Constructive analysis

Helmut Schwichtenberg

Mathematisches Institut der Universität München

MAP, Genua, August/September 2006

Logic and arithmetic in finite types

- Free algebras
- Constants, terms, formulas
- Axioms of HA^{ω}, natural deduction
- Realizability interpretation, soundness
- Majorization and realizability

Types, free algebras

Our type system is defined by two type forming operations: arrow types $\rho \to \sigma$, and the formation of inductively generated types (free algebras) $\mu\vec{\alpha}\,\vec{\kappa}$, where $\vec{\alpha} = (\alpha_j)_{j=1,\ldots,N}$ is a list of distinct "type variables", and $\vec{\kappa} = (\kappa_i)_{i=1,\ldots,k}$ is a list of "constructor types", whose argument types contain $\alpha_1, \ldots, \alpha_N$ in strictly positive positions only.

Definition

Let $\vec{\alpha} = (\alpha_j)_{j=1,\ldots,N}$ be a list of distinct type variables. *Types* $\rho, \sigma, \tau, \mu \in \mathrm{Ty}$ and *constructor types* $\kappa \in \mathrm{KT}(\vec{\alpha})$ are defined inductively by

$$\frac{\vec{\rho}, \vec{\sigma}_1, \ldots, \vec{\sigma}_n \in \mathrm{Ty}}{\vec{\rho} \to (\vec{\sigma}_1 \to \alpha_{j_1}) \to \ldots \to (\vec{\sigma}_n \to \alpha_{j_n}) \to \alpha_j \in \mathrm{KT}(\vec{\alpha})} \quad (n \geq 0)$$

$$\frac{\kappa_1, \ldots, \kappa_n \in \mathrm{KT}(\vec{\alpha})}{(\mu\vec{\alpha}\,(\kappa_1, \ldots, \kappa_n))_j \in \mathrm{Ty}} \quad (n \geq 1) \qquad \frac{\rho, \sigma \in \mathrm{Ty}}{\rho \to \sigma \in \mathrm{Ty}}$$

Examples of free algebras

$\mathbf{U}$	$:= \mu\alpha\,\alpha,$	Unit
$\mathbf{B}$	$:= \mu\alpha\,(\alpha, \alpha),$	Booleans
$\mathbf{N}$	$:= \mu\alpha\,(\alpha, \alpha \to \alpha),$	Natural numbers
$\mathbf{L}(\rho)$	$:= \mu\alpha\,(\alpha, \rho \to \alpha \to \alpha),$	Lists
$\rho \otimes \sigma$	$:= \mu\alpha\,(\rho \to \sigma \to \alpha),$	(Tensor) product
$\rho + \sigma$	$:= \mu\alpha\,(\rho \to \alpha, \sigma \to \alpha),$	Sum
$(\mathrm{tree}, \mathrm{tlist})$	$:= \mu(\alpha, \beta)\,(\mathbf{N} \to \alpha, \beta \to \alpha,$ $\beta, \alpha \to \beta \to \beta),$	
$\mathbf{Bin}$	$:= \mu\alpha\,(\alpha, \alpha \to \alpha \to \alpha),$	Binary trees
$\mathcal{O}$	$:= \mu\alpha\,(\alpha, \alpha \to \alpha, (\mathbf{N} \to \alpha) \to \alpha),$	Ordinals
$\mathcal{T}_0$	$:= \mathbf{N},$	
$\mathcal{T}_{n+1}$	$:= \mu\alpha\,(\alpha, (\mathcal{T}_n \to \alpha) \to \alpha).$	Trees

Finitary algebras

A type is called finitary if it is a μ-type with all its parameter types $\vec{\rho}$ finitary, and in all its constructor types

$$\vec{\rho} \to (\vec{\sigma}_1 \to \alpha_{j_1}) \to \ldots \to (\vec{\sigma}_n \to \alpha_{j_n}) \to \alpha_j$$

the $\vec{\sigma}_1, \ldots, \vec{\sigma}_n$ are all empty. In the examples above **U**, **B**, **N**, tree, tlist and **Bin** are all finitary, whereas $\mathcal{O}$ and $\mathcal{T}_{n+1}$ are not. $\mathbf{L}(\rho)$, $\rho \otimes \sigma$ and $\rho + \sigma$ are finitary provided their parameter types are. An argument position in a type is called finitary if it is occupied by a finitary type.

