 1, \vec{z}) = h(f(x, \vec{z}), x, \vec{z})$$

Can define pairing and sequencing, and then handle operations on integers, rational numbers, lists, graphs, trees, finite sets, etc.

Viewpoints:

- Set theory: very weak model of computation
- Computer science: far too strong
- Proof theory: just right

Primitive recursive arithmetic

Primitive recursive arithmetic is an axiomatic theory, with

- defining equations for the primitive recursive functions
- quantifier-free induction

$$\frac{\varphi(0) \qquad \varphi(x) \to \varphi(x+1)}{\varphi(t)}$$

In *PRA*, can handle most (all?) “finitary” proofs.

PRA can be presented either as a first-order theory (classical or intuitionistic) or as a quantifier-free calculus.

Theorem (Herbrand). Suppose first-order classical *PRA* proves $\forall x \; \exists y \; \varphi(x, y)$, with φ quantifier-free. Then for some function symbol f, quantifier-free *PRA* proves $\varphi(x, f(x))$.

First-order arithmetic

First-order arithmetic is essentially *PRA* plus induction. *Peano arithmetic* (*PA*) is classical, *Heyting arithmetic* (*HA*) is intuitionistic.

Language: 0, S, $+$, $\times$, $<$.

Axioms: quantifier-free defining axioms, induction.

A formula is

- Δ_0 if every quantifier is bounded
- Σ_1 if of the form $\exists \vec{x}\ \varphi$, $\varphi \in \Delta_0$
- Π_1 if of the form $\forall \vec{x}\ \varphi$, $\varphi \in \Delta_0$
- Δ_1 if equivalent to Σ_1 and Π_1

Primitive recursive functions / relations have Σ_1 / Δ_1 definitions.

First-order arithmetic

$I\Sigma_1$ is the restriction of *PA* with induction for only Σ_1 formulas.

This theory suffices to define the primitive recursive functions, and hence interpret *PRA*. Conversely:

Theorem (Parsons, Mints, Takeuti). $I\Sigma_1$ is conservative over *PRA* for Π_2 sentences: if

$$I\Sigma_1 \vdash \forall x\ \exists y\ \varphi(x, y),$$

with φ quantifier-free, then

$$PRA \vdash \varphi(x, f(x))$$

for some function symbol f.

Higher-type recursion

The *finite types* are defined as follows:

- N is a finite type
- If σ and τ are finite types, so are $\sigma \times \tau$ and $\sigma \to \tau$

For example, the reals can be represented as type $\mathsf{N} \to \mathsf{N}$ functionals. Functions from $\mathbb{R}$ to $\mathbb{R}$ can be represented by type $(\mathsf{N} \to \mathsf{N}) \to (\mathsf{N} \to \mathsf{N})$.

The *primitive recursive* functionals *of finite type* allow:

- λ abstraction, application, pairing, projection
- Higher-type primitive recursion:

$$F(0) = G, \quad F(n+1) = H(F(n), n)$$

The theory PRA^{ω} (i.e. Gödel's theory T) axiomatizes these.

Higher-type recursion

Ackermann's functions can be defined using primitive recursive functionals:

$$\begin{array}{rcl} F(0) & = & S \\ F(n+1) & = & \mathit{Iterate}(n+2, F(n)) \\ \mathit{Iterate}(0, G) & = & \lambda x\ x \\ \mathit{Iterate}(n+1, G) & = & \lambda x\ (G(\mathit{Iterate}(n, G)(x))) \end{array}$$

Restricted higher-type primitive recursion:

$$F(0, \vec{z}) = G(\vec{z}), \quad F(n+1, \vec{z}) = H(n, F(n, \vec{z}), \vec{z})$$

where $F(n, \vec{z})$ has type N.

The theory $\widehat{PRA}^{\omega}$ axiomatizes these functionals, and is a conservative extension of PRA.

Higher-type arithmetic

Higher-type arithmetic:

- $PA^{\omega} = PRA^{\omega} + \text{induction}$
- $HA^{\omega} = PRA_i^{\omega} + \text{induction}$

These are conservative extensions of *PA* and *HA* respectively.

In fact, one can add quantifier-free choice axioms (QF-AC) to PA^{ω}, and full choice (AC) to HA^{ω}.

Restricted versions are conservative over *PRA* and PRA_i:

- $\widehat{PRA}^{\omega} + (QF\text{-}AC)$
- $\widehat{PRA_i}^{\omega} + (AC)$

There are weaker / stronger analogues.

Higher-type arithmetic

Some useful notation:

- N is “type 0”
- $\mathsf{N} \to \mathsf{N}$ is “type 1”
- $(\mathsf{N} \to \mathsf{N}) \to \mathsf{N}$ is “type 2”
- $\ldots$

An object of type $n + 1$ is, essentially (with currying and pairing), a functional $F(x_1, \ldots, x_k)$, taking arguments of type n, and returning a natural number.

This corresponds to $V_{\omega+n}$ in the set-theoretic hierarchy.

Extensionality

One can take type N equality to be basic, then define higher-type equality extensionally:

$$f = g \equiv \forall x \ (fx = gx)$$

The extensionality axiom says that functions respect this equality:

$$f = g \to Ff = Fg.$$

Alternatively, one can take $=$ to be basic at each type, and have axioms asserting that this corresponds to the extensional notion.

Theorem (Luckhardt). Extensionality can be interpreted, preserving the second-order (type $0/1$) fragment.

More generally, though, the issues are subtle.

Second / higher-order arithmetic

Functions vs. sets:

- With function types, interpret sets via characteristic functions:

$$x \in S \equiv \chi_S(x) = 1.$$

- With relation types, interpret functions as functional relations, $\forall x \; \exists! y \; R(x, y)$

Can add comprehension axioms,

$$\exists S \; \forall x \; (x \in S \leftrightarrow \varphi),$$

and choice axioms

$$\forall x \; \exists y \; \varphi(x, y) \to \exists f \; \forall y \; (\varphi(x, f(x))).$$

Classically, choice implies comprehension, and the latter are very strong.

