Translating nonstandard proofs to constructive ones

Jeremy Avigad

Department of Philosophy

Carnegie Mellon University

avigad@cmu.edu

http://andrew.cmu.edu/∼avigad

Conservation theorems in proof theory

A *conservation theorem* is one of the following form: if T_1 proves φ for some φ in Γ, then T_2 proves φ as well (or perhaps a translation, φ').

These provide foundational reductions:

- Infinitary to finitary
- Nonconstructive to constructive
- Impredicative to predicative
- Nonstandard to standard

Kreisel's "unwinding" program: find constructive content in classical proofs.

Contemporary work in "proof mining" by Kohlenbach and students, Schwichtenberg, Berger, Coquand, Lombardi, et al.

Nonstandard analysis

Robinson (1966): Reason about saturated elementary extensions of a suitable mathematical universe

Kreisel (1969): Axiomatic nonstandard second-order and higher-order arithmetic

Friedman: Nonstandard Peano arithmetic

Nelson (1977): Axiomatic nonstandard set theory

Others have considered weaker theories, constructive theories, etc.

Nonstandard first-order arithmetic

Add to the language of first-order (Peano) arithmetic:

- a predicate, $st(x)$ ("x is standard")
- a constant, ω

Axioms of nonstandard PA:

- All the axioms of first-order arithmetic
- $\neg st(\omega)$, and $st(x) \wedge y < x \rightarrow st(y)$
- Transfer: $st(\vec{z}) \rightarrow (\varphi(\vec{z}) \leftrightarrow \varphi^{st}(\vec{z}))$ for φ in the original language
- Standard induction:

$$\varphi(0) \wedge \forall x \, (\varphi(x) \rightarrow \varphi(x+1)) \rightarrow \forall^{st} x \, \varphi(x)$$

Theorem (Friedman). *NPA* is a conservative extension of *PA*.

Note: the saturation principle

$$\forall^{st} x \, \exists y \, \varphi(x, y) \rightarrow \exists y \, \forall^{st} x \, \varphi(x, y_x)$$

raises the strength to second-order arithmetic.

A weak theory of nonstandard arithmetic

Start with *Primitive recursive arithmetic* (*PRA*):

- Defining equations for the primitive recursive functions
- Quantifier-free induction

A nonstandard version, *NPRA*:

- $\neg st(\omega)$
- $st(x) \wedge y < x \to st(y)$
- $st(x_1) \wedge \ldots \wedge st(x_k) \to st(f(x_1, \ldots, x_k))$, for each function symbol f
- A very restricted transfer principle ($\forall$ sentences without parameters)

A short model-theoretic argument shows:

Theorem (Avigad). Suppose *NPRA* proves $\forall^{st} x \, \exists y \, \varphi(x, y)$, with φ quantifier-free in the language of *PRA*. Then *PRA* proves $\forall x \, \exists y \, \varphi(x, y)$.

In particular, the conclusion holds if *NPRA* proves either $\forall x \, \exists y \, \varphi(x, y)$ or $\forall^{st} x \, \exists^{st} y \, \varphi(x, y)$.

An explicit translation

In fact, an explicit "forcing" translation interprets the nonstandard theory in a conservative extension (i.e. with variables and quantifiers ranging over functions).

- The translation is efficient.
- It extends smoothly to higher types.
- It works for weaker theories (elementary arithmetic, polynomial time computable arithmetic).
- The strongest version gives *constructive* proofs.
- Stronger transfer, saturation, and induction principles can be added "gingerly."
- Standard induction translates to ordinary induction.
- Can add Skolem functions to obtain more transfer.

Weak theories of nonstandard arithmetic

Benefits:

- Can formalize arguments in ordinary analysis
- Real numbers are type 0 objects (bounded nonstandard rationals)
- Can formalize measure theoretic arguments
- Can formalize nonstandard arguments in combinatorics, probability theory
- Weak König's lemma (compactness) holds on the standard part.

In the translation, for example:

- The standard natural numbers correspond to bounded sequences of natural numbers.
- Reals correspond to bounded sequences of rationals.
- Nonstandardly large intervals translate to sequences of arbitrarily large intervals.

Two small applications

Henry Towsner used the translation to:

1. Obtain a standard version of a nonstandard theorem by Renling Jin
2. Obtain a standard version of Wilkie's nonstandard proof of a result, due to Ajtai

The translations were fairly straightforward.

