Unifying functional interpretations of nonstandard/uniform arithmetic

Chuangjie Xu

Ludwig-Maximilians-Universität München

Special session on Proof Theory and Constructivism at Logic Colloquium 2018
Udine, Italy, 24th July 2018

Motivation: computational content of mathematical proofs

- Efficiency of program extraction
  Observation: shorter proof $\Rightarrow$ faster extraction & simpler term
  Proofs in Nonstandard Analysis are usually shorter.
- Scope of mathematics to extract
  We want to extract computational content from more mathematics
  Program extraction of classical Nonstandard Analysis has a large scope[1].
- Computer implementation/formalisation
  Goals: verified proofs & efficient programs

[1] S. Sanders. *The computational content of Nonstandard Analysis*, in Proceedings CL&C 2016, arXiv:1606.05820, 2016.

In this talk, we

- Reformulate van den Berg *et al.*'s Herbrand functional interpretations[2] for nonstandard arithmetic in a way that is suitable for a type-theoretic development.
- Introduce a parametrised functional interpretation, following Oliva[3]
  - unifying both the Herbrand functional interpretations (for nonstandard arithmetic) as well as the usual ones (for uniform Heyting arithmetic[4])
  - with a single, parametrised soundness proof (and term extraction algorithm).
- Implement it in the Agda proof assistant using Agda's parameterised module system (and rewriting).

[2] B. van den Berg, E. Briseid, and P. Safarik, *A functional interpretation for nonstandard arithmetic*, Annals of Pure and Applied Logic 163 (2012), no. 12, 1962–1994.

[3] P. Oliva, *Unifying functional interpretations*, Notre Dame J. Formal Logic 47 (2006), no. 2, 263–290.

[4] U. Berger, *Uniform Heyting arithmetic*, Annals of Pure and Applied Logic 133 (2005), no. 1, 125–148.

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-HA^ω)
  so that its Dialectica interpretation is sound
- Can be embedded as 4 inductive datatypes within dependent type theory

A constructive system of nonstandard arithmetic

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

$$\Phi(0) \wedge \forall^{\mathrm{st}} n\ (\Phi(n) \to \Phi(\mathsf{s}n)) \ \to\ \forall^{\mathrm{st}} n\ \Phi(n)$$

We add $\forall^{\mathrm{st}}, \exists^{\mathrm{st}}$ and axioms $\forall^{\mathrm{st}} x A \leftrightarrow \forall x(\mathsf{st}(x) \to A)$, $\exists^{\mathrm{st}} x A \leftrightarrow \exists x(\mathsf{st}(x) \wedge A)$

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

Herbrand Dialectica interpretation

Idea: Each formula $\Phi(\underline{a})$ in $\mathsf{HA}^{\omega*}_{\mathsf{st}}$ is interpreted as $\exists^{\mathsf{st}}\underline{x}\,\forall^{\mathsf{st}}\underline{y}\,\varphi_{\mathsf{D_{st}}}(\underline{a},\underline{x},\underline{y})$ where $\underline{x}$ is a finite sequence of potential realisers, and $\varphi_{\mathsf{D_{st}}}(\underline{a},\underline{x},\underline{y})$ 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 (type-theoretic) development, we calculate the types $\mathsf{d}^+\Phi$ of (actual) realisers and $\mathsf{d}^-\Phi$ of counterexamples for formula Φ:

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

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

Our formulation of the Herbrand Dialectica interpretation

For every formula Φ and terms $r : (\mathsf{d}^+\Phi)^*$ and $u : \mathsf{d}^-\Phi$, we define an internal formula $\Phi_{\mathsf{D_{st}}}(r, u)$ by induction on Φ:

$$
\begin{aligned}
(a =_\sigma b)_{\mathsf{D_{st}}}(r,u) &:\equiv a =_\sigma b \\
(\mathsf{st}^\sigma(t))_{\mathsf{D_{st}}}(r,u) &:\equiv t \in_\sigma r \\
(A \wedge B)_{\mathsf{D_{st}}}(r,(u,v)) &:\equiv A_{\mathsf{D_{st}}}(r_1,u) \wedge B_{\mathsf{D_{st}}}(r_2,v) \\
(A \vee B)_{\mathsf{D_{st}}}(r,(u,v)) &:\equiv A_{\mathsf{D_{st}}}(r_1,u) \vee B_{\mathsf{D_{st}}}(r_2,v) \\
(A \to B)_{\mathsf{D_{st}}}(W,(r,v)) &:\equiv \forall u \in W_2[r,v]\, A_{\mathsf{D_{st}}}(r,u) \to B_{\mathsf{D_{st}}}(W_1[r],u) \\
(\forall z^\sigma \Phi(z))_{\mathsf{D_{st}}}(r,u) &:\equiv \forall z^\sigma (\Phi(z))_{\mathsf{D_{st}}}(r,u) \\
(\exists z^\sigma \Phi(z))_{\mathsf{D_{st}}}(r,u) &:\equiv \exists z^\sigma \forall v \in u\, (\Phi(z))_{\mathsf{D_{st}}}(r,v) \\
(\forall^{\mathsf{st}} z^\sigma \Phi(z))_{\mathsf{D_{st}}}(R,(a,u)) &:\equiv (\Phi(a))_{\mathsf{D_{st}}}(R[a],u) \\
(\exists^{\mathsf{st}} z^\sigma \Phi(z))_{\mathsf{D_{st}}}(r,u) &:\equiv \exists z \in r_1 \forall v \in u\, (\Phi(z))_{\mathsf{D_{st}}}(r_2,v)
\end{aligned}
$$