Subsystems of second-order arithmetic

Restrict induction to Σ^0_1 formulas with parameters, and restrict set existence principles:

- *RCA$_0$*: recursive (Δ^0_1) comprehension
  Recursive analysis.
- *WKL$_0$*: paths through infinite binary trees
  Compactness.
- *ACA$_0$*: arithmetic comprehension
  Analytic principles like the least-upper bound principle.
- Π^1_1-*CA$_0$*: Π^1_1 comprehension
  Strong analytic principles.

RCA$_0$ and *WKL$_0$* are conservative over *PRA* for Π_2 sentences. *ACA$_0$* is conservative over *PA*.

Summary

Languages of interest:

- Variables ranging over natural numbers
- Variables ranging over finite types
- Variables ranging over sets
- Constants for primitive recursive functionals

Logic: may be classical or intuitionistic

Axioms:

- Various types of primitive recursion
- Various types of induction
- Various set (and function) existence principles

Overview of lectures

1. General frameworks: theories of arithmetic
2. General methods and results
3. Formalizing analysis
4. Weak König's lemma and monotone functional interpretation
5. Applications to functional analysis

Cut elimination and normalization

The idea: suppose you prove a lemma,

$$\forall x \ (\varphi(x) \to \psi(x) \wedge \theta(x)).$$

Later, in a proof, knowing $\varphi(t)$, you use the lemma to conclude $\theta(t)$.

You could have proceeded more directly, though perhaps less efficiently, by deriving $\theta(t)$ directly from $\varphi(t)$.

A sequent calculus

$$\Gamma, A \Rightarrow A \qquad\qquad \Gamma, \bot \Rightarrow A$$

$$\frac{\Gamma, \varphi_i \Rightarrow \psi}{\Gamma, \varphi_0 \wedge \varphi_1 \Rightarrow \psi} \qquad\qquad \frac{\Gamma \Rightarrow \varphi \qquad \Gamma \Rightarrow \psi}{\Gamma \Rightarrow \varphi \wedge \psi}$$

$$\frac{\Gamma, \varphi \Rightarrow \psi \qquad \Gamma, \theta \Rightarrow \psi}{\Gamma, \varphi \vee \theta \Rightarrow \psi} \qquad\qquad \frac{\Gamma \Rightarrow \varphi_i}{\Gamma \Rightarrow \varphi_0 \vee \varphi_1}$$

$$\frac{\Gamma, \Rightarrow \varphi \qquad \Gamma, \theta \Rightarrow \psi}{\Gamma, \varphi \to \theta \Rightarrow \psi} \qquad\qquad \frac{\Gamma, \varphi \Rightarrow \psi}{\Gamma \Rightarrow \varphi \to \psi}$$

$$\frac{\Gamma, \varphi[t/x] \Rightarrow \psi}{\Gamma, \forall x\, \varphi \Rightarrow \psi} \qquad\qquad \frac{\Gamma, \Rightarrow \psi}{\Gamma \Rightarrow \forall x\;\psi}$$

$$\frac{\Gamma, \varphi \Rightarrow \psi}{\Gamma, \exists x\;\varphi \Rightarrow \psi} \qquad\qquad \frac{\Gamma, \Rightarrow \psi[t/x]}{\Gamma \Rightarrow \exists x\;\psi}$$

The cut elimination theorem

The cut rule:

$$\frac{\Gamma \Rightarrow \varphi \qquad \Gamma, \varphi \Rightarrow \psi}{\Gamma \Rightarrow \psi}$$

For classical logic, allow sets of formulas on the right:

$$\Gamma \Rightarrow \Delta$$

Read: the conjunction of Γ implies the disjunction of Δ. For example,

$$\cfrac{\cfrac{\varphi \Rightarrow \varphi, \bot}{\Rightarrow \varphi, \neg\varphi}}{\Rightarrow \varphi \vee \neg\varphi}$$

Theorem (Gentzen). Cuts can be eliminated from proofs in pure first-order logic.

Applications of cut-elimination

Theorem (Gentzen). Suppose $\exists x\ \varphi(x)$ is provable intuitionistically. Then for some term t, so is $\varphi(t)$.

Proof. Consider the last inference in a cut-free proof.

Theorem (Herbrand). Suppose $\exists x\ \varphi(x)$ is provable classically, and φ is quantifier-free. Then for some sequence of terms,

$$\varphi(t_1) \vee \ldots \vee \varphi(t_k)$$

has a quantifier-free proof.

Proof. Again, consider the structure of a cut-free proof.

The latter (slightly generalized) shows that quantifiers can be eliminated in *PRA*.

Applications of cut-elimination

$I\Sigma_1$ is the variant of Peano arithmetic in which induction is restricted to Σ_1 formulas.

Theorem (Parsons, Mints, Takeuti). $I\Sigma_1$ is conservative over *PRA* for Π_2 sentences.

Proof (Sieg, Buss): Express induction as a rule:

$$\frac{\Gamma \Rightarrow \exists y\ \varphi(0, y), \Delta \qquad \Gamma, \exists y\ \varphi(x, y) \Rightarrow \exists y\ \varphi(x+1, y), \Delta}{\Gamma \Rightarrow \exists y\ \varphi(x, y), \Delta}$$

Eliminate cuts except on axioms. “Read off” witnessing information.

Conservation results

RCA_0 has Σ_1 induction and a Δ_1 comprehension axiom:

$$\forall x\ (\varphi(x) \leftrightarrow \psi(x)) \rightarrow \exists Z\ \forall x\ (x \in Z \leftrightarrow \varphi(x)),$$

where φ is Σ_1 and ψ is Π_1.

Theorem. RCA_0 is conservative over $I\Sigma_1$.

Proof. Interpret sets by indices for computable sets.

ACA_0 adds arithmetic comprehension.

Theorem. ACA_0 is conservative over PA.

Proof. Add names for arithmetically definable sets, and eliminate cuts.

In fact, one can eliminate an arithmetic choice principle, (Σ^1_1-AC).

Shortcomings of cut elimination

Proofs can get longer: there are superexponential lower bounds, originally due to Statman and Orevkov, independently.

Also, the procedure is not *modular*: you can't eliminate cuts one lemma at a time.

Modified realizability

Kleene: one can assign to every constructive theorem of arithmetic a computable “realizer.”

Kreisel: for theorems of HA^{ω}, one can take realizers to be primitive recursive functionals.

