Applications of proof theory in mathematics

Thomas Powell

CARMIN postdoc

Séminaire de Mathématiques, IHÉS
14 January 2014

Theorem. There are infinitely many prime numbers.

Theorem. For all $n \in \mathbb{N}$ there exists $p \geq n$ such that $\texttt{prime}(p)$

Theorem. For all $n \in \mathbb{N}$ there exists $p \geq n$ such that $\texttt{prime}(p)$

Proof (Euclid). Suppose that $p_1, \ldots, p_m$ are the primes less than n. Then $p_1 \ldots p_m + 1$ has a prime factor p distinct from the p_i, and so in particular $p \geq n$.

Theorem. For all $n \in \mathbb{N}$ there exists $p \geq n$ such that $\texttt{prime}(p)$

Proof (Euclid). Suppose that $p_1, \ldots, p_m$ are the primes less than n. Then $p_1 \ldots p_m + 1$ has a prime factor p distinct from the p_i, and so in particular $p \geq n$.

Strengthened Theorem. For all $n \in \mathbb{N}$ there exists $p \in [n, p_1 \ldots p_m + 1]$ such that $\texttt{prime}(p)$, where $p_1, \ldots, p_m$ are the prime numbers less than n.

Theorem. For all $n \in \mathbb{N}$ there exists $p \geq n$ such that $\texttt{prime}(p)$

Proof (Euclid). Suppose that $p_1, \ldots, p_m$ are the primes less than n. Then $p_1 \ldots p_m + 1$ has a prime factor p distinct from the p_i, and so in particular $p \geq n$.

Strengthened Theorem. For all $n \in \mathbb{N}$ there exists $p \in [n, p_1 \ldots p_m + 1]$ such that $\texttt{prime}(p)$, where $p_1, \ldots, p_m$ are the prime numbers less than n.

Moral: A proof of an existential statement contains more information than simply the truth of the statement - often it comes with additional 'computational content'.

G. Kreisel, 1950s: "What more do we know if we have proved a theorem by restricted means than if we merely know that it is true"

G. Kreisel, 1950s: "What more do we know if we have proved a theorem by restricted means than if we merely know that it is true"

Proof interpretations help answer this question.

- A mapping: $A \mapsto \exists x \forall y |A|^x_y$ where $|A|^x_y$ is quantifier-free;
- A soundness proof: $\mathcal{T} \vdash A \Rightarrow \mathsf{T} \vdash \forall y |A|^t_y$ where $t \in \mathsf{T}$ can be recursively extracted from the proof of A.

G. Kreisel, 1950s: "What more do we know if we have proved a theorem by restricted means than if we merely know that it is true"

Proof interpretations help answer this question.

- A mapping: $A \mapsto \exists x \forall y |A|^x_y$ where $|A|^x_y$ is quantifier-free;
- A soundness proof: $\mathcal{T} \vdash A \Rightarrow \mathrm{T} \vdash \forall y |A|^t_y$ where $t \in \mathrm{T}$ can be recursively extracted from the proof of A.

For a logical theory $\mathcal{T}$ interpreted in a class of functionals T:

Mathematical proof of A

$\rightsquigarrow$ Formal proof of A in $\mathcal{T}$ (restricted means)

$\mapsto$ Extracted $t \in \mathrm{T}$ satisfying $\forall y |A|^t_y$ (computational content)

Metatheorem A (Gödel, 1958). If $\mathsf{PA} \vdash A$ then we can extract a (higher-type) primitive recursive functional t satisfying $\forall y |A|^t_y$.

Metatheorem A (Gödel, 1958). If PA $\vdash A$ then we can extract a (higher-type) primitive recursive functional t satisfying $\forall y |A|^t_y$.

<u>$\forall\exists$ statements</u>: $\forall n \exists m A_0(n, m)$ interpreted as $\exists f^{\mathbb{N}\to\mathbb{N}} \forall n A_0(n, fn)$

Metatheorem A'. If PA $\vdash \forall n \exists m A_0(n, m)$ then we can extract a primitive recursive function f satisfying $\forall n A_0(n, fn)$.

Metatheorem A (Gödel, 1958). If $\mathrm{PA} \vdash A$ then we can extract a (higher-type) primitive recursive functional t satisfying $\forall y |A|^t_y$.