Recursion operators, by example

$$\mathrm{tt}^{\mathbf{B}} := \mathrm{C}_1^{\mathbf{B}}, \quad \mathrm{ff}^{\mathbf{B}} := \mathrm{C}_2^{\mathbf{B}},$$
$$\mathcal{R}_{\mathbf{B}}^{\tau} \colon \tau \to \tau \to \mathbf{B} \to \tau,$$

$$0^{\mathbf{N}} := \mathrm{C}_1^{\mathbf{N}}, \quad \mathrm{S}^{\mathbf{N}\to\mathbf{N}} := \mathrm{C}_2^{\mathbf{N}},$$
$$\mathcal{R}_{\mathbf{N}}^{\tau} \colon \tau \to (\mathbf{N} \to \tau \to \tau) \to \mathbf{N} \to \tau,$$

$$\mathrm{nil}^{\mathbf{L}(\alpha)} := \mathrm{C}_1^{\mathbf{L}(\alpha)}, \quad \mathrm{cons}^{\alpha\to\mathbf{L}(\alpha)\to\mathbf{L}(\alpha)} := \mathrm{C}_2^{\mathbf{L}(\alpha)},$$
$$\mathcal{R}_{\mathbf{L}(\alpha)}^{\tau} \colon \tau \to (\alpha \to \mathbf{L}(\alpha) \to \tau \to \tau) \to \mathbf{L}(\alpha) \to \tau,$$

$$\left(\mathrm{Inl}_{\rho\sigma}\right)^{\rho\to\rho+\sigma} := \mathrm{C}_1^{\rho+\sigma},$$
$$\left(\mathrm{Inr}_{\rho\sigma}\right)^{\sigma\to\rho+\sigma} := \mathrm{C}_2^{\rho+\sigma},$$
$$\mathcal{R}_{\rho+\sigma}^{\tau} \colon (\rho \to \tau) \to (\sigma \to \tau) \to \rho + \sigma \to \tau,$$

$$\left(\otimes_{\rho\sigma}^{+}\right)^{\rho\to\sigma\to\rho\otimes\sigma} := \mathrm{C}_1^{\rho\otimes\sigma},$$
$$\mathcal{R}_{\rho\otimes\sigma}^{\tau} \colon (\rho \to \sigma \to \tau) \to \rho\otimes\sigma \to \tau.$$

Example: decidable equality

We can define decidable equality $=_\mu : \mu \to \mu \to \mathbf{B}$, for finitary base types μ.

$$
\begin{aligned}
(0 = 0) &:= \mathrm{tt}, \\
(0 = \mathrm{S}(n)) &:= \mathrm{ff}, \\
(\mathrm{S}(m) = 0) &:= \mathrm{ff}, \\
(\mathrm{S}(m) = \mathrm{S}(n)) &:= \ (n = m).
\end{aligned}
$$

Conversion

To define the conversion relation, it will be helpful to use the following notation. Let $\vec{\mu} = \mu\vec{\alpha}\,\vec{\kappa}$ and

$$\kappa_i = \rho_1 \to \dots \rho_m \to (\vec{\sigma}_1 \to \alpha_{j_1}) \to \dots (\vec{\sigma}_n \to \alpha_{j_n}) \to \alpha_j \in \mathrm{KT}(\vec{\alpha}),$$

and consider $\mathrm{C}_i^{\vec{\mu}}\vec{N}$. Then we write $\vec{N}^P = N_1^P, \dots, N_m^P$ for the parameter arguments $N_1^{\rho_1}, \dots, N_m^{\rho_m}$ and $\vec{N}^R = N_1^R, \dots, N_n^R$ for the recursive arguments $N_{m+1}^{\vec{\sigma}_1 \to \mu_{j_1}}, \dots, N_{m+n}^{\vec{\sigma}_n \to \mu_{j_n}}$, and n^R for the number n of recursive arguments.
We define a conversion relation $\mapsto_\rho$ between terms of type ρ by

$$(\lambda x M)N \mapsto M[x := N],$$
$$\lambda x.Mx \mapsto M \quad \text{if } x \notin \mathrm{FV}(M) \ (M \text{ not an abstraction}),$$
$$(\mathcal{R}_j\vec{M})^{\mu_j \to \tau_j}(\mathrm{C}_i^{\vec{\mu}}\vec{N}) \mapsto M_i\vec{N}\big((\mathcal{R}_{j_1}\vec{M}) \circ N_1^R\big) \dots \big((\mathcal{R}_{j_n}\vec{M}) \circ N_n^R\big).$$

Here we have written $\mathcal{R}_j$ for $\mathcal{R}_{\mu_j}^{\vec{\mu},\vec{\tau}}$.