Define “a realizes η” inductively:

1. If θ is atomic, a realizes θ if and only if θ is true.
2. a realizes $\varphi \wedge \psi$ if and only if $(a)_0$ realizes φ and $(a)_1$ realizes ψ.
3. a realizes $\varphi \vee \psi$ if and only if $(a)_0 = 0$ and $(a)_1$ realizes φ, or $(a)_0 = 1$ and $(a)_2$ realizes ψ.
4. a realizes $\varphi \to \psi$ if and only if when b realizes φ, $a(b)$ realizes ψ.
5. a realizes $\forall z\ \varphi(z)$ if and only if for every b, $a(b)$ realizes $\varphi(b)$.
6. a realizes $\exists z\ \varphi(z)$ if and only if $(a)_1$ realizes $\theta((a)_0)$.

Theorem. If $HA^{\omega} \vdash \eta$, then for some t, $HA^{\omega} \vdash$ “t realizes η”.

Modified realizability

In fact, the axiom of choice (AC), is also realized:

$$\forall x\ \exists y\ \varphi(x, y) \to \exists f\ \exists x\ \varphi(x, f(x)).$$

Also a certain “independence of premise” axiom (IP') for ∃-free formulas.

Theorem (Troelstra). We have

$$HA^{\omega} + (AC) + (IP') \vdash \varphi$$

if and only if, for some term t,

$$HA^{\omega} \vdash t \text{ realizes } \varphi.$$

One can also add extensionality and ∃-free axioms to both sides of the equivalence.

Shortcomings of realizability

Realizers do not always carry useful computational information.
For example,

$$\forall x\ A(x) \to \forall x\ B(x)$$

is realized (by anything) if and only if it is true.

Also, a negation never carries any computational information: $\neg\varphi$ is realized (by anything) if and only if φ is not realized.

There are tricks to recover information from negation:

- The Friedman-Dragalin trick
- The Coquand-Hoffmann trick

These are useful in conjunction with double-negation interpretations.

Double-negation translations

The Gödel-Gentzen double-negation translation interprets classical logic in minimal logic:

- $A^N \equiv \neg\neg A$ for atomic A
- $(\varphi \vee \psi)^N \equiv \neg(\neg\varphi^N \wedge \neg\psi^N)$
- $(\exists x\ \varphi)^N \equiv \neg\forall x\ \neg\varphi^N$.

The translation commutes with $\forall$, $\wedge$, $\to$.

Theorem. If $\Gamma \vdash \varphi$ classically, $\Gamma^N \vdash \varphi^N$ in minimal logic.

Corollary. If $PA^\omega \vdash \varphi$, then $HA^\omega \vdash \varphi^N$

The Kuroda translation, instead, adds $\neg\neg$ after each universal quantifier.

Theorem. $\neg\neg\varphi^K$ and φ^N are intuitionistically equivalent.

The Dialectica interpretation

Assigns to every formula φ in the language of PRA^ω a formula

$$\varphi^D \equiv \exists x\ \forall y\ \varphi_D(x, y)$$

where x and y are sequences of variables and φ_D is quantifier-free.

Idea: $\forall y\ \varphi_D(x, y)$ asserts that x is a “strong” realizer for φ.

Note: in contrast to realizability, $\forall y\ \varphi_D(x, y)$ is universal.

Inductively one shows:

Theorem (Gödel). If HA^ω proves φ, there is a sequence of terms t such that quantifier-free PRA^ω proves $\varphi_D(t, y)$.

The Dialectica interpretation

Define the translation inductively, assuming

$$\varphi^D = \exists x\ \forall y\ \ \varphi_D \quad \text{and} \quad \psi^D = \exists u\ \forall v\ \psi_D.$$

1. For θ an atomic formula, $\theta^D = \theta_D = \theta$.
2. $(\varphi \wedge \psi)^D = \exists x, u\ \forall y, v\ (\varphi_D \wedge \psi_D)$.
3. $(\varphi \vee \psi)^D = \exists z, x, u\ \forall y, v\ ((z = 0 \wedge \varphi_D) \vee (z = 1 \wedge \psi_D))$.
4. $(\forall z\ \varphi(z))^D = \exists X\ \forall z, y\ \varphi_D(X(z), y, z)$.
5. $(\exists z\ \varphi(z))^D = \exists z, x\ \forall y\ \varphi_D(x, y, z)$.
6. $(\varphi \to \psi)^D = \exists U, Y\ \forall x, v\ (\varphi_D(x, Y(x, v)) \to \psi_D(U(x), v))$.

The last clause is a Skolemization of the formula

$$\forall x\ \exists u\ \forall v\ \exists y\ (\varphi_D(x, y) \to \psi_D(u, v)).$$

Interpreting modus ponens

Consider the rule “from φ and $\varphi \to \psi$ conclude ψ.”

We are given terms a, b, and c such that PRA^ω proves

$$\varphi_D(a, y)$$

and

$$\varphi_D(x, b(x, v)) \to \psi_D(c(x), v).$$

We need a term d such that PRA^ω proves

$$\psi_D(d, v).$$

Substituting $b(a, v)$ for y in the first hypothesis and a for x in the second, we see that taking $d = c(a)$ works.

Advantages of the Dialectica interpretation

For example, the formula

$$\forall x\ A(x) \to \forall x\ B(x)$$

translates to

$$\exists f\ \forall x\ (A(f(x)) \to B(x))$$

Markov's principle is verified:

$$\neg\neg\exists x\ A(x) \to \exists x\ A(x)$$

So is the axiom of choice:

$$\forall x\ \exists y\ \varphi(x, y) \to \exists f\ \forall x\ \varphi(x, f(x)).$$

An "independence of premise" principle ($IP_\forall$) is also interpreted.

The Dialectica interpretation

The D-interpretation works well on restricted theories.

Theorem. If $\widehat{PRA}_i^{\omega} + (AC) + (MP)$ proves φ, then $\widehat{PRA}_i^{\omega} \vdash \varphi^D$.

Corollary. If $\widehat{PRA}^{\omega} + (QF\text{-}AC)$ proves $\forall x\ \exists y\ \varphi(x, y)$, with φ q.f. in the language of PRA, then so does $\widehat{PRA}^{\omega}$.