The Herbrand Dialectica interpretation $\Phi^{\mathsf{D_{st}}}$ of a formula Φ is defined by

$$\Phi^{\mathsf{D_{st}}} :\equiv \exists^{\mathsf{st}} x^{(\mathsf{d}^+\Phi)^*} \forall^{\mathsf{st}} y^{\mathsf{d}^-\Phi}\, \Phi_{\mathsf{D_{st}}}(x,y)$$

Soundness of the Herbrand Dialectica interpretation

Theorem (van den Berg *et al*. 2012). Let Φ be a formula of system H and let Δ_{int} be a set of internal formulas. If

$$\mathsf{H} + \Delta_{\text{int}} \ \vdash \ \Phi$$

then from the proof one can extract a closed term $t : (\mathsf{d}^{+}\Phi)^{*}$ in T^{*} such that

$$\mathsf{HA}^{\omega *} + \Delta_{\text{int}} \ \vdash \ \forall y^{\mathsf{d}^{-}\Phi} \Phi_{\mathsf{D_{st}}}(t, y).$$

Proof. By induction on the length of the derivation.

Another functional interpretation of H: Herbrand realisability

We firstly work out the types $\tau(\Phi)$ of (acutal) realisers for formula Φ. Then for each formula Φ and term $s : (\tau\Phi)^*$ we define s hr Φ

$\tau(a =_\sigma b)$	$:\equiv$	$\mathbb{1}$	$s \,\mathsf{hr}\, a = b$	$:\equiv$	$a = b$
$\tau(\mathsf{st}^\sigma(t))$	$:\equiv$	σ	$s \,\mathsf{hr}\, \mathsf{st}(t)$	$:\equiv$	$t \in s$
$\tau(A \wedge B)$	$:\equiv$	$\tau A \times \tau B$	$s \,\mathsf{hr}\, (A \wedge B)$	$:\equiv$	$s^1 \,\mathsf{hr}\, A \;\wedge\; s^2 \,\mathsf{hr}\, B$
$\tau(A \vee B)$	$:\equiv$	$\tau A \times \tau B$	$s \,\mathsf{hr}\, (A \vee B)$	$:\equiv$	$s^1 \,\mathsf{hr}\, A \vee\; s^2 \,\mathsf{hr}\, B$
$\tau(A \to B)$	$:\equiv$	$(\tau A)^* \to (\tau B)^*$	$s \,\mathsf{hr}\, (A \to B)$	$:\equiv$	$\forall^{\mathrm{st}} u(u \,\mathsf{hr}\, A \to s[u] \,\mathsf{hr}\, B)$
$\tau(\forall x^\sigma A)$	$:\equiv$	τA	$s \,\mathsf{hr}\, \forall x A(x)$	$:\equiv$	$\forall x(s \,\mathsf{hr}\, A(x))$
$\tau(\exists x^\sigma A)$	$:\equiv$	τA	$s \,\mathsf{hr}\, \exists x A(x)$	$:\equiv$	$\exists x(s \,\mathsf{hr}\, A(x))$
$\tau(\forall^{\mathrm{st}} x^\sigma A)$	$:\equiv$	$\sigma \to (\tau A)^*$	$s \,\mathsf{hr}\, \forall^{\mathrm{st}} x A(x)$	$:\equiv$	$\forall^{\mathrm{st}} x(s[x] \,\mathsf{hr}\, A(x))$
$\tau(\exists^{\mathrm{st}} x^\sigma A)$	$:\equiv$	$\sigma \times (\tau A)^*$	$s \,\mathsf{hr}\, \exists^{\mathrm{st}} x A(x)$	$:\equiv$	$\exists z \in s^1(s^2 \,\mathsf{hr}\, A(z))$

Similar to the situation of (standard) Dialectica and modified realisability, their Herbrand variants differ in the interpretation of implication.