Reduction

The one step reduction relation $\longrightarrow$ can now be defined as follows. $M \longrightarrow N$ if N is obtained from M by replacing a subterm M' in M by N', where $M' \mapsto N'$. The reduction relations $\longrightarrow^+$ and $\longrightarrow^*$ are the transitive and the reflexive transitive closure of $\longrightarrow$, respectively. For $\vec{M} = M_1, \ldots, M_n$ we write $\vec{M} \longrightarrow \vec{M'}$ if $M_i \longrightarrow M'_i$ for some $i \in \{1, \ldots, n\}$ and $M_j = M'_j$ for all $i \neq j \in \{1, \ldots, n\}$. A term M is normal (or in normal form) if there is no term N such that $M \longrightarrow N$. Clearly normal closed terms are of the form $\mathrm{C}_i^{\vec{\mu}} \vec{N}$.

Definition

The set SN of strongly normalizing terms is inductively defined by

$$(\forall N. M \longrightarrow N \Rightarrow N \in \mathrm{SN}) \Rightarrow M \in \mathrm{SN}.$$

Theorem

Every term is strongly normalizing.

Atomic formulas, formulas

There is a syntactically defined definitional equality on terms: two terms are called definitionally equal when they reduce to the same normal form.
Recall that we have a decidable equality $=_\mu : \mu \to \mu \to \mathbf{B}$, for finitary base types μ. Every every atomic formula has the form $\mathrm{atom}(r^{\mathbf{B}})$, i.e., is built from a boolean term $r^{\mathbf{B}}$. In particular, there is no need for (logical) falsity $\bot$, since we can take the atomic formula $F := \mathrm{atom}(\mathrm{ff})$ – called arithmetical falsity – built from the boolean constant ff instead.
The formulas of HA^ω are built from atomic ones by the connectives $\to$, $\forall$, $\wedge$ and $\exists$. We define negation $\neg A$ by $A \to F$.

Natural deduction

derivation	term
$u : A$	u^A
$\dfrac{\begin{array}{c}[u : A] \\ \vert\ M \\ B\end{array}}{A \to B} \to^+ u$	$(\lambda u^A M^B)^{A \to B}$
$\dfrac{\vert\ M \quad A \to B \qquad \vert\ N \quad A}{B} \to^-$	$(M^{A \to B} N^A)^B$

Natural deduction: $\forall$-rules

derivation	term
$\dfrac{\begin{array}{c}\vert\, M \\ A\end{array}}{\forall_x A}\,\forall^+ x \quad \text{(VarC)}$	$(\lambda x M^A)^{\forall_x A}$ (VarC)
$\dfrac{\begin{array}{cc}\vert\, M & \\ \forall_x A(x) & r\end{array}}{A(r)}\,\forall^-$	$(M^{\forall_x A(x)}\, r)^{A(r)}$

Induction axioms, by examples

$$\mathrm{Ind}_{p,A}\colon A(\mathtt{tt}) \to A(\mathtt{ff}) \to \forall_{p^{\mathsf{B}}}\, A(p),$$
$$\mathrm{Ind}_{n,A}\colon A(0) \to \forall_n(A(n) \to A(\mathrm{S}n)) \to \ \forall_{n^{\mathsf{N}}}\, A(n),$$
$$\mathrm{Ind}_{l,A}\colon A(\mathrm{nil}) \to \forall_{x,l}(A(l) \to A(\mathrm{cons}(x,l))) \to \forall_{l^{\mathsf{L}(\alpha)}}\, A(l),$$
$$\mathrm{Ind}_{x,A}\colon \forall_{y_1}\, A(\mathrm{Inl}(y_1)) \to \forall_{y_2}\, A(\mathrm{Inr}(y_2)) \to \forall_{x^{\rho_1+\rho_2}}\, A(x).$$

Logical axioms

The logical axioms are $\wedge^+$, $\wedge^-$, $\exists^+$ and $\exists^-$, and the truth axiom $\mathrm{Ax}_{\mathrm{tt}} : \mathrm{atom}(\mathrm{tt})$.
We postulate the compatibility axioms for f of a type of level ≤ 1:

$$x_1 =_{\mu_1} y_1 \to \cdots \to x_n =_{\mu_n} y_n \to f\vec{x} =_{\mu} f\vec{y}.$$

Let HA^ω be the theory based on the axioms above including the induction axioms, and ML^ω be the (many-sorted) minimal logic, where the induction axioms are left out.

