Extracting Computer Programs from Nonstandard Proofs

Chuangjie Xu (j.w.w. Sam Sanders)

Ludwig-Maximilians-Universität München

Workshop Logic and its Applications

12 Oct 2016, Mathematisches Institut der Universität München

Heyting Arithmetic with finite types HA^ω

Term language T:

Simply typed lambda calculus (or SKI) + natural numbers and recursor

Logic language:

Intuitionistic logic + arithmetic axioms (incl. the induction axiom)

- Equality of natural numbers only ($I\text{-}HA^\omega$)
  so that its Dialectica interpretation is sound
- Can be embedded as inductive datatypes within dependent type theory

The Dialectica interpretation of HA^ω

Idea: Each formula A is interpreted as $\exists x \forall y |A|^x_y$ where $|A|^x_y$ is quantifier-free.

For a formal (type-theoretic) development[1],

- we firstly calculate the types $\tau^+(A)$ and $\tau^-(A)$ by induction on A,
- for each A, $x^{\tau^+(A)}$, $y^{\tau^-(A)}$, define a quantifier-free formula $|A|^x_y$, and then
- prove the soundness of the interpretation

$$\mathsf{D}(A) \;:\equiv\; \exists x^{\tau^+(A)} \forall y^{\tau^-(A)} |A|^x_y$$

by induction on (the length of) derivations.

Characterisation:

$$\mathsf{AC} + \mathsf{IP} + \mathsf{MP} \;\vdash\; A \leftrightarrow \mathsf{D}(A)$$

[1] H. Schwichtenberg and S.S. Wainer. Proofs and computations. Perspectives in Logic. Cambridge University Press, 2012.

An Agda implementation

`http://cj-xu.github.io/agda/dialectica/Dialectica.html`

- Instead of T, we work with a (simply-typed) subset of Agda.
- Predicates are represented as functions, and hence
- realisers and counterexamples of $\forall, \exists$ are dependently-typed.
- One can directly extract Agda terms from proofs in HA^{ω}.

Our goal

A constructive fragment of Nelson's Internal Set Theory, based on HA^ω, and a nonstandard variant of the Dialectica interpretation have been introduced in

> *Benno van den Berg, Eyvind Briseid, and Pavol Safarik, A functional interpretation for nonstandard arithmetic, Annals of Pure and Applied Logic 163 (2012), no. 12, 1962–1994.*

We adapt our previous development for the above to

extract computer programs from nonstandard proofs.

A nonstandard arithmetic – System H

Term language T^*: T + finite sequences σ^*

to simulate finite sets for formulating the nonstandard axioms

$\mathsf{HA}^{\omega *} :\equiv \mathsf{HA}^\omega$ + axioms for finite sequences

$\mathsf{HA}^{\omega *}_{\mathsf{st}} :\equiv \mathsf{HA}^{\omega *}$ + st predicate + axioms for st + external induction principle

$$\Phi(0) \land \forall^{\mathrm{st}} n\ (\Phi(n) \Rightarrow \Phi(\mathsf{s}n)) \Rightarrow \forall^{\mathrm{st}} n\ \Phi(n)$$

We add $\forall^{\mathrm{st}}, \exists^{\mathrm{st}}$ and axioms $\forall^{\mathrm{st}} x A \Leftrightarrow \forall x(\mathsf{st}(x) \Rightarrow A),\ \exists^{\mathrm{st}} x A \Leftrightarrow \exists x(\mathsf{st}(x) \land A)$

System H $:\equiv \mathsf{HA}^{\omega *}_{\mathsf{st}}$ + 5 nonstandard axioms (characterisation of Dialectica)

A nonstandard variant of the Dialectica interpretation

Idea: Each formula $\Phi(\underline{a})$ in $\mathsf{HA}^{\omega *}_{\mathrm{st}}$ is interpreted as $\exists^{\mathrm{st}}\underline{x}\forall^{\mathrm{st}}\underline{y}\,\varphi_{\mathsf{D}_{\mathrm{st}}}(\underline{a},\underline{x},\underline{y})$ where $\underline{x}$ is a finite list of potential realisers, and $\varphi_{\mathsf{D}_{\mathrm{st}}}(\underline{a},\underline{x},\underline{y})$ is internal.