$(QF\text{-}AC)$ can be used to prove Σ_1 induction:

$$\exists y\ \varphi(0, y) \wedge \forall x\ (\exists y\ \varphi(x, y) \rightarrow \exists y\ \varphi(x + 1, y)) \rightarrow \forall x\ \exists y\ \varphi(x, y).$$

So this shows that $I\Sigma_1$ is Π_2 conservative over PRA.

Applying the Dialectica interpretation

Recipe:

- Start with a nonconstructive proof.
- Formalize it in $PA^\omega + (QF\text{-}AC)$.
- Apply a double-negation translation.
- Get a proof in $HA^\omega + (MP) + (IP) + (AC)$.
- Apply the Dialectica interpretation

Later: we will see that certain nonconstructive principles, like weak König's lemma, can also be eliminated.

We will also consider a modification of the D-interpretation, due to Kohlenbach, that makes it easier to extract bounds instead of witnesses.

The no-counterexample translation

Consider, for example, what the ND-translation does to prenex arithmetic formulas, such as

$$\forall x \; \exists y \; \forall z \; A(x, y, z). \qquad (*)$$

Negate, Skolemize, and negate again:

$$\forall x, Z \; \exists y \; A(x, y, Z(y))$$

If (*) is true, one can compute a y from x and Z. Such a y foils the putative counterexample function, Z.

The Skolemization of this formula is Kreisel's *no-counterexample* interpretation:

$$\exists Y \; \forall x, Z \; A(x, Y(x, Z), z(Y(x, Z))),$$

This works for any number of quantifiers.

An example: the Hilbert basis theorem

Hilbert's basis theorem says that every polynomial ideal in $Q[x_1, \ldots, x_k]$ is finitely generated.

Let F range over sequences of polynomials. Formalize this as:

$$\forall F \; \exists n \; \forall m \; (F(m) \in \langle F(0), \ldots, F(n) \rangle).$$

This is computationally false.

The Hilbert basis theorem is classically equivalent to

$$\forall F \; \exists n \; (F(n+1) \in \langle F(0), \ldots, F(n) \rangle).$$

This is computationally true, but hard to prove constructively.

Consider the no-counterexample interpretation:

$$\forall F, M \; \exists n \; (F(M(n)) \in \langle F(0), \ldots, F(n) \rangle).$$

The Hilbert basis theorem

Consider the no-counterexample interpretation:

$$\forall F, M\ \exists n\ (F(M(n)) \in \langle F(0), \ldots, F(n)\rangle).$$

Hertz (2004) shows:

- Translating a standard proof that uses Dickson's lemma yields a constructive proof.
- The translation is modular, i.e. proceeds lemma by lemma.
- Translating a different proof (by Simpson) yields sharp bounds.

This is just an illustration. In real proof mining:

- Proofs are more complex.
- More elaborate principles get eliminated (e.g. compactness).
- Information obtained is genuinely useful.

Overview of lectures

1. General frameworks: theories of arithmetic
2. General methods and results
3. Formalizing analysis
4. Weak König's lemma and monotone functional interpretation
5. Applications to functional analysis

Conservation results summarized

The following theories are “finitary”:

- PRA
- $I\Sigma_1$
- RCA_0, WKL_0
- $\widehat{PRA}^{\omega} + (QF\text{-}AC) + (WKL)$

The following theories are “arithmetic”:

- PA
- ACA_0, $\Sigma^1_1\text{-}AC_0$
- PRA^{ω}
- $PA^{\omega} + (QF\text{-}AC) + (WKL)$
- $HA^{\omega} + (AC) + (MP) + (WKL)$.

These suffice for many purposes!

Formalizing mathematics

In the language of *PRA*, one can define integers, rational numbers, and other finitary objects in natural ways.

Define the real numbers to be Cauchy sequences of rationals with a fixed rate of convergence:

$$\forall n \, \forall m \geq n \; (|a_n - a_m| < 2^{-n}).$$

Equality is a Π_1 notion:

$$a = b \equiv \forall n \; (|a_n - b_n| \leq 2^{-n+1}).$$

Less-than is a Σ_1 notion:

$$a < b \equiv \exists n \; (a_n + 2^{-n+1} < b_n).$$

Complete separable metric spaces

Definition. A *complete separable metric space* $X = \hat{A}$ consists of a set A together with a function $d : A \times A \to \mathbb{R}$ satisfying:

- $d(x, x) = 0$
- $d(x, y) = d(y, x)$
- $d(x, z) \leq d(x, y) + d(y, z)$.

A *point of* $\hat{A}$ is a sequence $\langle a_n \mid n \in \mathbb{N} \rangle$ of elements of A such that for every n and $m > n$ we have $d(a_n, a_m) < 2^{-n}$.

Examples of complete separable metric spaces:

- $\mathbb{R}$, $\mathbb{C}$, $\mathbb{Q}_p$
- Infinite products: Baire Space, Cantor Space
- $C(X)$, for X a compact space
- Measure spaces: $L_1(X)$, $L_2(X)$, $L_p(X)$
- l_p, c_0

Compactness

Three notions of compactness for a CSM:

- Totally bounded: for every rational $\varepsilon > 0$, there is a finite ε-net.
- Heine-Borel compact: every covering by open sets has a finite subcover.
- Sequentially compact: every sequence has a convergent subsequence.

In weak theories:

- RCA_0 proves e.g. $[0, 1]$ is totally bounded.
- Totally bounded $\Rightarrow$ Heine-Borel requires weak König's lemma.
- Totally bounded $\Rightarrow$ sequentially compact requires arithmetic comprehension.

In constructive mathematics, one usually uses "totally bounded."

Continuity

A function f between CSM's is *uniformly continuous* if

$$\forall \varepsilon > 0 \, \exists \delta > 0 \, \forall x, y \; (d(x, y) < \delta \rightarrow d(f(x), f(y)) < \varepsilon).$$

A *modulus of uniform continuity* for f is a function $g(\varepsilon)$ returning such a δ for each ε:

$$\forall x, y, \varepsilon > 0 \; (d(x, y) < g(\varepsilon) \rightarrow d(f(x), f(y)) < \varepsilon).$$

Theorem. In *RCA*$_0$, the statement that every continuous function from a compact space to $\mathbb{R}$ has modulus of uniform continuity is equivalent to (*WKL*).