Extensionality

We define pointwise equality $=_\rho$, by induction on the type. $x_1 =_\mu x_2$ is already defined, and

$$(x_1 =_{\rho\to\sigma} x_2) := \forall_y(x_1 y =_\sigma x_2 y).$$

The extensionality axioms are

$$y_1 =_\rho y_2 \to xy_1 =_\sigma xy_2.$$

We write E-HA$^\omega$ when the extensionality axioms are present. Howard proved that already the first non trivial instance of the extensionality scheme

$$y_1 =_1 y_2 \to xy_1 =_{\mathbf{N}} xy_2$$

does not have a Dialectica realizer. In fact, he introduced the majorizing relation as a tool to prove this result. This is in contrast to the realizability interpretation, where extensionality axioms are unproblematic, since they are $\exists$-free.

Weak extensionality rule

It is customary to try to alleviate the difficulty of not being able to use extensionality when formalizing mathematical arguments (when an application of the Dialectica interpretation is envisaged) by adding a so-called weak extensionality rule

$$\frac{A_0 \to r =_\rho s}{A_0 \to t(r) =_\sigma t(s)} \quad (A_0 \text{ quantifier-free})$$

to the formal system considered. Since the conclusion is (equivalent to) a purely universal formula, this does not change the behaviour of the formal system w.r.t. the Dialectica interpretation. We write WE-HA$^\omega$ when the weak extensionality rule is present, but not the extensionality axioms.

Other useful equality notions

Later we will consider some more equality notions: extensional equality $=^e_\rho$, hereditary extensional equality $\approx_\rho$, and Leibniz equality, where the latter is defined inductively, by the introduction axiom

$$\mathrm{Eq}^+ : \forall_x \mathrm{Eq}(x, x)$$

and the elimination axiom

$$\mathrm{Eq}^- : \forall_{x,y}\big(\forall_x A(x, x) \to \mathrm{Eq}(x, y) \to A(x, y)\big).$$

Notice that Leibniz equality introduces additional atomic formulas, which are not any more given by boolean terms. For types of level ≤ 1, pointwise and extensional equality will coincide.

Further axiom schemes

The axiom of choice $(\mathrm{AC}_{\rho,\sigma})$ is the scheme

$$\forall_{x^\rho}\exists_{y^\sigma} A(x,y) \to \exists_{f^{\rho\to\sigma}}\forall_{x^\rho} A(x,f(x)).$$

(AC) is the collection of all $(\mathrm{AC}_{\rho,\sigma})$. By independence of premise $(\mathrm{IP}^{\omega}_{\exists\text{-free}})$ we mean the scheme

$$(A \to \exists_{x^\rho} B) \to \exists_{x^\rho}(A \to B) \quad \text{with } A \ \exists\text{-free and } x \notin \mathrm{FV}(A).$$

The type of a realizer

$A \mapsto \tau(A)$ (a type or the symbol ε). In case $\tau(A) = \varepsilon$ proofs of A have no computational content; then A is called Harrop formula.

$$\tau(P(\vec{s})) := \varepsilon,$$

$$\tau(\exists_{x^\rho} A) := \begin{cases} \rho & \text{if } \tau(A) = \varepsilon \\ \rho \otimes \tau(A) & \text{otherwise,} \end{cases}$$

$$\tau(\forall_{x^\rho} A) := \begin{cases} \varepsilon & \text{if } \tau(A) = \varepsilon \\ \rho \to \tau(A) & \text{otherwise,} \end{cases}$$

$$\tau(A \to B) := \begin{cases} \tau(B) & \text{if } \tau(A) = \varepsilon \\ \varepsilon & \text{if } \tau(B) = \varepsilon \\ \tau(A) \to \tau(B) & \text{otherwise,} \end{cases}$$

$$\tau(A \wedge B) := \begin{cases} \tau(B) & \text{if } \tau(A) = \varepsilon \\ \tau(A) & \text{if } \tau(B) = \varepsilon \\ \tau(A) \otimes \tau(B) & \text{otherwise.} \end{cases}$$

Extracted terms

We define $[\![M]\!]$, for a derivation M using axioms $\exists^{\pm}$, induction axioms, (AC) and $(\mathrm{IP}^{\omega}_{\exists\text{-free}})$ and some $\exists$-free axioms.
Assume first that M derives a formula A with $\tau(A) \neq \varepsilon$. Then its extracted term $[\![M]\!]$ of type $\tau(A)$ is

$$[\![u^A]\!] := x_u^{\tau(A)} \quad (x_u^{\tau(A)} \text{ uniquely associated with } u^A),$$

$$[\![\lambda u^A M]\!] := \begin{cases} [\![M]\!] & \text{if } \tau(A) = \varepsilon \\ \lambda x_u^{\tau(A)} [\![M]\!] & \text{otherwise,} \end{cases}$$

$$[\![M^{A\to B} N]\!] := \begin{cases} [\![M]\!] & \text{if } \tau(A) = \varepsilon \\ [\![M]\!][\![N]\!] & \text{otherwise,} \end{cases}$$

$$[\![(\lambda x^\rho M)^{\forall_x A}]\!] := \lambda x^\rho [\![M]\!],$$

$$[\![M^{\forall_x A} t]\!] := [\![M]\!] t.$$

For derivations M^A where $\tau(A) = \varepsilon$ (i.e., A is a Harrop formula) we define $[\![M]\!] := \varepsilon$ (ε some new symbol).