First attempt to unify Herbrand functional interpretations

As in Oliva 2006, we introduced an uninterpreted bounded universal quantifier

$$\forall x \sqsubset t\, A(x)$$

where $x : \sigma$ is a variable and $t : \sigma^*$ is a term.

Then the parametrised formula interpretation $|A|^x_y$ is almost the same as the $\mathsf{D_{st}}$-interpretation except the case of implication

$$|A \to B|^R_{s,u} \;:\equiv\; \forall v \sqsubset R^2[s,u]\, |A|^s_v \to |B|^{R^1[s]}_u.$$

- Take $\forall x \sqsubset t\, A(x)$ to be $\forall x \in t\, A(x)$, then we get the Herbrand Dialectica.
- Take $\forall x \sqsubset t\, A(x)$ to be $\forall^{\mathrm{st}} x A(x)$, then we get the Herbrand realisability (because $s \text{ hr } A \;\leftrightarrow\; \forall^{\mathrm{st}} u |A|^s_u$).

Parametrised formula interpretation

We want a more general parametrised formula interpretation to obtain also the standard functional interpretations via its instantiations.

The interpreted system: $\mathsf{HA}^{\omega*}_{\mathsf{st}} \equiv \mathsf{HA}^{\omega*} + \mathsf{st}$

The verifying system: $\mathsf{HA}^{\circ} \equiv \mathsf{HA}^{\omega*} + \sigma^{\circ} + t \in w + \forall x \sqsubset t\, A(x)$

- σ° behaves as the type of finite sequences, e.g.
  - 'singleton' $\sigma \to \sigma^{\circ}$
  - 'concatenation' $\sigma^{\circ} \times \sigma^{\circ} \to \sigma^{\circ}$
  - 'pairing' $\sigma^{\circ} \times \rho^{\circ} \to (\sigma \times \rho)^{\circ}$
  - 'projections' $(\sigma_0 \times \sigma_1)^{\circ} \to \sigma_i$
  - 'application' $(\sigma \to \rho^{\circ})^{\circ} \times \sigma^{\circ} \to \rho^{\circ}$
- $t \in w$ behaves as the membership relation for $t : \sigma$ and $w : \sigma^{\circ}$
- $\forall x \sqsubset w\, A(x)$ behaves as a bounded, universal quantifier for $x : \sigma$ and $w : \sigma^{\circ}$

Parametrised formula interpretation (cont.)

Each formula Φ is associated with types $\tau^+\Phi$ and $\tau^-\Phi$:

$$\begin{aligned}
&\tau^+(a =_\sigma b) :\equiv \mathbb{1} && \tau^-(a =_\sigma b) :\equiv \mathbb{1}\\
&\tau^+(\mathsf{st}^\sigma(t)) :\equiv \sigma && \tau^-(\mathsf{st}(t)) :\equiv \mathbb{1}\\
&\tau^+(A \wedge B) :\equiv \tau^+A \times \tau^+B && \tau^-(A \wedge B) :\equiv \tau^-A \times \tau^-B\\
&\tau^+(A \vee B) :\equiv \tau^+A \times \tau^+B && \tau^-(A \vee B) :\equiv \tau^-A \times \tau^-B\\
&\tau^+(A \to B) :\equiv ((\tau^+A)^\circ \to (\tau^+B)^\circ) \times ((\tau^+A)^\circ \times \tau^-B \to (\tau^-A)^\circ) && \tau^-(A \to B) :\equiv (\tau^+A)^\circ \times \tau^-B\\
&\tau^+(\forall x^\sigma A) :\equiv \tau^+A && \tau^-(\forall x^\sigma A) :\equiv \tau^-A\\
&\tau^+(\exists x^\sigma A) :\equiv \tau^+A && \tau^-(\exists x^\sigma A) :\equiv (\tau^-A)^\circ\\
&\tau^+(\forall^{\mathrm{st}} x^\sigma A) :\equiv \sigma \to (\tau^+A)^\circ && \tau^-(\forall^{\mathrm{st}} x^\sigma A) :\equiv \sigma \times \tau^-A\\
&\tau^+(\exists^{\mathrm{st}} x^\sigma A) :\equiv \sigma \times \tau^+A && \tau^-(\exists^{\mathrm{st}} x^\sigma A) :\equiv (\tau^-A)^\circ
\end{aligned}$$

