Sheaf semantics and nonstandard intuitionistic arithmetic

Jeremy Avigad (with Jeffrey Helzner)

Carnegie Mellon University

avigad@cmu.edu

http://andrew.cmu.edu/∼avigad

Outline

Why nonstandard intuitionistic arithmetic?

Why sheaf semantics?

Conservation results

Nonstandard analysis

External (Robinson):

- Start with a classical structure (e.g. $\mathbb{N}$, $\mathbb{R}$, $V_\omega(\mathbb{R})$, a model of set theory)
- Use compactness, or an ultrapower construction, to find an elementary extension with saturation properties
- Reason about what is true in the extension
- Transfer to the original structure

Internal (Nelson):

- Start with a classical theory
- Add a predicate $\mathrm{st}(x)$, for "x is standard", and appropriate axioms
- Show the new theory is a conservative extension of the old one

Nonstandard arithmetic

Why nonstandard *arithmetic*?

- (Chauqui, Suppes, Sommer) Can carry out parts of real analysis
- (Nelson) Can carry out probability theory
- (Wilkie, Ajtai, Woods) Can carry out combinatorial arguments

Why nonstandard *intuitionistic* arithmetic?

In the nonstandard setting, many arguments have a constructive flavor.

Thesis: Nonstandard theories of intuitionistic arithmetic provides a natural framework for formalizing a number of interesting mathematical arguments.

Palmgren (BSL 99) develops intuitionistic nonstandard analysis.

Sheaf semantics

Background:

- Grothendieck: spaces of sheaves (topoi) are useful in algebraic topology, algebraic geometry, etc.
- Lawvere, Tierney: topoi provide an algebraic semantics for (higher type) intuitionsitic logic
- Joyal: this semantics can be seen as a generalization of Kripke semantics and Beth semantics

We will actually use a slight generalization of sheaf semantics, due to Palmgren 97.

Kripke semantics

Start with a first-order language L. A Kripke model consists of

- A poset
- A "universe" at each node of the poset
- An interpretation of the function symbols at each node
- An interpretation of the relation symbols at each node

where the universes are increasing, the interpretation of the function symbols agree between nodes, and the interpretation of the relations is monotone.

Kripke semantics (continued)

Truth at each node is determined by a forcing relation:

- $p \Vdash (\theta \wedge \eta)[\vec{a}]$ if and only if $p \Vdash \theta[\vec{a}]$ and $p \Vdash \eta[\vec{a}]$
- $p \Vdash (\theta \vee \eta)[\vec{a}]$ if and only if $p \Vdash \theta[\vec{a}]$ or $p \Vdash \eta[\vec{a}]$
- $p \Vdash (\theta \rightarrow \eta)[\vec{a}]$ if and only if for every $q \leq p$, if $q \Vdash \theta[\vec{a}]$, then $q \Vdash \eta[\vec{a}]$.
- $p \Vdash (\forall x\ \theta(x))[\vec{a}]$ if and only if for every $q \leq p$ and $u \in D(q)$, $p \Vdash \theta(z)[\vec{a}, b]$
- $p \Vdash (\exists x\ \theta(x))[\vec{a}]$ if and only if there is a b in $D(p)$ such that $p \Vdash \theta(z)[\vec{a}, b]$

Beth semantics

For Beth semantics:

- Make the poset a tree
- Say "$q_1, \ldots, q_k$ covers p" if every maximal path passing through p passes through one of the q_i as well
- Make the interpretation of the relations satisfy a covering condition.

The forcing definition is analogous, except for $\vee$ and $\exists$:

- $p \Vdash (\theta \vee \eta)[\vec{a}]$ if and only if there is a cover $\{q_1, \ldots, q_k\}$ of p such that for each i either $q_i \Vdash \theta[\vec{t}]$ or $q_i \Vdash \eta[\vec{a}]$
- $p \Vdash (\exists x\, \theta(x))[\vec{a}]$ if and only if there are a cover $\{q_1, \ldots, q_k\}$ of p and a sequence of elements $b_1 \in D(q_1), \ldots, b_l \in D(q_l)$, such that for each i, $q_i \Vdash \theta(z)[\vec{a}, b_i]$

Sheaf semantics

For sheaf semantics:

- Use an arbitrary category $\mathcal{C}$
- Use a basis for a Grothendieck topology (that is, a generalized notion of covering)

For standard sheaf semantics, one uses a sheaf to interpret the universe. Palmgren notes that for first-order logic, one only needs a presheaf (but the interpretation of the relations must still obey the covering condition).

Completeness

The completeness of Kripke, Beth, and sheaf semantics (and so, a fortiori, Palmgren's semantics) is well known.

For Palmgren's semantics, however, the construction is almost trivial.

Given a theory T,

- Let the objects of $\mathcal{C}$ be formulas
- An arrow $\varphi \xrightarrow{f} \psi$ is a renaming f of the variables of ψ such that $\varphi \vdash_T \psi^f$
- For the notion of covering, take the smallest basis for a Grothendieck topology satisfying:
  1. If $\varphi \vdash_T \theta \vee \eta$, then $\{\varphi \wedge \theta \to \varphi, \varphi \wedge \eta \to \psi\}$ covers ψ.
  2. If $\varphi \vdash_T \exists x\, \theta(x)$, and y is not a free variable of φ or $\exists x\, \theta(x)$, then $\{\varphi \wedge \theta(y) \to \varphi\}$ covers φ.
- Take the universe at φ to be the set of terms with free variables among those of φ.
- Interpret function symbols in the obvious way.
- Interpret R at φ by $\varphi \vdash_T R(t_1, \ldots, t_k)$.