Extracted terms for the axioms

For the axioms

$$\exists^{+}_{x,A} : \forall_{x^\rho}(A \to \exists_{x^\rho} A)$$
$$\exists^{-}_{x,A,B} : \exists_{x^\rho} A \to \forall_{x^\rho}(A \to B) \to B$$

we set

$$[\![\exists^{+}_{x^\rho,A}]\!] := \begin{cases} \lambda x^\rho x & \text{if } \tau(A) = \varepsilon \\ \lambda x^\rho \lambda y^{\tau(A)} \langle x, y \rangle & \text{otherwise} \end{cases}$$

$$[\![\exists^{-}_{x^\rho,A,B}]\!] := \begin{cases} \lambda x^\rho \lambda f^{\rho \to \tau(B)}.fx & \text{if } \tau(A) = \varepsilon \\ \lambda z^{\rho \otimes \tau(A)} \lambda f^{\rho \to \tau(A) \to \tau(B)}.f(z0)(z1) & \text{otherwise.} \end{cases}$$

Extracted terms for the axioms (continued)

For the axioms

$$\wedge^+ : A \to B \to A \wedge B$$
$$\wedge^- : (A \to B \to C) \to A \wedge B \to C$$

we set

$$[\![\wedge^+]\!] := \begin{cases} \lambda x^{\tau(A)} x & \text{if } \tau(B) = \varepsilon \\ \lambda y^{\tau(B)} y & \text{if } \tau(A) = \varepsilon \\ \lambda x^{\tau(A)} \lambda y^{\tau(B)} \langle x, y \rangle & \text{otherwise} \end{cases}$$

$$[\![\wedge^-]\!] := \begin{cases} \lambda z^{\tau(C)} z & \text{if } \tau(A){=}\varepsilon,\ \tau(B){=}\varepsilon \\ \lambda f^{\tau(A)\to\tau(C)} \lambda y^{\tau(B)}.fy & \text{if } \tau(A){=}\varepsilon,\ \tau(B){\neq}\varepsilon \\ \lambda f^{\tau(A)\to\tau(C)} \lambda x^{\tau(A)}.fx & \text{if } \tau(A){\neq}\varepsilon,\ \tau(B){=}\varepsilon \\ \lambda f^{\tau(A)\to\tau(B)\to\tau(C)} \lambda z^{\tau(A)\otimes\tau(B)}.f(z0)(z1) & \text{if } \tau(A){\neq}\varepsilon,\ \tau(B){\neq}\varepsilon. \end{cases}$$

Extracted terms for the axioms (continued)

The extracted term $[\![\mathrm{Ind}_j]\!]$ of an induction axiom is defined to be the recursion operator $\mathcal{R}^{\vec{\mu},\vec{\tau}}_{\mu_j}$. Here $\vec{\mu},\vec{\tau}$ list only the types μ_j,τ_j with $\tau_j := \tau(A_j) \neq \varepsilon$, i.e., the recursion operator is simplified accordingly.

Example

For the induction scheme

$$\mathrm{Ind}_{n,A}\colon A(0) \to \forall_n(A(n) \to A(n+1)) \to \forall_n A(n)$$

we have

$$[\![\mathrm{Ind}_{n,A}]\!] := \mathcal{R}^{\tau}_{\mathbf{N}}\colon \tau \to (\mathbf{N} \to \tau \to \tau) \to \mathbf{N} \to \tau,$$

where $\tau := \tau(A) \neq \varepsilon$.

As extracted terms of (AC) and $(\mathrm{IP}^{\omega}_{\exists\text{-free}})$ we can take identities of the appropriate types.