In constructive mathematics, functions are usually assumed to come with such moduli.

Closed sets

Two notions of a closed set:

- *closed* = complement of a sequence of basic open balls
- *separably closed* = the closure of a sequence of points

A set with a distance function is said to be *located*.

Many of the relationships between these notions have been worked out by Simpson, Brown, Giusto, Marcone, Avigad, Simic, and others.

For example, in the base theory RCA_0, Brown shows:

1. In compact spaces, “closed ⇒ separably closed” is equivalent to (ACA)
2. In general, “closed ⇒ separably closed” is equivalent to $(\Pi^1_1\text{-}CA)$
3. “separably closed ⇒ closed” is equivalent to (ACA)

Distance

Theorem (Avigad, Simic): Over RCA_0, the following are equivalent to (ACA):

1. In a compact space, if C is any closed set and x is any point, then $d(x, C)$ exists.
2. If C is any closed subset of $[0, 1]$, then $d(0, C)$ exists.

The following are equivalent to $(\Pi^1_1\text{-}CA)$:

1. In an arbitrary space, if C is any closed set and x is any point, then $d(x, C)$ exists.
2. In a compact space, if S is any G_δ set and x is any point, then $d(x, S)$ exists.
3. If S is a G_δ subset of $[0, 1]$, then $d(0, S)$ exists.

In constructive mathematics, sets are often assumed to be located.

Hilbert space and Banach spaces

A *Hilbert space* $H = \hat{A}$ consists of a countable vector space A over $\mathbb{Q}$ together with a function $\langle \cdot, \cdot \rangle : A \times A \to \mathbb{R}$ satisfying

1. $\langle x, x \rangle \geq 0$
2. $\langle x, y \rangle = \langle y, x \rangle$
3. $\langle ax + by, z \rangle = a\langle x, z \rangle + b\langle y, z \rangle$

Define $||x|| = \langle x, x \rangle^{\frac{1}{2}}$ and $d(x, y) = ||x - y||$, and think of H as the completion of A.

Similarly, a *Banach space* is represented as the completion of a countable vector space under a norm.

Overview of lectures

1. General frameworks: theories of arithmetic
2. General methods and results
3. Formalizing analysis
4. Weak König's lemma and monotone functional interpretation
5. Applications to functional analysis

Weak König’s lemma

In the language of second-order arithmetic, a *tree on {0, 1}* is a set of finite binary sequences closed under initial segments.

A *path* through T is a function $f : \mathbb{N} \to \{0, 1\}$ such that for every n, $\bar{f}(n) = \langle f(0), \ldots, f(n-1)\rangle$ is in T.

Weak König’s lemma is the assertion that every infinite tree on $\{0, 1\}$ has a path.

Weak König's lemma

A basis for $\{0, 1\}^\omega$ is given by (clopen) sets of the form

$$[\sigma] = \{f \mid f \supset \sigma\}.$$

Trees code closed sets:

- If T is a tree, the set of paths through T is closed.
- If C is closed, $\{\sigma \mid [\sigma] \cap C \neq \emptyset\}$ is a tree.

One can also effectively represent the intersection of a sequence of closed sets as the set of paths through a tree.

Theorem. In RCA_0, the following are equivalent to (WKL):

1. $\{0, 1\}^\omega$ is compact.
2. $[0, 1]$ is compact.
3. First-order logic is compact.

Weak König's lemma

Theorem (Friedman). $\mathit{WKL_0}$ is conservative over PRA for Π_2 sentences.

First syntactic proof, using cut-elimination, by Sieg.

Theorem (Harrington). $\mathit{WKL_0}$ is conservative over $\mathit{RCA_0}$ for Π^1_1 sentences.

Syntactic proofs by Hájek, Avigad.

Theorem (Kohlenbach). $\widehat{\mathit{PRA}}^{\omega} + (\mathit{QF}\text{-}\mathit{AC}) + (\mathit{WKL})$ is conservative over PRA for Π_2 sentences.

Kohlenbach used the Dialectica interpretation.

All proofs and results extend to weaker theories.

Weak König's lemma

The idea behind Sieg's proof:

- Note that because $\{0, 1\}^\omega$ is compact, if $G(f, \vec{z})$ is continuous, it is bounded by an $H(\vec{z})$.
- Observation (Howard): if G is a primitive recursive functional, there is a primitive recursive H, and one can verify this axiomatically.
- Put *WKL* in contrapositive form: if there is no path through T, then T is finite.
- Express (*WKL*) as a rule:

$$\frac{\Gamma \Rightarrow \forall f\ \exists n\ (\bar{f}(n) \notin T), \Delta}{\Gamma \Rightarrow \exists n\ \forall \sigma\ (length(\sigma) = n \rightarrow \sigma \notin T), \Delta}$$

- Eliminate cuts.
- From a witness for the top, obtain a witness for the bottom.

Majorizability

A similar idea works for higher types.

Definition (Howard). Define $a \leq^*_\tau b$, read *a is hereditarily majorized by b*, by induction on the type of τ:

- $a \leq^*_{\mathsf{N}} b \equiv a \leq b$
- $a \leq^*_{\rho \to \sigma} b \equiv \forall x, y\ (x \leq^*_\rho y \to a(x) \leq^*_\sigma b(y))$

For example, g majorizes f at type $\mathsf{N} \to \mathsf{N}$ if for every x, $g(x)$ is greater than or equal to $f(0), f(1), \ldots, f(x)$.

Proposition. If $x \geq^* y$ and $y \geq z$ (pointwise) then $x \geq^* z$.

Proposition. Every term of PRA^ω has a majorant.

Weak König's lemma

Saying $f \in \{0, 1\}$ is equivalent to $f \leq^* \lambda x \,.\, 1$.

So, if $\lambda f \; G(f, \vec{x})$ is majorized by $\lambda f \; H(f, \vec{x})$, then for each f, $G(f, \vec{x}) \leq H(\lambda x \,.\, 1, \vec{x})$.

Theorem. $\widehat{PRA}^{\omega} + (QF\text{-}AC) + (WKL)$ is conservative over $\widehat{PRA}^{\omega}$ for Π_2 sentences.

