Functional interpretation and the finite convergence principle

Philipp Gerhardy
Department of Mathematics
University of Oslo

Workshop on Proof Theory, Bern, Switzerland, July 9-11.

Outline

Introduction

The Finite Convergence Principle

The Finite Convergence Principle as a Lemma

Proof-theoretic Results and an Example

Conclusions and Future Work

Computational content

Proof mining is concerned with extracting the computational content of theorems and proofs.

Every true statement $\forall x \in \mathbb{N} \exists y \in \mathbb{N} A_{qf}(x, y)$ with A_{qf} decidable has a trivial computational interpretation: For a given $x \in \mathbb{N}$, find y by unbounded search.

Better: Functional interpretations allow one to obtain for proofs in e.g. Peano arithmetic a closed term t in Gödel's $\mathbf{T}$ that witnesses the theorem, i.e.$\forall x \in \mathbb{N} A_{qf}(x, tx)$.

Computational content

Combining functional interpretations with majorization, one obtains monotone functional interpretations and may extract (uniform) bounds instead of realizers:

$$\forall x \in \mathbb{N} \forall y \leq sx \exists z \leq tx A_{qf}(x, y, z).$$

With monotone functional interpretations we may treat principles that admit no computable realizers but admit computable bounds, such as weak König's Lemma.

These results extend to many formal systems based on Peano arithmetic, even interpreting e.g. dependent choice.

Computational content of $\forall\exists\forall$-statements

Example: Let $(x_n)_{n\in\mathbb{N}}$ be a sequence of real numbers, then

$$\forall k \exists n \forall m > n (|x_n - x_m| \leq 2^{-k}),$$

expresses the (Cauchy) convergence of $(x_n)_{n\in\mathbb{N}}$.

Problem: We can construct bounded monotone sequences of rationals – so-called Specker sequences – where there exists no computable rate of convergence, i.e. no bounds on $\exists n$.

Computational content of ∀∃∀-statements

On the other hand,

$$\forall k \exists n \forall m > n (|x_n - x_m| \leq 2^{-k}),$$

is classically equivalent to

$$\forall k \forall M : \mathbb{N} \to \mathbb{N} \exists n (M(n) > n \to |x_n - x_{M(n)}| \leq 2^{-k}).$$

This statement is constructively weaker and a ∀∃-statement.

This is the 'Dialectica' transform of the negative translation of the full convergence statement. For ∀∃∀-statements, this is known as the no-counterexample interpretation.

Applications of no-counterexample weakenings

In this talk, I want to discuss two applications of this weakened form of $\forall\exists\forall$-statements:

- Establishing infinitary statements ($\forall\exists\forall$-statements) by establishing their (classically equivalent, constructively weaker) no-counterexample interpretation.
- Computational interpretation (using Gödel's Dialectica interpretation) of $\forall\exists\forall$-lemmas in proofs of $\forall\exists$-theorems.

Principle of convergence for bounded monotone sequences

The principle of convergence for bounded monotone sequences, *PCM*, is the following statement:

$$\forall (x_n)_{n\in\mathbb{N}}(\forall i(x_i \le x_{i+1} \le b) \to \forall k \exists n \forall m > n(|x_n - x_m| \le 2^{-k})),$$

where $(x_n)_{n\in\mathbb{N}}$ is a sequence of real numbers.

We write $PCM(a_n)$ for the instantiation of *PCM* with a particular sequence $(a_n)_{n\in\mathbb{N}}$.

Principle of convergence for bounded monotone sequences

We may not be able to compute a bound for $\exists n$, even for a computable sequence $(a_n)_{n\in\mathbb{N}}$, e.g. Specker sequences.

But: One easily computes bounds for the (classically equivalent) no-counterexample version of PCM:

$$\forall (x_n)_{n\in\mathbb{N}}(\forall i(x_i \leq x_{i+1} \leq b) \to \\ \forall k \forall M \exists n(M(n) > n \to |x_n - x_{M(n)}| \leq 2^{-k})).$$

We call this constructively weaker form PCM^-. T.Tao has dubbed this the "finite convergence principle".