<u>$\forall\exists$ statements</u>: $\forall n \exists m A_0(n,m)$ interpreted as $\exists f^{\mathbb{N}\to\mathbb{N}} \forall n A_0(n, fn)$

Metatheorem A'. If $\mathrm{PA} \vdash \forall n \exists m A_0(n,m)$ then we can extract a primitive recursive function f satisfying $\forall n A_0(n, fn)$.

Example. $\mathrm{PA} \vdash \forall n \exists p(\underbrace{p \geq n \wedge \mathtt{prime}(p)}_{A_0(n,p)})$

The primitive recursive function we extracted from Euclid's proof is

$$f(n) = \text{least } p \in [n, p_1 \dots p_m + 1] \text{ such that } \mathtt{prime}(p)$$

Metatheorem A (Gödel, 1958). If $\mathrm{PA} \vdash A$ then we can extract a (higher-type) primitive recursive functional t satisfying $\forall y |A|^t_y$.

<u>∀∃ statements</u>: $\forall n \exists m A_0(n,m)$ interpreted as $\exists f^{\mathbb{N}\to\mathbb{N}} \forall n A_0(n, fn)$

Metatheorem A'. If $\mathrm{PA} \vdash \forall n \exists m A_0(n,m)$ then we can extract a primitive recursive function f satisfying $\forall n A_0(n, fn)$.

Example. $\mathrm{PA} \vdash \forall n \exists p (\underbrace{p \geq n \wedge \mathtt{prime}(p)}_{A_0(n,p)})$

The primitive recursive function we extracted from Euclid's proof is

$$f(n) = \text{least } p \in [n, p_1 \dots p_m + 1] \text{ such that } \mathtt{prime}(p)$$

But Euclid's proof is (a) constructive, and (b) trivial!
Metatheorem A' doesn't tell us anything new here.

Theorem. For $f \in C[0,1]$ let $E_{n,f} := \inf_{p \in P_n} \|f - p\|_\infty$. Then for all $p_1, p_2 \in P_n$ we have

$$\bigcap_{i=1}^{2} \|f - p_i\|_\infty = E_{n,f} \rightarrow p_1 \equiv p_2.$$

Theorem. For $f \in C[0,1]$ let $E_{n,f} := \inf_{p \in P_n} \|f - p\|_\infty$. Then for all $p_1, p_2 \in P_n$ we have

$$\bigcap_{i=1}^{2} \|f - p_i\|_\infty = E_{n,f} \to p_1 \equiv p_2.$$

Standard proofs due to Young (1907) and de La Vallée Poussin (1919) can be formalised in sufficiently weak theory to guarantee extractability of Gödel primitive recursive functional Φ satisfying

$$\bigcap_{i=1}^{2} \|f - p_i\|_\infty - E_{n,f} < 2^{-\Phi(f,n,k)} \to \|p_1 - p_2\|_\infty < 2^{-k}.$$

Theorem. For $f \in C[0,1]$ let $E_{n,f} := \inf_{p \in P_n} \|f - p\|_\infty$. Then for all $p_1, p_2 \in P_n$ we have

$$\bigcap_{i=1}^{2} \|f - p_i\|_\infty = E_{n,f} \rightarrow p_1 \equiv p_2.$$

Standard proofs due to Young (1907) and de La Vallée Poussin (1919) can be formalised in sufficiently weak theory to guarantee extractability of Gödel primitive recursive functional Φ satisfying

$$\bigcap_{i=1}^{2} \|f - p_i\|_\infty - E_{n,f} < 2^{-\Phi(f,n,k)} \rightarrow \|p_1 - p_2\|_\infty < 2^{-k}.$$

Proof theoretic studies by U. Kohlenbach in the 1990s led to the formal extraction of several explicit numerical results in approximation theory that improved previously discovered bounds.

Proof interpretations reveal their power when applied to complex, non-constructive proofs, where they are used to extract numerical information ‘hidden’ in the implicit logical structure of those proofs.

This application of proof interpretations to reveal new numerical results in mathematics is the basis of the ‘proof mining’ program.

Proof interpretations reveal their power when applied to complex, non-constructive proofs, where they are used to extract numerical information 'hidden' in the implicit logical structure of those proofs.

This application of proof interpretations to reveal new numerical results in mathematics is the basis of the 'proof mining' program.