Modified realizability

We define formulas $r \operatorname{mr} A$, where A is a formula and r is a term of type $\tau(A)$ if the latter is a type, or the symbol ε if $\tau(A) = \varepsilon$.

$$r \operatorname{mr} P(\vec{s}) := P(\vec{s}),$$

$$r \operatorname{mr} (\exists_x A(x)) := \begin{cases} \varepsilon \operatorname{mr} A(r) & \text{if } \tau(A) = \varepsilon \\ r1 \operatorname{mr} A(r0) & \text{otherwise,} \end{cases}$$

$$r \operatorname{mr} (\forall_x A) := \begin{cases} \forall_x \varepsilon \operatorname{mr} A & \text{if } \tau(A) = \varepsilon \\ \forall_x rx \operatorname{mr} A & \text{otherwise,} \end{cases}$$

$$r \operatorname{mr} (A \to B) := \begin{cases} \varepsilon \operatorname{mr} A \to r \operatorname{mr} B & \text{if } \tau(A) = \varepsilon \\ \forall_x (x \operatorname{mr} A \to \varepsilon \operatorname{mr} B) & \text{if } \tau(A) \neq \varepsilon = \tau(B) \\ \forall_x (x \operatorname{mr} A \to rx \operatorname{mr} B) & \text{otherwise,} \end{cases}$$

$$r \operatorname{mr} (A \wedge B) := \begin{cases} r \operatorname{mr} B & \text{if } \tau(A) = \varepsilon \\ r \operatorname{mr} A & \text{if } \tau(B) = \varepsilon \\ (r0 \operatorname{mr} A) \wedge (r1 \operatorname{mr} B) & \text{otherwise.} \end{cases}$$

Realizability (continued)

Formulas which do not contain the existence quantifier $\exists$ play a special role in this context; we call them $\exists$-free (or invariant); in the literature such formulas are also called "negative". Their crucial property is that for an $\exists$-free formula A we have $\varepsilon \operatorname{mr} A = A$. Notice also that every formula $r \operatorname{mr} A$ is $\exists$-free.

Soundness

Theorem

Let $M\colon A$ be a derivation in $\mathsf{HA}^\omega + \mathrm{AC} + \mathrm{IP}^\omega_{\exists\text{-free}} + \mathrm{Ax}_{\exists\text{-free}}$ from assumptions $u_i\colon C_i$ $(i = 1, \ldots, n)$. Then we can find a derivation $\mu(M)$ in $\mathsf{HA}^\omega + \mathrm{Ax}_{\exists\text{-free}}$ of

$$[\![M]\!]\ \mathrm{mr}\ A$$

from assumptions $\bar{u}_i\colon x_{u_i}\ \mathrm{mr}\ C_i$.

Proof.

Induction on M. □

Characterization

We consider the question under what conditions a formula A and its modified realizability interpretation $\exists_x x \text{ mr } A$ are equivalent.

Theorem (Characterization)

$$\mathrm{AC} + \mathrm{IP}^{\omega}_{\exists\text{-free}} \vdash A \leftrightarrow \exists_x x \text{ mr } A.$$

Proof.

Induction on A; case $A \to B$ with $\tau(A) \neq \varepsilon$ and $\tau(B) \neq \varepsilon$.

$$\begin{aligned}
(A \to B) &\leftrightarrow (\exists_x x \text{ mr } A \to \exists_y y \text{ mr } B) && \text{by IH} \\
&\leftrightarrow \forall_x (x \text{ mr } A \to \exists_y y \text{ mr } B) && \text{by } \mathrm{ML}^{\omega} \\
&\leftrightarrow \forall_x \exists_y (x \text{ mr } A \to y \text{ mr } B) && \text{by } (\mathrm{IP}^{\omega}_{\exists\text{-free}}) \\
&\leftrightarrow \exists_f \forall_x (x \text{ mr } A \to f(x) \text{ mr } B) && \text{by } (\mathrm{AC}) \\
&\leftrightarrow \exists_f f \text{ mr } (A \to B).
\end{aligned}$$

Extraction

Theorem (Extraction)

Assume $\mathsf{HA}^\omega + \mathrm{AC} + \mathrm{IP}^\omega_{\exists\text{-free}} + \mathrm{Ax}_{\exists\text{-free}} \vdash \forall_x \exists_y A(x,y)$ *with* $A(x,y)$ *an arbitrary formula with at most the displayed variables free. Then we can find a closed* HA^ω*-term* t *such that*

$$\mathsf{HA}^\omega + \mathrm{AC} + \mathrm{IP}^\omega_{\exists\text{-free}} + \mathrm{Ax}_{\exists\text{-free}} \vdash \forall_x A(x, tx).$$

Proof.