Combinatorial proofs of infinitary statements

While a full convergence statement is an infinitary statement, the no-counterexample is a finitary, combinatorial statement.

T.Tao recently established the following theorem:

Theorem (T.Tao)

Let $l \geq 1$ be an integer. Assume that $T_1, \dots T_l : X \to X$ are commuting invertible measure-preserving transformations of a measure space (X, χ, μ). Then for any $f_1, \dots, f_l \in L^\infty(X, \chi, \mu)$, the averages

$$\frac{1}{N} \sum_{n=0}^{N-1} f_1(T_1^n x) \dots f_l(T_l^n x)$$

are convergent in $L^2(X, \chi, \mu)$.

Combinatorial proofs of infinitary statements

This theorem was established by proving the following:

Theorem (T.Tao)

Let $l \geq 1$ be an integer, let $F : \mathbb{N} \to \mathbb{N}$ be a function and let $\varepsilon > 0$. Then there exists an integer $M^* > 0$ with the following property: If $P \geq 1$ and $f_1, \ldots, f_l : \mathbb{Z}_P^l \to [0,1]$ are arbitrary functions on $\mathbb{Z}_P^l$, then there exists an integer $1 \leq M \leq M^*$ such that we have "L^2 metastability"

$$\|A_N(f_1, \ldots, f_l) - A_{N'}(f_1, \ldots, f_l)\|_{L^2(\mathbb{Z}_P^l)} \leq \varepsilon$$

for all $M \leq N, N' \leq F(M)$ where we give $\mathbb{Z}_P^l$ the uniform probability measure.

This theorem has an essentially combinatorial constructive proof.

Conclusions so far

- The no-counterexample interpretation (negative translation + Dialectica) is an approximation to the computational content of a (classically provable) $\forall\exists\forall$-statement.
- The no-counterexample interpretation may have a simpler, more direct proof than the original infinitary statement.
- We may consider the no-counterexample interpretation to establish the truth of the infinitary ($\forall\exists\forall$-)statement.

$\forall\exists\forall$-statements as lemmas

Let $A \equiv \forall x \exists y \forall z P(x, y, z)$ and $B \equiv \forall u \exists v Q(u, v)$ with P, Q decidable and all variables ranging over $\mathbb{N}$.

Assume A is a lemma in a (classical) proof of B, i.e. we have a proofs of A and $A \to B$ and the final proof step

$$\frac{A \quad A \to B}{B}.$$

How do we interpret this proof using Gödel's ('Dialectica') functional interpretation (combined with negative translation)?

$\forall\exists\forall$-statements as lemmas

The Dialectica transform of the negative translation of A is just the no-counterexample version $A^- \equiv \forall x, Z \exists y P(x, y, Z(y))$.

Using negative translation and functional interpretation the proof of A is transformed into a (constructive) proof of A^-.

This is in our favour as proofs/realizers of the no-counterexample version of $\forall\exists\forall$-statements are often simpler.

$\forall\exists\forall$-statements as lemmas

Informally, the proof of $A \rightarrow B$ is transformed into a (constructive) proof of $A^- \rightarrow B$; this is a proof of B from a weaker premise A^-, but soundness of negative translation and functional interpretation guarantees that it can be done.

Combining the realizers for the proofs of A^- and $A^- \rightarrow B$, we get realizers for B, i.e. a computable term t_v such that $\forall u Q(u, t_v u)$ holds (similarly for bounds).

Proof theory vs. mathematical reality

What is the normal use of $\forall\exists\forall$-lemmas, e.g. of $PCM(a_n)$ for some bounded monotone sequence $(a_n)_{n\in\mathbb{N}}$?

Normal use: "Let $\varepsilon > 0$ be given, choose $n \in \mathbb{N}$ s.t. $|a_n - a_m| \leq \varepsilon/4$ for all $m \geq n$ . . . [use this n later in the proof]."

In "ordinary mathematics", when using e.g. PCM as a lemma there is often an implicit appeal to the axiom of choice (for arithmetical formulas).