- **1940s** Proof interpretations developed for foundational purposes (relative consistency proofs).
- **1950s** Kreisel suggests reorientation of proof theory for extracting numerical information from proofs.
- **1990s** First non-trivial results obtained in numerical analysis.
- **2000-** Methods become increasingly sophisticated, have an impact in ergodic theory, combinatorics, etc. No longer restricted to 'direct' computational content.

How do we give a computational interpretation $\exists x \forall y |A|^x_y$ to A when it is a $\forall\exists\forall$-statement? *We cannot directly interpret it as*

$$\forall n \exists m \forall k A_0(n, m, k) \mapsto \exists f^{\mathbb{N}\to\mathbb{N}} \forall n, k A_0(n, f(n), k).$$

How do we give a computational interpretation $\exists x \forall y |A|^x_y$ to A when it is a $\forall\exists\forall$-statement? *We cannot directly interpret it as*

$$\forall n \exists m \forall k A_0(n, m, k) \mapsto \exists f^{\mathbb{N}\to\mathbb{N}} \forall n, k A_0(n, f(n), k).$$

Example. Cauchy convergence of some sequence (x_i) can be expressed as:

$$\forall n \exists m \forall i, j \geq m(\|x_i - x_j\| < 2^{-n}).$$

How do we give a computational interpretation $\exists x \forall y |A|^x_y$ to A when it is a $\forall\exists\forall$-statement? *We cannot directly interpret it as*

$$\forall n \exists m \forall k A_0(n, m, k) \mapsto \exists f^{\mathbb{N}\to\mathbb{N}} \forall n, k A_0(n, f(n), k).$$

Example. Cauchy convergence of some sequence (x_i) can be expressed as:

$$\forall n \exists m \forall i, j \geq m (\|x_i - x_j\| < 2^{-n}).$$

Theorem (Specker, 1949). There exists a computable, increasing sequence (x_i) of rationals in $[0, 1]$ whose rate of convergence is non-computable i.e. there is no computable function $f : \mathbb{N} \to \mathbb{N}$ satisfying

$$\forall n \forall i, j \geq f(n)(|x_i - x_j| < 2^{-n}).$$

One solution is to interpret, instead, a double negated version of the statement. Let $A \equiv \forall n \exists m \forall k A_0(n, m, k)$...

One solution is to interpret, instead, a double negated version of the statement. Let $A \equiv \forall n \exists m \forall k A_0(n, m, k)$...

1. $\neg A$ is logically equivalent to $\exists n \forall m \exists k \neg A_0(n, m, k)$.

One solution is to interpret, instead, a double negated version of the statement. Let $A \equiv \forall n \exists m \forall k A_0(n, m, k)$...

1. $\neg A$ is logically equivalent to $\exists n \forall m \exists k \neg A_0(n, m, k)$.

2. Interpreting the inner $\forall\exists$-statement, $\neg A$ is equivalent to $\exists n, f^{\mathbb{N} \to \mathbb{N}} \forall m \neg A_0(n, m, f(m))$.

One solution is to interpret, instead, a double negated version of the statement. Let $A \equiv \forall n \exists m \forall k A_0(n, m, k)$...

1. $\neg A$ is logically equivalent to $\exists n \forall m \exists k \neg A_0(n, m, k)$.

2. Interpreting the inner $\forall\exists$-statement, $\neg A$ is equivalent to $\exists n, f^{\mathbb{N}\to\mathbb{N}} \forall m \neg A_0(n, m, f(m))$.

3. Then $A \leftrightarrow \neg\neg A$ is equivalent to the $\forall\exists$ statement $\forall n, f \exists m A_0(n, m, f(m))$.

One solution is to interpret, instead, a double negated version of the statement. Let $A \equiv \forall n \exists m \forall k A_0(n, m, k)$...

1. $\neg A$ is logically equivalent to $\exists n \forall m \exists k \neg A_0(n, m, k)$.

2. Interpreting the inner $\forall\exists$-statement, $\neg A$ is equivalent to $\exists n, f^{\mathbb{N}\to\mathbb{N}} \forall m \neg A_0(n, m, f(m))$.

3. Then $A \leftrightarrow \neg\neg A$ is equivalent to the $\forall\exists$ statement $\forall n, f \exists m A_0(n, m, f(m))$.

4. This is interpreted by a higher-type functional $F \colon \mathbb{N} \times (\mathbb{N} \to \mathbb{N}) \to \mathbb{N}$ satisfying $\forall n, f A_0(n, F(n, f), f(F(n, f)))$.