We assume $\tau(A(x,y)) \neq \varepsilon$. $\mathsf{HA}^\omega + \mathrm{Ax}_{\exists\text{-free}}$ proves

$$\begin{aligned}
&[\![M]\!] \operatorname{mr} \forall_x \exists_y A(x,y) && \text{by the Soundness Theorem}\\
&\forall_x([\![M]\!]x \operatorname{mr} \exists_y A(x,y))\\
&\forall_x([\![M]\!]x1 \operatorname{mr} A(x,[\![M]\!]x0)).
\end{aligned}$$

Hence $\mathsf{HA}^\omega + \mathrm{AC} + \mathrm{IP}^\omega_{\exists\text{-free}} + \mathrm{Ax}_{\exists\text{-free}} \vdash \forall_x A(x,[\![M]\!]x0)$ by the Characterization Theorem. □

Majorization

We assume here that all base types are finitary, and that $\geq_\mu$ is a given reflexive and transitive relation on the total ideals of base type μ such that

- for every $y \in G_\mu$ there are only finitely many $x \in G_\mu$ with $y \geq x$;
- there is a max-operation on G_μ such that

$$\max(x, y) \geq x, y,$$
$$z \geq x \to z \geq y \to z \geq \max(x, y).$$

Majorization (continued)

We extend $\geq_\mu$ to higher types, in a pointwise fashion (as for $=_\mu$)

$$x_1 \geq_{\rho\to\sigma} x_2 := \forall_y(x_1 y \geq_\sigma x_2 y).$$

Following Howard, we define a relation $x^* \operatorname{maj}_\rho x$ (x^* hereditarily majorizes x) for $x^*, x \in G_\rho$, by induction on the type ρ:

$$\begin{aligned} x^* \operatorname{maj}_\mu x &:= x^* \geq_\mu x, \\ x^* \operatorname{maj}_{\rho\to\sigma} x &:= \forall_{y^*,y}(y^* \operatorname{maj}_\rho y \to x^* y^* \operatorname{maj}_\sigma xy). \end{aligned}$$

Lemma

(a) $\vdash x^* =_\rho \tilde{x}^* \to x =_\rho \tilde{x} \to x^* \operatorname{maj}_\rho x \to \tilde{x}^* \operatorname{maj}_\rho \tilde{x}$.

(b) $\vdash x^* \operatorname{maj}_\rho x \to x \geq_\rho \tilde{x} \to x^* \operatorname{maj}_\rho \tilde{x}$.

Proof.

Induction on ρ. We argue informally, and only treat (b). *Case* $\rho \to \sigma$. Assume $y^* \operatorname{maj}_\rho y$. Then $x^* y^* \operatorname{maj}_\sigma xy$ and $xy \geq_\sigma \tilde{x}y$, hence by IH $x^* y^* \operatorname{maj}_\sigma \tilde{x}y$.

Majorization of closed HA^ω-terms

x is called hereditarily majorizable if there is an x^* such that $x^* \operatorname{maj} x$.
Let 1 denote the type $\mathbf{N} \to \mathbf{N}$. Clearly, for every monotone function D of type 1 we have $D \operatorname{maj} D$. Moreover, $\mathcal{R}^\tau_\mu$ is hereditarily majorizable:

Lemma (Majorization)

(a) *Define* $M\colon (\mu \to \tau) \to \mu \to \tau$ *with* $\tau = \vec{\rho} \to \mu'$ *by*

$$Mfn\vec{x} := \max_{i \le n} fi\vec{x}.$$

Then $\mathrm{HA}^\omega \vdash \forall_n \bar{f}\, n \operatorname{maj} fn \to M\bar{f} \operatorname{maj} f$.

(b) $\mathrm{HA}^\omega \vdash f^*, g^* \operatorname{maj} f, g \to \mathcal{R}_\mu f^* g^* n \operatorname{maj} \mathcal{R}_\mu fgn$.

(c) *Define* $\mathcal{R}^*_\mu fg := M(\mathcal{R}_\mu fg)$. *Then* $\mathrm{HA}^\omega \vdash \mathcal{R}^*_\mu \operatorname{maj} \mathcal{R}_\mu$.

Majorization of closed HA^ω-terms

Lemma

Let $r(\vec{x})$ be a HA^ω-term with free variables among $\vec{x}$. Assume that $\mathrm{HA}^\omega \vdash c^* \operatorname{maj} c$ for all constants c in r. Let r^* be r with all constants c replaced by c^*. Then $\mathrm{HA}^\omega \vdash \vec{x}^* \operatorname{maj} \vec{x} \to r(\vec{x}^*) \operatorname{maj} r(\vec{x})$.

Proof.