In van den Berg *et al.*, it is (informally) defined as follows

(i) $\varphi(\underline{a})^{D_{\mathrm{st}}} :\equiv \varphi_{D_{\mathrm{st}}}(\underline{a}) :\equiv \varphi(\underline{a})$ for internal atomic formulas $\varphi(\underline{a})$,
(ii) $\mathrm{st}^{\sigma}(u^{\sigma})^{D_{\mathrm{st}}} :\equiv \exists^{\mathrm{st}}x^{\sigma *}\ u \in_{\sigma} x$.

Let $\Phi(\underline{a})^{D_{\mathrm{st}}} \equiv \exists^{\mathrm{st}}\underline{x}\ \forall^{\mathrm{st}}\underline{y}\ \varphi_{D_{\mathrm{st}}}(\underline{x},\underline{y},\underline{a})$ and $\Psi(\underline{b})^{D_{\mathrm{st}}} \equiv \exists^{\mathrm{st}}\underline{u}\ \forall^{\mathrm{st}}\underline{v}\ \psi_{D_{\mathrm{st}}}(\underline{u},\underline{v},\underline{b})$. Then

(iii) $(\Phi(\underline{a}) \wedge \Psi(\underline{b}))^{D_{\mathrm{st}}} :\equiv \exists^{\mathrm{st}}\underline{x},\underline{u}\ \forall^{\mathrm{st}}\underline{y},\underline{v}\ (\varphi_{D_{\mathrm{st}}}(\underline{x},\underline{y},\underline{a}) \wedge \psi_{D_{\mathrm{st}}}(\underline{u},\underline{v},\underline{b}))$,
(iv) $(\Phi(\underline{a}) \vee \Psi(\underline{b}))^{D_{\mathrm{st}}} :\equiv \exists^{\mathrm{st}}\underline{x},\underline{u}\ \forall^{\mathrm{st}}\underline{y},\underline{v}\ (\varphi_{D_{\mathrm{st}}}(\underline{x},\underline{y},\underline{a}) \vee \psi_{D_{\mathrm{st}}}(\underline{u},\underline{v},\underline{b}))$,
(v) $(\Phi(\underline{a}) \to \Psi(\underline{b}))^{D_{\mathrm{st}}} :\equiv \exists^{\mathrm{st}}\underline{U},\underline{Y}\ \forall^{\mathrm{st}}\underline{x},\underline{v}\ (\forall \underline{y} \in \underline{Y}[\underline{x},\underline{v}]\ \varphi_{D_{\mathrm{st}}}(\underline{x},\underline{y},\underline{a}) \to \psi_{D_{\mathrm{st}}}(\underline{U}[\underline{x}],\underline{v},\underline{b}))$.

Let $\Phi(z,\underline{a})^{D_{\mathrm{st}}} \equiv \exists^{\mathrm{st}}\underline{x}\ \forall^{\mathrm{st}}\underline{y}\ \varphi_{D_{\mathrm{st}}}(\underline{x},\underline{y},z,\underline{a})$, with the free variable z not occurring among the $\underline{a}$. Then

(vi) $(\forall z\ \Phi(z,\underline{a}))^{D_{\mathrm{st}}} :\equiv \exists^{\mathrm{st}}\underline{x}\ \forall^{\mathrm{st}}\underline{y}\ \forall z\ \varphi_{D_{\mathrm{st}}}(\underline{x},\underline{y},z,\underline{a})$,
(vii) $(\exists z\ \Phi(z,\underline{a}))^{D_{\mathrm{st}}} :\equiv \exists^{\mathrm{st}}\underline{x}\ \forall^{\mathrm{st}}\underline{y}\ \exists z\ \forall \underline{y'} \in \underline{y}\ \varphi_{D_{\mathrm{st}}}(\underline{x},\underline{y'},z,\underline{a})$,
(viii) $(\forall^{\mathrm{st}}z\ \Phi(z,\underline{a}))^{D_{\mathrm{st}}} :\equiv \exists^{\mathrm{st}}\underline{X}\ \forall^{\mathrm{st}}z,\underline{y}\ \varphi_{D_{\mathrm{st}}}(\underline{X}[z],\underline{y},z,\underline{a})$,
(ix) $(\exists^{\mathrm{st}}z\ \Phi(z,\underline{a}))^{D_{\mathrm{st}}} :\equiv \exists^{\mathrm{st}}\underline{x},z\ \forall^{\mathrm{st}}\underline{y}\ \exists z' \in z\ \forall \underline{y'} \in \underline{y}\ \varphi_{D_{\mathrm{st}}}(\underline{x},\underline{y'},z',\underline{a})$.