Theorem (Jin). Let U be a cut in a nonstandard model of arithmetic, with $H \notin U$. Let A and B be subsets of $\{0, 1, \ldots, H\}$. If $0 < st(|A|/H)$, and $0 < st(|B|/H)$, then $A + B$ is not U-nowhere dense.

Corollaries:

- If A and B are sequences of natural numbers with positive upper Banach density, then $A + B$ is piecewise syndetic.
- Steinhaus' theorem…

Steinhaus' theorem

Theorem (Steinhaus 1920): Let A and B be subsets of $\mathbb{R}$ with positive Lebesgue measure. Then $A + B$ includes an interval.

Corollary: If A has positive Lebesgue measure, $A - A$ includes an interval.

Steinhaus' theorem is an easy consequence of the Lebesgue density theorem, which, in turn, is usually proved using Vitali's theorem.

Find a constructive version:

- Rework Jin's argument, to make it as direct as possible.
- Translate.
- Tinker.

A constructive rewording

Without loss of generality, we can assume that A and B are compact (even subsets of $[0, 1/2]$).

Theorem. Suppose A and B are compact subsets of $[0, 1/2]$, and $A + B$ is nowhere dense. Then $\min(\mu(A), \mu(B)) = 0$.

Read:

- *Compact:* closed, and for every $\varepsilon > 0$, there is a finite ε-net.
- *Nowhere dense:* for every $(x, y) \subseteq [0, 1]$, there is a $(u, v) \subseteq (x, y)$ such that $(x, y) \cap (A + B) = \emptyset$.

An explicit proof

Lemma. Suppose n is a multiple of 4,

- $S \subseteq \{0, \ldots, n\}$
- $T \subseteq \{0, \ldots, n\}$
- $\{n, \ldots, 3n/2\} \not\subseteq S + T$

Then $|S| + |T| \leq 3n/2 + 1$.

In particular, either $|S| \leq \frac{3}{4}n$ or $|T| \leq \frac{3}{4}n$.

Proof. Suppose $z \in \{n, \ldots, 3n/2\}$, but $z \notin S + T$.

Then for every x in S, $z - x$ is not in T. So $x \mapsto z - x$ is an injection from S to $\{0, \ldots, 3n/2\} \setminus T$.

An explicit proof

For every n:

- divide $[0, 1/2]$ into 2^n subintervals.
- Find a $1/2^{n+1}$-net for A, and rationals $q_1, \ldots, q_k$ approximating these to within $1/2^{n+1}$.
- Put an "x" in each interval containing or adjacent to a q_i.

Then

- A is covered by the intervals with x's.
- If there is an x in an interval, there is a point of A in that interval or the one adjacent.

Do the same for B.

An explicit proof

Take the "sumset," based on left endpoints. So if there is an "x" in the sumset, there is a point of $A+B$ nearby.

Since $A+B$ is nowhere dense, we can find a $(u,v) \subseteq [1/2, 3/4]$ disjoint from $A+B$.

For n large enough, (u,v) will have an interval without an "x."

By the lemma, either A or B is covered by less than three quarters of the intervals.

Iterate.

Future work

The applications would be much more impressive if:

- a constructive proof, or constructive information, had been explicitly sought,
- the “unwinding” had been more difficult, making the translation-heuristic indispensible.

There are plenty of places to look for such applications: anywhere nonconstructive or analytic methods are used to obtain “concrete” results, e.g. in number theory or combinatorics.

Extra slides…

The forcing interpretation (simplest version)

Names:

- Replace the constant ω by a variable.
- Replace each variable x_i by a term $\tilde{x}_i(\omega)$.
- Replace terms $t[\omega, x_1, \ldots, x_k]$ by $t[\omega, \tilde{x}_1(\omega), \ldots, \tilde{x}_k(\omega)]$. (Call this $\widehat{t}$.)

Conditions: A condition is a unary relation $\alpha(\omega)$, satisfying

$$\forall z \, \exists \omega \geq z \; \alpha(\omega).$$

A condition α is *stronger than* β, written $\alpha \preceq \beta$, if $\forall \omega \; (\alpha(\omega) \to \beta(\omega))$.

The atomic case: Say $\alpha \Vdash t_1 = t_2$ if and only if

$$\exists z \, \forall \omega \geq z \; (\alpha(\omega) \to \widehat{t}_1 = \widehat{t}_2).$$

In other words, $\alpha \Vdash t_1 = t_2$ on all but a finite subset of α.