Proof. Use the Dialectica interpretation. If the source theory proves $\forall x \; \exists y \; A(x, y)$, then $\widehat{PRA}^{\omega}$ proves the ND-translation of

$$(WKL) \to \forall x \; \exists y \; A(x, y).$$

Majorizability can be used to eliminate the dependence on the hypothesis.

The same result holds for conclusions of the form $\forall x \; \forall f \; \exists y \; A(f, x, y)$.

Weak König's lemma

Observations:

- Often, one only cares about bounds, not witnesses.
- Some principles, like (*WKL*), don't affect bounds.

A "monotone" variant of the Dialectica interpretation, due to Kohlenbach, interprets every formula in by one of the form

$$\exists x^* \, \exists x \leq^* x^* \, \forall y \; A(x, y)$$

Let ($QF\text{-}AC^{0,1}$) denote, for quantifier-free φ,

$$\forall x^1 \, \exists y^0 \, \varphi(x, y) \rightarrow \exists f^{1 \rightarrow 0} \, \forall x^1 \, \varphi(x, f(x))$$

Let T^ω denote, e.g., the theory $\widehat{PRA}^\omega + (QF\text{-}AC^{1,0})$.

A metatheorem

Theorem (Kohlenbach 1992). Let Δ be a set of closed axioms of the form

$$\forall u^1 \, \exists v^1 \leq tu \, \forall w^0 \; A(u, v, w)$$

where t is closed and A is q.f. Suppose

$$T^\omega + \Delta \vdash \forall x^1 \, \forall y^1 \leq sx \, \exists z^0 \; B(x, y, z).$$

with B q.f. Then there is a term Φ such that

$$T^\omega_i + \Delta_\varepsilon \vdash \forall x^1 \, \forall y^1 \leq sx \, \exists z^0 \leq \Phi x \;\; B(x, y, z)$$

where Δ_ε are "ε-weakenings" of formulas in Δ,

$$\forall u^1, w^0 \, \exists v^1 \leq tu \, \forall i \leq w \; A(u, v, i).$$

Weak König's lemma

In particular, T_i^ω proves (WKL_ε).

Corollary. If

$$T^\omega + (WKL) \vdash \forall x^1 \, \forall y^1 \leq sx \; \exists z^0 \; B(x, y, z)$$

then for some Φ

$$T_i^\omega \vdash \forall x^1 \, \forall y^1 \leq sx \; \exists z^0 < \Phi x \; B(x, y, z).$$

Similar results hold if we vary the theories on the left, e.g. with PA^ω and HA^ω.

There are even more general versions of the metatheorem, though handling extensionality is subtle.

A mathematical reading

Let X and K range over CSM's, K compact.

Theorem (Kohlenbach 1993). Let Δ be a set of closed axioms of the form

$$\forall x' \in X' \, \exists y' \in K'_{x'} \, \forall m' \in \mathbb{N} \; A(x', y', m')$$

Suppose

$$T^{\omega} + \Delta \vdash \forall n \in \mathbb{N} \, \forall x \in X \, \forall y \in K_x \, \exists m \in \mathbb{N} \; B(n, x, y, m).$$

Then there is a term Φ such that

$$T^{\omega}_{i} + \Delta_{\varepsilon} \vdash \forall n \in \mathbb{N} \, \forall x \in X \, \forall y \in K_x \, \exists m \leq \Phi(n, x) \; B(n, x, y, m)$$

where Δ_{ε} are "ε-weakenings" of formulas in Δ,

$$\forall x \in X' \, \forall m' \in \mathbb{N} \, \exists y \in K' \, \forall i \leq m' \; A(x', y', m').$$

A mathematical reading

Again, the bound doesn't depend on parameters from K_x. But note that x and y really range over *representatives* of elements of x and K_x. (Note, e.g., that every *extensional* computable functional $\Phi : \mathbb{R} \to \mathbb{N}$ is constant!)

Allowable axioms include a wide range of principles from analysis, e.g.:

1. Basic properties of continuous functions, integrals, sups, trig functions, etc.
2. The fundamental theorem of calculus.
3. The Heine-Borel theorem.
4. Uniform continuity of continous functions on a compact interval.
5. The extreme value theorem.

Principles based on *sequential compactness* are not included. But even restricted uses of these can be eliminated, in certain contexts (Kohlenbach 1998).

A mathematical reading

Let K be compact, and let $\langle a_0, a_1, \ldots \rangle$ be a countable dense sequence.

Consider, for example, the extreme value theorem for K:

$$\forall f \in C(K)\ \exists y \in K\ \forall z \in K\ (f(y) \geq f(z)).$$

This is equivalent to

$$\forall f \in C(K)\ \exists y \in K\ \forall m\ ((f(y))_m \geq (f(a_m))_m - 2^{-m+1}).$$

The ε-weakening is

$$\forall f \in C(K), m\ \exists y \in K\ \forall i < m\ ((f(y))_i \geq (f(a_i)) - 2^{-i+1}).$$

But this is easily obtained, choosing y from among $\langle a_0, \ldots, a_m \rangle$ so that $f(y)$ is within 2^{-m} of the maximum.

Overview of lectures

1. General frameworks: theories of arithmetic
2. General methods and results
3. Formalizing analysis
4. Weak König's lemma and monotone functional interpretation
5. Applications to functional analysis

Applications to functional analysis

Three types of applications to date:

- Approximation theory: extract a rate of strong uniqueness from a proof of the uniqueness of a best approximant
- Fixed points of nonexpansive mappings: extract a rate of asymptotic regularity from proof of convergence
- Uniformity results: determine independence of rates from various parameters

Applications to functional analysis

General picture: given K compact, and $F : K \to \mathbb{R}$ continuous (with parameters), want to construct solutions to

$$\exists x \in K \ (F(x) = 0),$$

Often one proceeds in two steps:

1. Construct approximate solutions, $|F(x_n)| < 2^{-n}$
2. By compactness, obtain a convergent subsequence.

Varying scenarios:

1. Solution is unique; want rate of convergence (yields effective solution)
2. Solution is not unique; solution ineffective; but want other information
3. Want to determine dependence / independence on parameters.