Theorem. For every θ, $\Vdash \theta$ iff $\vdash_T \theta$.

Back to nonstandard arithmetic

Let L be the language of arithmetic, and let L^{st} be L with a new predicate symbol, $\mathrm{st}(x)$.

Nonstandard Peano arithmetic consists of the following axioms:

- All the axioms of Peano arithmetic.
- $\exists x \, \neg \mathrm{st}(x)$
- $\mathrm{st}(x_1) \wedge \ldots \wedge \mathrm{st}(x_n) \rightarrow \mathrm{st}(f(x_1, \ldots, x_n))$, for each function symbol f
- *External induction:* For each formula $\varphi(x)$ in L^{st} (possibly with other free variables),

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

- *Transfer:* For each formula φ in L with free variables $x_1, \ldots, x_n$,

  $$\mathrm{st}(x_1) \wedge \ldots \wedge \mathrm{st}(x_n) \rightarrow (\varphi \leftrightarrow \varphi^{\mathrm{st}})$$

Theorem: (Friedman, unpublished, c. 1967) Nonstandard *PA* is a conservative extension of *PA*.

Nonstandard Heyting arithmetic

Take nonstandard Heyting arithmetic, HAI, to be given by the following axioms:

- All the axioms of HA
- $\mathrm{st}(x_1) \wedge \ldots \wedge \mathrm{st}(x_n) \to \mathrm{st}(f(x_1, \ldots, x_n))$ for each function symbol f
- $\neg\neg\mathrm{st}(x) \to \mathrm{st}(x)$
- *External induction:* for each formula $\varphi(x)$ of L^{st},

  $$\varphi(0) \wedge \forall^{\mathrm{st}} x\ (\varphi(x) \to \varphi(x+1)) \to \forall^{\mathrm{st}} x\ \varphi(x).$$

- *Overspill:* for each formula $\varphi(x)$ of L,

  $$\forall^{\mathrm{st}} x\ \varphi(x) \to \exists x\ (\neg\mathrm{st}(x) \wedge \varphi(x)).$$

- *Underspill:* for each formula $\varphi(x)$ of L,

  $$\forall x\ (\neg\mathrm{st}(x) \to \varphi(x)) \to \exists^{\mathrm{st}} x\ \varphi(x).$$

Theorem (Moerdijk and Palmgren 97) HAI is a conservative extension of HA. Also, the transfer principles imply the law of the excluded middle.

History

- Palmgren (95), and also Coquand and Smith (96), obtain conservation results for weaker nonstandard versions of HA
- Moerdijk (95): Presents a nonstandard model of arithmetic, using a sheaf construction over a category of filters
- Moerdijk and Palmgren (97): Obtains the conservation result, using a category of provable filter bases
- Avigad and Helzner:
  - A slight modification of the completeness proof above yields the same result
  - Internalizing the argument yields an additional transfer rule
  - This transfer rule is optimal
- Butz independently presents another proof of the M-P conservation result.

A nonstandard model

Modifying the completeness proof:

- Use types Γ instead of formulas φ
- For the notion of covering, take the smallest basis for a Grothendieck topology satisfying the following:
  1. If $\Gamma \vdash \theta \vee \eta$, then $\{\Gamma \cup \{\theta\} \to \Gamma, \Gamma \cup \{\eta\} \to \Gamma\}$ covers Γ.
  2. If $\Gamma \vdash \exists x\, \theta(x)$, and y is not free in Γ or $\exists x\, \theta(x)$, then

     $$\Gamma \cup \{\theta(y)\} \cup \{y \geq \bar{n} \mid \Gamma \vdash_T \theta(\bar{n})\}$$

     covers Γ.
- Interpret st at the node Γ by the relation

  $$\exists n\, (\Gamma \vdash_T t \leq \bar{n}).$$

Theorem. For θ in L, $\Vdash \theta$ iff $\vdash_{HA} \theta$.

Theorem. Each axiom of HAI is forced.

Transfer principles

Positive results:

- If HAI proves φ, φ in L, then HAI proves φ^{st}
- If HAI proves φ^{st}, φ in L negative, then HA proves φ
- If HAI proves $\forall^{\mathrm{st}} x\, \varphi$, φ in L, then HA proves φ.
- If HAI proves φ^{st}, φ in L and Π_2, then HA proves φ.

Negative results: there are primitive recursive $A(x)$, $B(x)$, $C(x)$, $D(x)$, and $E(x, y)$ such that

- $HAI + \exists x\, A(x) \to \exists^{\mathrm{st}} x\, A(x)$ is not conservative over HA
- $HAI + \forall^{\mathrm{st}} x\, B(x) \to \forall x\, B(x)$ is not conservative over HA
- HAI proves $\forall^{\mathrm{st}} x\, C(x) \vee \forall^{\mathrm{st}} x\, D(x)$, but HA does not prove $\forall x\, C(x) \vee \forall x\, D(x)$
- HAI proves $\exists^{\mathrm{st}} x\, \forall y\, E(x, y)$ but HA does not prove $\exists x\, \forall y\, E(x, y)$

Corollary: HAI has neither the existence property nor the disjunction property.

Future work

Things to do:

- Find nicer translations for classical theories.
- Obtain the right conservation results for weak theories.
- See what kinds of mathematics can be developed in these theories.