Types of realisers and counterexamples

For a formal development, we calculate the types $\mathsf{d}^+(A)$ of (actual) realisers and $\mathsf{d}^-(A)$ of counterexamples for formula A:

$$\mathsf{d}^+(a =_\sigma b) :\equiv \mathbb{1}$$
$$\mathsf{d}^+(\mathsf{st}^\sigma(t)) :\equiv \sigma$$
$$\mathsf{d}^+(A \wedge B) :\equiv \mathsf{d}^+A \times \mathsf{d}^+B$$
$$\mathsf{d}^+(A \vee B) :\equiv \mathsf{d}^+A \times \mathsf{d}^+B$$
$$\mathsf{d}^+(A \Rightarrow B) :\equiv ((\mathsf{d}^+A)^* \to (\mathsf{d}^+B)^*) \times ((\mathsf{d}^+A)^* \to \mathsf{d}^-B \to (\mathsf{d}^-A)^*)$$
$$\mathsf{d}^+(\forall x^\sigma A) :\equiv \mathsf{d}^+A$$
$$\mathsf{d}^+(\exists x^\sigma A) :\equiv \mathsf{d}^+A$$
$$\mathsf{d}^+(\forall^{\mathrm{st}} x^\sigma A) :\equiv \sigma \to (\mathsf{d}^+A)^*$$
$$\mathsf{d}^+(\exists^{\mathrm{st}} x^\sigma A) :\equiv \sigma \times \mathsf{d}^+A$$

$$\mathsf{d}^-(a =_\sigma b) :\equiv \mathbb{1}$$
$$\mathsf{d}^-(\mathsf{st}(t)) :\equiv \mathbb{1}$$
$$\mathsf{d}^-(A \wedge B) :\equiv \mathsf{d}^-A \times \mathsf{d}^-B$$
$$\mathsf{d}^-(A \vee B) :\equiv \mathsf{d}^-A \times \mathsf{d}^-B$$
$$\mathsf{d}^-(A \Rightarrow B) :\equiv (\mathsf{d}^+A)^* \times \mathsf{d}^-B$$
$$\mathsf{d}^-(\forall x^\sigma A) :\equiv \mathsf{d}^-A$$
$$\mathsf{d}^-(\exists x^\sigma A) :\equiv (\mathsf{d}^-A)^*$$
$$\mathsf{d}^-(\forall^{\mathrm{st}} x^\sigma A) :\equiv \sigma \times \mathsf{d}^-A$$
$$\mathsf{d}^-(\exists^{\mathrm{st}} x^\sigma A) :\equiv (\mathsf{d}^-A)^*$$

- Compare to the original Dialectica interpretation ($\mathsf{st}, \forall^{\mathrm{st}}, \exists^{\mathrm{st}}, *$)
- Variables quantified by $\forall, \exists$ have no computational content