Proof theory vs. mathematical reality

Arithmetical choice is used to derive $PCM \Rightarrow PCM^+$, where PCM^+ is

$$\forall (x_n)_{n\in\mathbb{N}}(\forall i(x_i \leq x_{i+1} \leq b) \to \exists h \forall k \forall m > h(k)(|x_{h(k)} - x_m| \leq 2^{-k})).$$

Interpreting PCM^+ (and AC_{ar}) requires bar-recursion; for the interpretation of PCM^- a primitive recursive (in the sense of Kleene) realizer suffices.

Question: Can we eliminate this appeal to the axiom of choice?

Eliminating certain instances of choice

We can eliminate certain instances of arithmetical choice (like $PCM \Rightarrow PCM^+$) without causing additional growth, i.e. without resorting to bar-recursion.

Informal idea:

- Given $(a_n)_{n\in\mathbb{N}}$ and k, we ask for an n s.t. $|a_n - a_m| \leq 2^{-k}$ for all $m \geq n$ (using PCM^+).
- Looking at the proof, we find we only use particular m, e.g. $m = n + 1000$ or all $m \in [n, n^2]$ or some m definable in the other parameters and then PCM^- suffices.

Crucial that PCM^+ is used for a sequence $(a_n)_{n\in\mathbb{N}}$ given in the other parameters of the theorem.

Definitions

- The finite types $\mathbf{T}$ are:

$$(i)\mathbb{N} \in \mathbf{T},\ (ii)\rho, \tau \in \mathbf{T} \Rightarrow \rho \to \tau \in \mathbf{T}.$$

- For $n \in \mathbb{N}$, $G_n A^\omega$ is the subsystem of arithmetic in all finite types where the growth of the definable functions corresponds to the n-th level of the Grzegorzyk hierarchy.

- A formula $\forall x \exists y \forall z A_{qf}(x, y, z)$ is monotone if

$$x' \leq x \wedge y \leq y' \wedge z' \leq z \wedge A_{qf}(x, y, z) \Rightarrow A_{qf}(x', y', z').$$

We write 0 for the type $\mathbb{N}$ and 1 for $\mathbb{N} \to \mathbb{N}$.

U.Kohlenbach, " Elimination of Skolem Functions for Monotone Formulas in Analysis"(1998)

Theorem

Let $\forall x^0 \exists y^0 \forall z^0 A_{qf}(u^1, v^\tau, qx, y, z)$ *be provably monotone in* $G_nA^\omega + QF\text{-}AC$*, then the following rule holds:*

$$G_nA^\omega + QF\text{-}AC \vdash \forall u^1 \forall v \leq_\tau tu \; \big(\exists h^1 \forall x, z A_{qf}(u, v, x, h(x), z) \rightarrow \exists w^0 B_{qf}(u, v, w)\big)$$

Then one can extract a term $\chi \in G_nA^\omega + QF\text{-}AC$ *such that*

$$G_nA_i^\omega + QF\text{-}AC \vdash \forall u^1 \forall v \leq_\tau tu \forall \Psi^* \; \big((\Psi^* \textit{ satisfies mFI of } \forall x, M^1 \exists y A_{qf}(u, v, x, y, M(y))) \rightarrow \exists w^0 \leq \chi \Psi^* u B_{qf}(u, v, w)\big)$$

U.Kohlenbach, " Elimination of Skolem Functions for Monotone Formulas in Analysis"(1998)

Informally, in G_nA^ω we can make "some" use of the function h (modulus of convergence for PCM^+), but not "too much" to replace it with the no-counterexample version.

Due to the limited growth in G_nA^ω, we can from a proof using e.g. PCM^+ read off a counterexample function M and plug this function into PCM^-.

NB: Crucial that e.g. PCM^+ is used for a sequence $(a_n)_{n\in\mathbb{N}}$ given in the other parameters of the theorem (so that M is defined in these parameters).

U.Kohlenbach, " Elimination of Skolem Functions for Monotone Formulas in Analysis" (1998)

Principles that can be treated by this result:

- Principle of convergence for bounded monotone sequences.
- Every bounded (from above) sequence of real numbers has a least upper bound.
- The Bolzano-Weierstrass property for bounded sequences in $\mathbb{R}^d$ (for any fixed d).
- The existence of the $\limsup$ for bounded sequences of real numbers.