For each formula Φ and terms $r : (\tau^+\Phi)^\circ$ and $u : \tau^-\Phi$, we define formula $|\Phi|^r_u$

$$\begin{aligned}
&|a =_\sigma b|^r_u \;:\equiv\; a =_\sigma b && |\forall z^\sigma \Phi(z)|^r_u \;:\equiv\; \forall z^\sigma |\Phi(z)|^r_u\\
&|\mathsf{st}^\sigma(t)|^r_u \;:\equiv\; t \,\epsilon\, r && |\exists z^\sigma \Phi(z)|^r_u \;:\equiv\; \exists z^\sigma \forall v \,\epsilon\, u |\Phi(z)|^r_v\\
&|A \wedge B|^r_u \;:\equiv\; |A|^{r^1}_{u_1} \wedge |B|^{r^2}_{u_2} && |\forall^{\mathrm{st}} z^\sigma \Phi(z)|^R_{a,u} \;:\equiv\; |\Phi(a)|^{R[a]}_u\\
&|A \vee B|^r_u \;:\equiv\; |A|^{r^1}_{u_1} \vee |B|^{r^2}_{u_2} && |\exists^{\mathrm{st}} z^\sigma \Phi(z)|^r_u \;:\equiv\; \exists z \,\epsilon\, r^1 \forall v \,\epsilon\, u |\Phi(z)|^{r^2}_v\\
&|A \to B|^R_u \;:\equiv\; \forall v \sqsubset R^2[u]\, |A|^{u_1}_v \to |B|^{R^1[u_1]}_{u_2}
\end{aligned}$$

Parametrised formula interpretation $\mathsf{P}_{\mathsf{st}}(\Phi) \;:\equiv\; \exists^{\mathrm{st}} x^{(\tau^+\Phi)^\circ} \forall^{\mathrm{st}} y^{\tau^-\Phi}\; |\Phi|^x_y$

Soundness for the parametrised formula interpretation

Theorem. Let $\Delta_{\rm int}$ be a set of internal formula. If

$$\mathsf{HA}^{\omega *}_{\mathsf{st}} + \Delta_{\rm int} \ \vdash \ \Phi$$

then from the proof we can extract a closed term $t : (\tau^+ \Phi)^\circ$ in T° $(:\equiv \mathsf{T}^* + \circ)$ such that

$$\mathsf{HA}^\circ + \Delta_{\rm int} \ \vdash \ \forall y^{\tau^- \Phi} |\Phi|^t_y.$$

Proof. By induction on the length of the derivation.

Instantiations of the parametrised formula interpretation

σ°	$t \,\epsilon\, u$	$\forall x \sqsubset t\, A(x)$	Functional interpretations
σ	$t = u$	$A(t)$	(restricted) Dialectica interpretation
σ	$t = u$	$\forall^{\mathrm{st}} x A(x)$	modified realisability
σ	$t \leq^* u$	$\tilde{\forall} x \leq^* t\, A(x)$	bounded functional interpretation[5][6]
			⋮
σ^*	$t \in u$	$\forall x \in t\, A(x)$	Herbrand Dialectica interpretation
σ^*	$t \in u$	$\forall^{\mathrm{st}} x A(x)$	Herbrand realisability
			⋮

- One interpretation of "standardness" is totality.
- Then $\forall^{\mathrm{st}}$, $\exists^{\mathrm{st}}$ are the computational quantifiers in Berger's uniform HA.

[5] F. Ferreira and J. Gaspar, *Nonstandardness and the bounded functional interpretation*, Annals of Pure and Applied Logic 166 (2015), no. 6, 701–712.

[6] As pointed out by Paulo Oliva after the talk, the bounded functional interpretation may not be an instance but could be obtained by changing some conditions of the parameters.

Discussion I: Efficiency of term extraction via D_{st}

Motivation of the work: shorter proofs $\Rightarrow$ faster extraction & simpler terms

Extraction procedure may be faster, because

- nonstandard proofs, in many cases, are shorter than the usual ones,
- internal formulas and proofs are ignored.

Extracted terms may be computationally worse[7], because

- algorithms are hidden in external proofs,
- nonstandard axioms may introduced fake realisers.

[7] Examples: http://cj-xu.github.io/agda/nonstandard_dialectica/Examples.html

Discussion II: Implementation in intensional type theory

- Parametrised functional interpretation via Agda's parametrised modules.
- Difficulty: In intensional type theory, for arbitrary $\mathsf{HA}^{\omega *}_{\mathsf{st}}$ formula Φ, we have

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

  only up to identity type (similar to $\Pi(n, m : \mathbb{N}). n + m = m + n$).
  Then, given $r : \tau^{+/-}(\Phi)$ we have to transport it along the above equality/path to get an element of $\tau^{+/-}(\Phi[x := t])$, which makes proving the soundness theorem very difficult and the resulting proof unreadable.

  Solution: Add the above equation as a new rewriting rule to Agda.

Summary

- We reformulate Herbrand functional interpretations in a way that is suitable for a type-theoretic development.
- We extend Oliva's method to unify functional interpretations for nonstandard/uniform arithmetic.
- We implement the parametrised functional interpretation in Agda.

Thank you!