Our formulation of the nonstandard Dialectica interpretation

$$\text{(i)}\quad |a =_\sigma b|^r_u \;:\equiv\; a =_\sigma b$$

$$\text{(ii)}\quad |\mathsf{st}^\sigma(t)|^r_u \;:\equiv\; t \in_\sigma r$$

$$\text{(iii)}\quad |A \wedge B|^r_u \;:\equiv\; |A|^{r^1}_{u_1} \wedge |B|^{r^2}_{u_2}$$

$$\text{(iv)}\quad |A \vee B|^r_u \;:\equiv\; |A|^{r^1}_{u_1} \vee |B|^{r^2}_{u_2}$$

$$\text{(v)}\quad |A \Rightarrow B|^W_{r,v} \;:\equiv\; \forall u \in W^2[r,v]\, |A|^r_u \Rightarrow |B|^{W^1[r]}_v$$

$$\text{(vi)}\quad |\forall z^\sigma \Phi(z)|^r_u \;:\equiv\; \forall z^\sigma |\Phi(z)|^r_u$$

$$\text{(vii)}\quad |\exists z^\sigma \Phi(z)|^r_u \;:\equiv\; \exists z^\sigma \forall v \in u\, |\Phi(z)|^r_v$$

$$\text{(viii)}\quad |\forall^{\mathrm{st}} z^\sigma \Phi(z)|^R_{a,u} \;:\equiv\; |\Phi(a)|^{R[a]}_u$$

$$\text{(ix)}\quad |\exists^{\mathrm{st}} z^\sigma \Phi(z)|^r_u \;:\equiv\; \exists z \in r^1 \forall v \in u\, |\Phi(z)|^{r^2}_v$$

The nonstandard Dialectica interpretation $\mathsf{D}_{\mathsf{st}}(\Phi)$ of a formula Φ is defined by

$$\mathsf{D}_{\mathsf{st}}(\Phi) \;:\equiv\; \exists^{\mathrm{st}} r^{(\mathsf{d}^+\Phi)^*}\, \forall^{\mathrm{st}} u^{\mathsf{d}^-\Phi}\, |\Phi|^r_u.$$

System H within Agda

`http://cj-xu.github.io/agda/nonstandard_dislectica/H.html`

- As 4 dependently typed, inductive definitions
- Hilbert style rather than natural deduction (to be discussed later)
- Instead of implementing an evaluator/normaliser, we embed T^* into Agda.

Extracting Agda terms from proofs in H

`http://cj-xu.github.io/agda/nonstandard_dislectica/Dialectica.html`

- We implement only the term extraction algorithm behind the soundness proof of the nonstandard Dialectica interpretation.
- Using the embedding function, we obtain an Agda program from an extracted T^*-term.
- The complete soundness proof is too complicated to implement within intensional type theory (to be discussed later).

Decidability of atomic formulas

In the original Dialectica interpretation of HA^ω, decidability of atomic formulas is necessary for realising the contraction axiom $A \Rightarrow A \wedge A$.

In the D_{st}-interpretation of H, translated formulas $|A|^r_u$ may contain quantifiers and hence may not be decidable. However, decidability is not needed to realise the contraction axiom as follows

$A \Rightarrow A \wedge A$

We define

$$U :\equiv \lambda r.\, ((r_0, r_0) :: \cdots :: (r_{|r|-1}, r_{|r|-1}) :: []) \;:\; (\mathsf{d}^+A)^* \to (\mathsf{d}^+A \times \mathsf{d}^+A)^*$$

$$Y :\equiv \lambda r. \lambda v.\, (v_0 :: v_1 :: []) \;:\; (\mathsf{d}^+A)^* \to (\mathsf{d}^-A \times \mathsf{d}^-A) \to (\mathsf{d}^-A)^*.$$

For any $(r, v) : (\mathsf{d}^+A)^* \times (\mathsf{d}^-A \times \mathsf{d}^-A)$, we have

$$|A \Rightarrow A \wedge A|^{[U,Y]}_{r,v} \quad = \quad \forall u \in (v_0 :: v_1 :: [])\, |A|^r_u \Rightarrow |A|^r_{v_0} \wedge |A|^r_{v_1}.$$

Difficulty of the implementation in intensional type theory

In intensional type theory, for arbitrary formula Φ, we have

$$\mathsf{d}^{+/-}(\Phi) = \mathsf{d}^{+/-}(\Phi[x := t])$$

only up to identity type (similar to case $\Pi(n, m{:}\mathbb{N}).\, n + m = m + n$).