Chebycheff approximation

Consider $C[a, b]$ under the uniform norm,

$$\|f\| = \sup_{t \in [a,b]} |f(t)|,$$

and let $d(f, g) = \|f - g\|$. Let Y_n consist of the subspace of polynomials of degree n.

Given $f \in C[a, b]$, the function $d(f, \cdot)$ achieves its infimum on the compact set

$$K_f = \{h \in Y_n \mid \|h\| \leq 2\|f\|\},$$

so there is a $g \in Y_n$ such that

$$d(f, g) = d(f, Y_n) =_{\text{def}} \inf_{h \in Y_n} d(f, h).$$

Theorem (Chebycheff). There is a *unique* best approximant.

Chebycheff approximation

For every k, can find $g_k \in Y_n$ such that $d(f, g_k) \leq d(f, Y_n) + 2^{-k}$.

By compactness, some subsequence of $\langle g_k \rangle$ converges. By uniqueness, the sequence converges to the unique best approximant. But how fast?

Consider the statement that the best approximant is unique:

$$\forall f \in C[a, b]; g_1, g_2 \in K_f \; (\bigwedge_{i=1,2} d(f, g_i) = d(f, Y_n) \rightarrow g_1 = g_2).$$

Monotone functional interpretation produces a *rate of strong uniqueness*:

$$\forall f \in C[a, b]; g_1, g_2 \in K_f, \varepsilon > 0$$

$$(\bigwedge_{i=1,2} |d(f, g_i) - d(f, Y_n)| < \Phi(f, \varepsilon) \rightarrow d(g_1, g_2) < \varepsilon).$$

Chebycheff approximation

Monotone functional interpretation produces a *rate of strong uniqueness*:

$$\forall f \in C[a,b];\, g_1, g_2 \in K_f, \varepsilon > 0$$

$$\Big(\bigwedge_{i=1,2} |d(f,g_i) - d(f,Y_n)| < \Phi(f,\varepsilon) \to d(g_1,g_2) < \varepsilon\Big).$$

It is important that Φ does not depend on g_1, g_2.

Let g_1 be the (unknown) best approximant, h, and let g_2 be any $g \in Y_n$.

Then

$$|d(f,g) - d(f,Y_n)| < \Phi(f,\varepsilon) \to d(h,g) < \varepsilon.$$

This is the information we were after.

Chebycheff approximation

Theorem (Kohlenbach 1993). Let

$$\omega_n(\varepsilon) = \begin{cases} \min\left(\omega(\frac{\varepsilon}{2}), \frac{\varepsilon}{8n^2\lceil\frac{1}{\omega(1)}\rceil}\right)\} & \text{if } n \geq 1 \\ 1 & \text{if } n = 0 \end{cases}$$

and let

$$\Phi(\omega, n, \varepsilon) = \min\left(\varepsilon/4, \frac{\lfloor n/2\rfloor!\lceil n/2\rceil!}{2(n+1)} \cdot (\omega_n(\varepsilon/2))^n \cdot \varepsilon\right).$$

Then $\Phi(\omega, n, \varepsilon)$ is a uniform rate of strong uniqueness for the best uniform approximation from Y_n of functions f in $C[0, 1]$ with modulus of uniform continuity ω.

In other words, if g_1, g_2 are in Y_n, and $d(f, g_1)$ and $d(f, g_2)$ are within $\Phi(\omega, n, \varepsilon)$ of being optimal, then $d(g_1, g_2) < \varepsilon$.

Applications to approximation theory

Chebycheff approximation:

- Uniqueness of best approximant proved by Chebycheff (1859) (generalizes to “Haar subspaces” of $C[a, b]$).
- Ineffective proof of existence of constants of strong unicity by Newman and Shapiro (1963).
- Numerical bounds by Ko (1986) and, for arbitrary Haar subspaces, Bridges (1980/1982).
- Numerical bounds improved by Kohlenbach (1993), using proof mining.

L_1 approximation:

- Uniqueness of best approximant proved by Jackson (1921)
- Strong unicity studied by Björnestal (1975,1979) and Kroó (1981)
- First explicit bounds in all parameters obtained by Kohlenbach and Oliva (2003), using proof mining.

General observations

Applied to standard notions, monotone functional interpretation routinely provides useful enriched notions:

uniqueness	modulus of unicity rate of strong uniqueness
strict convexity	modulus of uniform convexity
extensionality	modulus of uniform continuity
monotonicity	modulus of monotonicity
pointwise convergence	uniform convergence

Fixed points

Let X be a complete metric space, $C \subseteq X$. A function $f : C \to C$ is called a *strict contraction* if, for some $\delta < 1$,

$$\|f(x) - f(y)\| \leq \delta\|x - y\|.$$

Theorem (Banach). Such a function has a unique fixed point. For any x_0, it is the limit of the sequence $x_{n+1} = f(x_n)$.

A function f is called *nonexpansive* if

$$\|f(x) - f(y)\| \leq \|x - y\|.$$

The theory of fixed points of nonexpansive mappings is more complex.

Fixed points

Difficulties:

- If C is unbounded, may have no fixed point.

  Example: $F(x) = x + 1$ on $\mathbb{R}$.

- If C is compact and convex, the Brouwer-Schauder fixed-point theorems imply fixed points exists; but they may not be unique.

  Example: $F(x) = x$ on $[0, 1]$.

- Even if the fixed point is unique, Banach iteration may not find it.

  Example: $F(x) = 1 - x$ on $[0, 1]$, with $x_0 = 0$.

Uniform convexity

A space is said to be *strictly convex* if

$$\|x\| = 1 \wedge \|y\| = 1 \wedge x \neq y \to \|\frac{1}{2}(x+y)\| < 1,$$

or, equivalently,

$$\|x\| \leq 1 \wedge \|y\| \leq 1 \wedge \|\frac{1}{2}(x+y)\| \geq 1 \to \|x-y\| = 0.$$

A space is said to be *uniformly convex* if for every $\varepsilon > 0$ there is a $\delta > 0$ such that

$$\|x\| \leq 1 \wedge \|y\| \leq 1 \wedge \|\frac{1}{2}(x+y)\| \geq 1-\delta \to \|x-y\| \leq \varepsilon.$$

A function η mapping ε to a such a $\delta = \eta(\varepsilon)$ is a *modulus of uniform convexity*.