One solution is to interpret, instead, a double negated version of the statement. Let $A \equiv \forall n \exists m \forall k A_0(n, m, k)$...

1. $\neg A$ is logically equivalent to $\exists n \forall m \exists k \neg A_0(n, m, k)$.

2. Interpreting the inner $\forall\exists$-statement, $\neg A$ is equivalent to $\exists n, f^{\mathbb{N}\to\mathbb{N}} \forall m \neg A_0(n, m, f(m))$.

3. Then $A \leftrightarrow \neg\neg A$ is equivalent to the $\forall\exists$ statement $\forall n, f \exists m A_0(n, m, f(m))$.

4. This is interpreted by a higher-type functional $F \colon \mathbb{N} \times (\mathbb{N} \to \mathbb{N}) \to \mathbb{N}$ satisfying $\forall n, f A_0(n, F(n, f), f(F(n, f)))$.

Formally, this interpretation corresponds to the negative translation combined with the Gödel *Dialectica* interpretation.

Theorem. All increasing sequences in $[0, 1]$ are Cauchy.

Theorem. All increasing sequences in $[0,1]$ are Cauchy.

Proof. If not then there is some number n and function f such that $\forall m(|x_{m+f(m)} - x_m| \geq 2^{-n})$. Define $\tilde{f}(a) = a + f(a)$. Then by the pigeonhole principle we must have $|x_{\tilde{f}^{(i+1)}(0)} - x_{\tilde{f}^{(i)}(0)}| < 2^{-n}$ for some $i \leq 2^n$, a contradiction.

Theorem. All increasing sequences in $[0,1]$ are Cauchy.

Proof. If not then there is some number n and function f such that $\forall m(|x_{m+f(m)} - x_m| \geq 2^{-n})$. Define $\tilde{f}(a) = a + f(a)$. Then by the pigeonhole principle we must have $|x_{\tilde{f}^{(i+1)}(0)} - x_{\tilde{f}^{(i)}(0)}| < 2^{-n}$ for some $i \leq 2^n$, a contradiction.

Interpretation (i). If (x_i) is an increasing sequence in $[0,1]$, then for all n the sequence experiences arbitrarily high-quality regions of *metastability* relative to functions $f : \mathbb{N} \to \mathbb{N}$ i.e. there is some $k \leq \tilde{f}^{(2^n)}(0)$ such that

$$|x_i - x_j| < 2^{-n}$$

for all $i, j \in [k, f(k)]$.

Theorem. All increasing sequences in $[0,1]$ are Cauchy.

Proof. If not then there is some number n and function f such that $\forall m(|x_{m+f(m)} - x_m| \geq 2^{-n})$. Define $\tilde{f}(a) = a + f(a)$. Then by the pigeonhole principle we must have $|x_{\tilde{f}^{(i+1)}(0)} - x_{\tilde{f}^{(i)}(0)}| < 2^{-n}$ for some $i \leq 2^n$, a contradiction.

Interpretation (i). If (x_i) is an increasing sequence in $[0,1]$, then for all n the sequence experiences arbitrarily high-quality regions of *metastability* relative to functions $f \colon \mathbb{N} \to \mathbb{N}$ i.e. there is some $k \leq \tilde{f}^{(2^n)}(0)$ such that

$$|x_i - x_j| < 2^{-n}$$

for all $i, j \in [k, f(k)]$.

Computational content in the form of higher-type (prim. rec.) functional $F \colon \mathbb{N} \times (\mathbb{N} \to \mathbb{N}) \to \mathbb{N}$ given by $F(n, f) = \tilde{f}^{(2^n)}(0)$.

Observation. The bound $\tilde{f}^{2^n}(0)$ is uniform over sequences (x_i) (not surprising since $[0,1]^\omega$ compact).

Observation. The bound $\tilde{f}^{2^n}(0)$ is uniform over sequences (x_i) (not surprising since $[0,1]^\omega$ compact).

Interpretation (ii). For all numbers n and functions $f : \mathbb{N} \to \mathbb{N}$ we have

$$|x_i - x_j| < 2^{-n}$$

for all $i, j \in [k, f(k)]$, where $k \leq \tilde{f}^{2^n}(0)$ and (x_i) is an arbitrary increasing sequence in $[0,1]$.