Then, given $r : \mathsf{d}^{+}(\Phi)$ we have to transport it along the above equality/path to get an element of $\mathsf{d}^{+}(\Phi[x := t])$.

Since the above equality is (implicitly) used in many places of the soundness proof, we have to apply the transport function many times and also need a few lemmas about those transported terms.

This makes proving the soundness theorem very difficult and the resulting proof unreadable.

Notice that for a closed/concrete formula Φ the above equations hold judgementally and hence no transport is needed.

Realisers of the nonstandard axioms

Nonstandard axioms are realised by "identity", *e.g.*

$$\text{(NCR)}\quad \forall y^\tau \exists^{\mathrm{st}} x^\sigma \Phi(x,y) \Rightarrow \exists^{\mathrm{st}} xs^{\sigma^*} \forall y^\tau \exists x \in_\sigma xs\, \Phi(x,y)$$

$$\mathsf{d}^+(\forall y^\tau \exists^{\mathrm{st}} x^\sigma \Phi(x,y)) = \sigma \times \mathsf{d}^+\Phi$$

$$\mathsf{d}^-(\forall y^\tau \exists^{\mathrm{st}} x^\sigma \Phi(x,y)) = (\mathsf{d}^-\Phi)^*$$

$$\mathsf{d}^+(\exists^{\mathrm{st}} xs^{\sigma^*} \forall y^\tau \exists x \in_\sigma xs\, \Phi(x,y)) = \sigma^* \times \mathsf{d}^+\Phi$$

$$\mathsf{d}^-(\exists^{\mathrm{st}} xs^{\sigma^*} \forall y^\tau \exists x \in_\sigma xs\, \Phi(x,y)) = ((\mathsf{d}^-\Phi)^*)^*$$

$$\mathsf{d}^+(\mathsf{NCR}) = ((\sigma \times \mathsf{d}^+\Phi)^* \to (\sigma^* \times \mathsf{d}^+\Phi)^*) \times ((\sigma \times \mathsf{d}^+\Phi)^* \to ((\mathsf{d}^-\Phi)^*)^* \to ((\mathsf{d}^-\Phi)^*)^*)$$

$$\mathsf{d}^-(\mathsf{NCR}) = (\sigma \times \mathsf{d}^+\Phi)^* \times ((\mathsf{d}^-\Phi)^*)^*$$

We define

$$U :\equiv \lambda r.((r^1, r_0^2) :: (r^1, r_1^2) :: \cdots :: (r^1, r_{|r|-1}^2) :: []) \;:\; (\sigma \times \mathsf{d}^+\Phi)^* \to (\sigma^* \times \mathsf{d}^+\Phi)^*$$

$$Y :\equiv \lambda r.\lambda v.v \;:\; (\sigma \times \mathsf{d}^+\Phi)^* \to ((\mathsf{d}^-\Phi)^*)^* \to ((\mathsf{d}^-\Phi)^*)^*$$

where we write r_i^2 for the ith element in the sequence $r^2 : (\mathsf{d}^+\Phi)^*$, and have

$$(U(r))^1 = [(r^1)^{|r|}] \qquad (U(r))^2 = r^2.$$

For any $(r, v) : (\sigma \times \mathsf{d}^+\Phi)^* \times ((\mathsf{d}^-\Phi)^*)^*$, we have

$$|\mathsf{NCR}|_{r,v}^{[(U,Y)]} \quad = \quad \forall u \in v \forall y^\tau \exists x \in r^1 \forall w \in u |\Phi(x,y)|_w^{r^2} \Rightarrow \exists xs \in [(r^1)^{|r|}] \forall u \in v \forall y^\tau \exists x \in xs \forall w \in u |\Phi(x,y)|_w^{r^2}.$$

Summary

Work done:

- We reformulate the nonstandard Dialectica interpretation in a way that is suitable for a type-theoretic development.
- Then we embed H within the Agda proof assistant and implement the term extraction algorithm behind the nonstandard Dialectica interpretation.

Work to do:

- Add examples and real numbers to the implementation.
- Adapt the development for natural deduction system.
- Complete the formalisation of the soundness proof.
- Generalise our development to other interpretations.