Example: Mean Ergodic Theorem

Theorem

Let $(X, \langle\cdot,\cdot\rangle)$ be a Hilbert space, let $T : X \to X$ be nonexpansive and for $f \in X$ define $A_n f :\equiv \frac{1}{n+1}\sum_{i=0}^{n} T^i f$. Then the ergodic averages $A_n f$ converge in the Hilbert space norm.

From a classical proof, Avigad, G and Towsner extracted bounds for the no-counterexample version of the Mean Ergodic Theorem of the Mean Ergodic Theorem.

Analysing the proof included treating an instance of PCM^+ for a bounded monotone sequence $(a_n)_{n\in\mathbb{N}}$ defined in the parameters of the theorem.

Example: Mean Ergodic Theorem

Stripping away the details, the applications of PCM^+ were:

Lemma

For all $f \in X$, nonexpansive $T : X \to X$, $\varepsilon > 0$, if $(a_n)_{n\in\mathbb{N}}$ (definable in the parameters f, T) converges then $\exists n_0 A_{qf}(\ldots)$.

Proof:

- Define a $\delta > 0$ in parameters.
- By convergence of $(a_n)_{n\in\mathbb{N}}$, find an i s.t. $|a_i - a_j| \leq \delta$ for all $j > i$.
- (then in particular $|a_i - a_{i+1}| \leq \delta$).
- Use i to construct n_0.

Observation: PCM^- for $M(n) = n + 1$ suffices.

Example: Mean Ergodic Theorem

Lemma

For all $f \in X$, nonexpansive $T : X \to X$, $\varepsilon > 0$, if $(a_n)_{n\in\mathbb{N}}$ (definable in the parameters f, T) converges then $\exists n_1 B_{qf}(\ldots)$.

Proof:

- Define a $\gamma > 0$ and $\gamma' > 0$ in parameters.
- By convergence of $(a_n)_{n\in\mathbb{N}}$, find an i s.t. $|a_i - a_j| \leq \gamma$ for all $j > i$.
- Some $j > i$ will satisfy previous lemma for γ'.
- Use i to construct n_0.

Observation: PCM^- for M' constructible in previous M suffices.

Further conclusions

When using $\forall\exists\forall$-statements as lemmas, we may:

- in certain cases formulate the proof with strong versions of the lemma, e.g. PCM^+, and then
- transform this proof into a constructive(!) proof using only a weak version, e.g. PCM^-, and
- thus eliminate appeals to arithmetical choice without causing additional growth.

This has applications in proof mining of proofs in e.g. fixed point theory and ergodic theory.

Summary

The combination of Gödel's Dialectica interpretation and negative translation to interpret classical proofs allows one to:

- Replace classical proofs of $\forall\exists\forall$-statements (infinitary) with constructive, combinatorial proofs of $\forall\exists$-statements (finitary).
- Reduce use of $\forall\exists\forall$-lemmas + arithmetical choice to (their no-counterexample interpretation) $\forall\exists$-lemmas.

Functional interpretation shows when and how this can be done.

Future Work

The use of no-counterexample interpretation of infinitary statements has found applications in mainstream mathematics: Bring together proof-theoretic and mathematical results.

Establish more "natural" conditions for e.g. eliminability of choice than "provable in G_nA^ω; and more rigorous than "I can see that from the proof".

Apply proof theoretic results to analysis of actual mathematical proofs (in fixed point theory, ergodic theory, etc).

References

- U.Kohlenbach, "Applied Proof Theory: Proof Interpretations and their Use in Mathematics". Springer Monographs in Mathematics (2008).
- U.Kohlenbach, "Elimination of Skolem functions for monotone formulas in analysis". Arch. Math. Logic 37, pp. 363-390 (1998).
- Avigad, G., Towsner, "Local stability of ergodic averages". To appear in: Trans. Am. Math. Soc.
- T.Tao, "Norm convergence of multiple ergodic averages for commuting transformations", Egodic Theory and Dynamical Systems 28, pp. 657-688 (2008).