Observation. The bound $\tilde{f}^{2^n}(0)$ is uniform over sequences (x_i) (not surprising since $[0,1]^\omega$ compact).

Interpretation (ii). For all numbers n and functions $f : \mathbb{N} \to \mathbb{N}$ we have

$$|x_i - x_j| < 2^{-n}$$

for all $i, j \in [k, f(k)]$, where $k \leq \tilde{f}^{2^n}(0)$ and (x_i) is an arbitrary increasing sequence in $[0,1]$.

This kind of thing is used (independently of proof interpretations) in mathematical analysis!

Observation. The bound $\tilde{f}^{2^n}(0)$ is uniform over sequences (x_i) (not surprising since $[0,1]^\omega$ compact).

Interpretation (ii). For all numbers n and functions $f : \mathbb{N} \to \mathbb{N}$ we have

$$|x_i - x_j| < 2^{-n}$$

for all $i, j \in [k, f(k)]$, where $k \leq \tilde{f}^{2^n}(0)$ and (x_i) is an arbitrary increasing sequence in $[0,1]$.

This kind of thing is used (independently of proof interpretations) in mathematical analysis!

'Finitary convergence principle' (T. Tao). If $n \in \mathbb{N}$, f is a function $\mathbb{N} \to \mathbb{N}$ and $0 \leq x_0 \leq \ldots \leq x_M \leq 1$ for M sufficiently large depending on n and f, then there exists $0 \leq k \leq k + f(k) \leq M$ such that $|x_i - x_j| \leq 2^{-n}$ for all $i, j \in [k, f(k)]$.

'Correspondence principle' (maths) $\Leftrightarrow$ **Proof interpretation** (logic)

Infinitary or qualitative statement	$\Leftrightarrow$	$\forall/\exists$ implicitly dependent
$\downarrow$		$\downarrow$
Finitary or quantitative statement	$\Leftrightarrow$	$\forall/\exists$ explicitly dependent

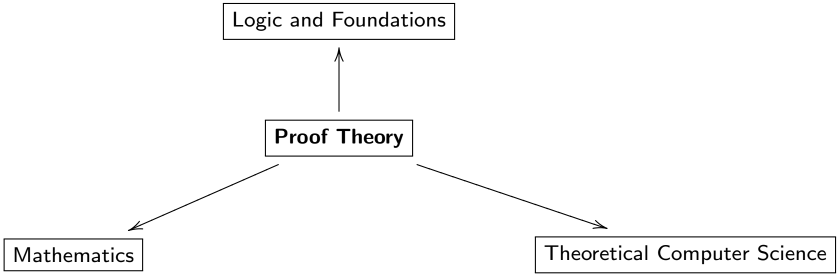
'Correspondence principle' (maths) $\Leftrightarrow$ **Proof interpretation** (logic)

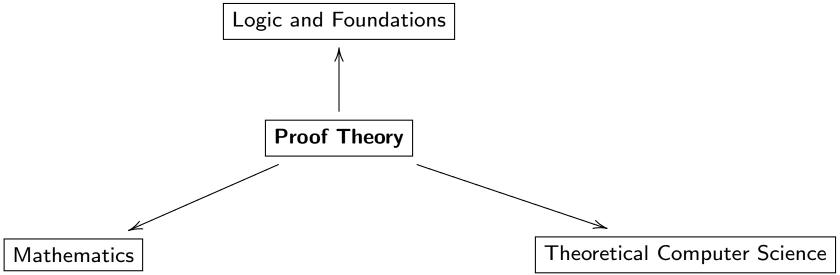
Infinitary or qualitative statement $\Leftrightarrow$ $\forall/\exists$ implicitly dependent

$\downarrow$ $\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad$ $\downarrow$

Finitary or quantitative statement $\Leftrightarrow$ $\forall/\exists$ explicitly dependent

From around 2008 onwards, there have been several new applications of proof interpretations in ergodic theory, where in particular they are used to obtain finitary convergence proofs with explicit rates of 'metastability'.


Logic and Foundations
Proof Theory
Mathematics
Theoretical Computer Science



My research on proof interpretations:

(i) Extensions of proof interpretations to strong theories of analysis that include countable choice axioms.

(ii) Trying to gain a better understanding of the mathematical, or semantic meaning of the action of proof interpretations.