The forcing interpretation (continued)

The full forcing relation is defined inductively, as follows:

1. $\alpha \Vdash t_1 = t_2 \equiv \exists z\, \forall \omega \geq z\, (\alpha(\omega) \to \widehat{t}_1 = \widehat{t}_2)$.
2. $\alpha \Vdash t_1 < t_2 \equiv \exists z\, \forall \omega \geq z\, (\alpha(\omega) \to \widehat{t}_1 < \widehat{t}_2)$.
3. $\alpha \Vdash st(t) \equiv \exists z\, \forall \omega \geq z\, (\alpha(\omega) \to \widehat{t} < z)$.
4. $\alpha \Vdash \varphi \wedge \psi \equiv (\alpha \Vdash \varphi) \wedge \alpha(\Vdash \psi)$.
5. $\alpha \Vdash \varphi \to \psi \equiv \forall \beta \preceq \alpha\, (\beta \Vdash \varphi \to \beta \Vdash \psi)$.
6. $\alpha \Vdash \neg\varphi \equiv \forall \beta \preceq \alpha\, \beta \not\Vdash \varphi$
7. $\alpha \Vdash \varphi \vee \psi \equiv \forall \alpha \preceq \beta\, \exists \gamma \preceq \beta\, ((\gamma \Vdash \varphi) \vee (\gamma \Vdash \psi))$
8. $\alpha \Vdash \forall x\, \varphi \equiv \forall \tilde{x}\, (\alpha \Vdash \varphi)$
9. $\alpha \Vdash \exists x\ \equiv \forall \alpha \preceq \beta\, \exists \gamma \preceq \beta\, \exists \tilde{x}\, (\gamma \Vdash \varphi)$

Theorem 1 *If $NPRA^\omega$ proves φ, $PRA^\omega + (\Sigma_1\text{-}IND)$ proves $\Vdash \varphi$.*

The conservation theorem follows from this.

The forcing interpretation (variations)

To translate $NPRA^\omega$ to PRA^ω, take conditions to be of the form $\langle \alpha, f \rangle$ satisfying

$$\forall z \, \exists \omega \, (\alpha(\omega) \wedge f(\omega) \geq z).$$

The relation $\langle \beta, g \rangle \preceq \langle \alpha, f \rangle$ is defined by

$$\langle \beta, g \rangle \preceq \langle \alpha, f \rangle \equiv \forall \omega \, (\beta(\omega) \rightarrow \alpha(\omega) \wedge g(\omega) \leq f(\omega)),$$

Define, for example,

$$\langle \alpha, f \rangle \Vdash t_1 = t_2 \equiv \exists z \, \forall \omega \, (\alpha(\omega) \wedge f(\omega) \geq z \rightarrow \widehat{t_1} = \widehat{t_2})$$

To translate $NPRA^\omega$ to *constructive* PRA^ω, something slightly more complicated works.

Developing real analysis

Definitions in $NPRA^\omega$:

- $\mathbb{N}^*$: the nonstandard natural numbers (type N)
- $\mathbb{N}$: the standard numbers (i.e. satisfying $st(x^{\mathrm{N}})$)
- $\mathbb{Z}^*, \mathbb{Z}$: the nonstandard / standard integers
- $\mathbb{Q}^*, \mathbb{Q}$: the nonstandard / standard rationals
- $q \in \mathbb{Q}^*$ is *bounded* if $\ulcorner q \urcorner$ is standard
- q is *infinitesimal* if it is zero or $1/q$ is unbounded
- $q \sim r$ if $q - r$ is infinitesimal
- $x \in \mathbb{R}$ means that $x \in \mathbb{Q}^*$ and x is bounded
- $x =_{\mathbb{R}} y$ means $x \sim y$

In other words, we are taking $\mathbb{R}$ to be $(\mathbb{Q}^*)^{bdd} / \sim$, and dispensing with $\mathbb{R}^*$ entirely.

The advantage: reals are type 0 objects.

A surprise

A *function* $f : \mathbb{R} \to \mathbb{R}$ is a function $\mathbb{Q}^* \to \mathbb{Q}^*$ satisfying $\forall r \in \mathbb{R}\ (f(r) \in \mathbb{R}) \wedge \forall r, s \in \mathbb{R}\ (r =_{\mathbb{R}} s \to f(r) =_{\mathbb{R}} f(s))$.