Fixed points of nonexpansive mappings

Theorem (Browder, Goehde, Kirk 1965). Let X be a uniformly convex Banach space, let $C \subseteq X$ be closed, bounded, and convex, let $f : C \to C$ nonexpansive. Then f has a fixed point.

Theorem (Krasnokelski 1955). If f maps C into a compact subset of C, then for any x_0 in C,

$$x_{k+1} = (x_k + f(x_k))/2$$

converges to a fixed point.

Theorem (Kohlenbach 2001). The rate of convergence is, in general, not computable in f (even for $C = [0, 1]$, $x_0 = 0$).

For the Krasnokelski iteration, $\|x_n - f(x_n)\|$ is *decreasing*, and $\|x_n - f(x_n)\| \to 0$ ("asymptotic regularity of $\langle x_n \rangle$").

The best one can do is determine when $\|x_n - f(x_n)\| < \varepsilon$.

Asymptotic regularity

To determine when $\|x_n - f(x_n)\| < \varepsilon$, it suffices to extract a bound from a proof of

$$\forall k \; \exists n \; (\|x_n - f(x_n)\| < \frac{1}{k+1}).$$

Theorem (Kirk, Yanez 1990). Let X be uniformly convex with modulus η, $C \subseteq X$ nonempty, convex, with diameter $< d_C$, and $f : C \to C$ nonexpansive. Then

$$\forall x \in C \; \forall \varepsilon > 0 \; \forall k \geq h(\varepsilon, d_C) \; (\|x_k - f(x_k)\| \leq \varepsilon),$$

where $h(\varepsilon, d_C) = \left\lceil \frac{\ln(4d_C) - \ln(\varepsilon)}{\eta(\varepsilon/(d_C+1))} \right\rceil$.

Using proof mining, Kohlenbach (2000) found a more elementary proof.

A generalization

There is a strong generalization of Krasnokelski's theorem due to Ishikawa:

- Krasnokelski's theorem holds for arbitrary Banach spaces.
- Asymptotic regularity holds for arbitrary bounded convex sets C.
- More general *Krasnokelski-Mann iterations* are allowed:

$$x_{k+1} = (1 - \lambda_k)x_k + \lambda_k f(x_k),$$

  where λ_k is a sequence of elements of $[0, 1]$ satisfying certain constraints.

Kohlenbach (2000) provides explicit and effective bounds rates of asymptotic regularity, yielding strong uniformity (independence from starting point, f, and C, except for a bound d_C on the diameter).

History

Proof mining vs. the experts:

- Ishikawa (1976): no uniformity for asymptotic regularity
- Edelstein/O'Brien (1978): uniformity w.r.t. x_0 for fixed $\lambda_k = \lambda$
- Goebel/Kirk (1982): uniformity w.r.t. x_0 and f (even for X a hyperbolic space)
- Goebel/Kirk (1990): conjecture no uniformity w.r.t. C
- Baillon/Bruck (1996): full uniformity for fixed $\lambda_k = \lambda$
- Kohlenbach (2000): full uniformity
- Kirk (2000): uniformity w.r.t. x_0 and f for fixed $\lambda_k = \lambda$, for f directionally nonexpansive
- Kohlenbach/Leustean (2002): full uniformity, even for hyperbolic spaces and f directionally nonexpansive.

In most cases, proof mining provides the only explicit bounds.

From applications to new metatheorems

Observations on the specific results:

- They hold for nonseparable spaces as well.
- Uniformities obtain for *bounded* sets, not just compact ones.
- For noncompact sets, it suffices to assume existence of *approximate* fixed points.

Kohlenbach (to appear, TAMS) develops very general metatheorems that explain this.

Kohlenbach and Leustean (2003) use this analysis to obtain explicit results for directionally nonexpansive mappings in hyperbolic spaces.

Kohlenbach and Lambov (to appear) obtain explicit results for asymptotically quasi-nonexpansive mappings in uniformly convex spaces, where neither quantitative nor qualitative uniformity results were previously known.

Final remarks

Proof mining combines ideas from

- Proof theory
- Set theory
- Recursion theory and descriptive set theory
- Constructive mathematics

and brings them to bear on *specific mathematical developments*.

It is well worth supporting:

- The ratio of successes to practitioners is high.
- The methods are elegant, and satisfying.
- It brings various branches of logic and mathematics together.

Range of applications

Markets for unwinding: anywhere

nonconstructive, abstract, or infinitary

methods are used, and

concrete, computational, or numerical

information is desired.

For example:

- analytic methods in number theory
- measure-theoretic methods in combinatorics
- maximality (Zorn's lemma) in algebra

References

Proof mining per se:

- Kohlenbach, *Proof Interpretations and the Computational Content of Proofs* (on his home page)
- Kohlenbach and Oliva, “Proof mining: a systematic way of analyzing proofs in mathematics”
- Odifreddi, editor, *Kreiseliana*

Reverse and constructive mathematics:

- Simpson, *Subsystems of Second-Order Arithmetic*
- Bishop and Bridges, *Constructive Analysis*

General proof theory:

- Buss, editor, *Handbook of Proof Theory*

References

For applications to functional analysis, see the list of references on Kohlenbach's home page:

http://www.mathematik.tu-darmstadt.de/∼kohlenbach

The articles with the strongest applications to functional analysis are in e.g. *Numerical and Functional Analysis and Optimization*, *Journal of Mathematical Analysis and Applications*, *Abstract and Applied Analysis*, *Proceedings of the International Conference on Fixed Point Theory*, etc.

These only sketch the underlying logical ideas.

Articles in logic journals provide more details regarding the logical methods and metatheorems, with simple examples from functional analysis.

References

For extraction of algorithms from proofs, and constructive algebra, see, for example:

- Ulrich Berger's home page

  http://www-compsci.swan.ac.uk/∼csulrich/
- Thierry Coquand's home page

  http://www.cs.chalmers.se/∼coquand/
- Susumu Hayashi's home page

  http://www.shayashi.jp/index-english.html
- The MAP (Mathematics = Algorithms + Programs) home page

  http://map.unican.es/