Theorem 2 (*NERA*$^\omega$) *Every function $f : \mathbb{R} \to \mathbb{R}$ is continuous.*

The point: variables range over *internal* functions. The function $f \in \mathbb{Q}^* \to \mathbb{Q}^*$ defined by

$$f(x) = \begin{cases} 0 & \text{if } x \leq_{\mathbb{Q}^*} 0 \\ 1 & \text{otherwise,} \end{cases}$$

is not a function from $\mathbb{R}$ to $\mathbb{R}$: for example, $1/\omega =_{\mathbb{R}} 0$ but $f(1/\omega) \neq_{\mathbb{R}} f(0)$.

On the other hand, the function $g \in \mathbb{Q}^* \to \mathbb{Q}^*$ defined by

$$g(x) = \begin{cases} 0 & \text{if } x \leq_{\mathbb{R}} 0 \\ 1 & \text{otherwise} \end{cases}$$

is not represented by a term of *NERA*$^\omega$, since $x \leq_{\mathbb{R}} 0$ is external.

The intermediate value theorem

Theorem 3 *Suppose* $f \in [0,1] \to \mathbb{R}$, $f(0) = -1$, *and* $f(1) = 1$. *Then there is an* $x \in [0,1]$ *such that* $f(x) = 0$.

Proof. Considering f as a function on $\mathbb{Q}^*$, let

$$j = \max\{i < \omega \mid f(i/\omega) <_{\mathbb{Q}^*} 0\}$$

and let $x = j/\omega$. Since $j/\omega \sim (j+1)/\omega$, we have

$$f((j+1)/\omega) =_{\mathbb{R}} f(j/\omega) \leq_{\mathbb{R}} 0 \leq_{\mathbb{R}} f((j+1)/\omega)$$

and so $f(x) =_{\mathbb{R}} 0$.

The extreme value theorem

Theorem 4 *If* $f \in [0,1] \to \mathbb{R}$, *then* f *attains a maximum value.*

Proof. Again considering f as a function on $\mathbb{Q}^*$, let

$$y = \max_{0 \le i \le \omega} f(i/\omega),$$

let $x = j/\omega$ satisfy $f(x) =_{\mathbb{Q}^*} y$. That y is a maximum is guaranteed by the fact that for any $x' \in [0,1]$, there is an i such that $x' \sim i/\omega$.

Lebesgue measure via Löb measure

Let ω be nonstandard, and let $A \subset \mathbb{Q}^*$ be the set

$$\{0, 1/\omega, 2/\omega, \ldots, 1-2/\omega, 1-1/\omega, 1\}$$

For any internal subset $B \subseteq A$, define

$$\mu(B) = |B|/\omega.$$

Say an *external* subset E is Löb measurable if

$$\mu(E) = \inf_{B \subseteq E} \mu(B) = \sup_{B \supseteq E} \mu(B).$$

If $X \subseteq [0,1]$ (possibly external) let

$$\widehat{X} = \{q \in A \mid \exists x \in X\ (q \sim x)\}.$$

Then X is Lebesgue measurable iff $\widehat{X}$ is Löb measurable, in which case $\lambda(X) = \mu(\widehat{X})$.

Lebesgue measure in our weak theories

Let $\varphi(x)$ be any property of reals, i.e. satisfying

$$r =_{\mathbb{R}} r' \wedge \varphi(r) \rightarrow \varphi(r').$$

Let $A = \{0, 1/\omega, 2/\omega, \ldots, 1 - 2/\omega, 1 - 1/\omega, 1\}$.

Say $\lambda(\varphi) = s$ iff for every standard $\varepsilon > 0$ there are sets B and C such that

- $\forall r \in A \ (r \in B \rightarrow \varphi(r) \wedge \varphi(r) \rightarrow r \in C)$
- $|B|/\omega > s - \varepsilon$
- $|C|/\omega < s + \varepsilon$

So, for example, φ holds almost surely on $[0, 1]$ if for every standard $\varepsilon > 0$, there is a set $B \subseteq A$ such that $|B|/\omega > 1 - \varepsilon$ and $\forall r \ (r \in B \rightarrow \varphi(r))$.

A theorem by Renling Jin

Define the (upper) Banach density of $A \subseteq \mathbb{N}$:

$$BD(A) = \lim_{n \to \infty} \sup_{b-a=n} \frac{|A \cap [a,b]|}{n+1}$$

A set A is *piecewise syndetic* if for some k there are arbitrarily long sequences $a_0, \ldots, a_n$ in A with $a_{i+1} - a_i \leq k$.

Theorem. If $BD(A) > 0$ and $BD(B) > 0$, $A + B$ is piecewise syndetic.