Induction on r. *Case* $\lambda y\, r(y, \vec{x})$. We argue informally. Assume $\vec{x}^* \operatorname{maj} \vec{x}$. We must show $y^* \operatorname{maj} y \to (\lambda y\, r(y, \vec{x}^*))y^* \operatorname{maj} (\lambda y\, r(y, \vec{x}))y$. So assume $y^* \operatorname{maj} y$. Then by IH $r(y^*, \vec{x}^*) \operatorname{maj} r(y, \vec{x})$, which is our claim. □

Hence every closed term r of HA^ω is hereditarily majorizable. In fact, we have constructed a closed term r^* of HA^ω such that $r^* \operatorname{maj} r$.

Extraction of uniform bounds

Theorem

Let s be a closed HA^ω-term, $A(x,y,z)$ a formula with at most the displayed variables free, and τ a type of level ≤ 2. Assume that

$$\mathsf{HA}^\omega + \mathrm{AC} + \mathrm{IP}^\omega_{\exists\text{-free}} \vdash \forall_{x^1} \forall_{y \leq_\rho sx} \exists_{z^\tau} A(x,y,z).$$

Then we can find a closed HA^ω-term t such that

$$\mathsf{HA}^\omega + \mathrm{AC} + \mathrm{IP}^\omega_{\exists\text{-free}} \vdash \forall_{x^1} \forall_{y \leq_\rho sx} \exists_{z \leq_\tau tx} A(x,y,z).$$

Extraction of uniform bounds: proof

Let $H^\omega := \mathrm{HA}^\omega + \mathrm{AC} + \mathrm{IP}^\omega_{\exists\text{-free}}$. By the Soundness Theorem we have a closed HA^ω-term r such that

$$H^\omega \vdash r \ \mathrm{mr}\ \forall_{x^1}\forall_{y\leq_\rho sx}\exists_{z^\tau} A(x,y,z).$$

Unfolding the definition of mr and using the fact that $y \leq_\rho sx$ is $\exists$-free we obtain (assuming $\tau(A(x,y,z)) \neq \varepsilon$)

$$H^\omega \vdash \forall_{x^1}\forall_{y\leq_\rho sx}\ rxy1 \ \mathrm{mr}\ A(x,y,rxy0)$$

and hence by the Characterization Theorem

$$H^\omega \vdash \forall_{x^1}\forall_{y\leq_\rho sx} A(x,y,rxy0).$$

Extraction of uniform bounds: proof (continued)

Recall

$$H^\omega \vdash \forall_{x^1} \forall_{y \leq_\rho sx} A(x, y, rxy0).$$

Let $r_1 := \lambda x \lambda y . rxy0$. Pick majorizing terms s^*, r_1^* for s, r_1. Writing x^M for Mx with the M from the Majorization Lemma we have $s^* x^M \ \mathrm{maj}_\rho \ sx$, hence

$$\mathsf{HA}^\omega \vdash \forall_{x^1} \forall_{y \leq_\rho sx} s^* x^M \ \mathrm{maj}_\rho \ y.$$

For simplicity assume $\tau = 2 := (\mathbf{N} \to \mathbf{N}) \to \mathbf{N}$. Then

$$\mathsf{HA}^\omega \vdash \forall_{x^1} \forall_{y \leq_\rho sx} \forall_f \ r_1^* x^M (s^* x^M) f^M \geq_{\mathbf{N}} r_1 xyf.$$

Hence we can take $t := \lambda x \lambda f . r_1^* x^M (s^* x^M) f^M$, because $tx \geq_2 t_1 xy =_{\mathrm{def}} rxy0$.

Fan rule

Corollary (Fan Rule)

Let $A(y, n)$ be a formula with at most the displayed variables free. Assume that

$$\mathrm{HA}^\omega + \mathrm{AC} + \mathrm{IP}^\omega_{\exists\text{-free}} \vdash \forall_{y^1} \exists_{n^{\mathbf{N}}} A(y, n).$$

Then

$$\mathrm{HA}^\omega + \mathrm{AC} + \mathrm{IP}^\omega_{\exists\text{-free}} \vdash \forall_{x^1} \exists_{m^{\mathbf{N}}} \forall_{y \leq_1 x} \exists_{n \leq_{\mathbf{N}} m} A(y, n).$$

Proof.

Let s be the identity in the theorem above. Take $m := tx$. □

Constructive real analysis

1. Motivation
2. Tools: Reals, continuous functions
3. Inverse functions
4. Implementation, and demonstration
5. ODE's: Cauchy-Euler method
6. Moore's Kth-order method

Motivation

- "Mathematics as a numerical language".
- Extract programs from proofs, for exact real numbers.
- Special emphasis on low type level witnesses (making use of separability).

Tools

. . . for algorithmically reasonable proofs: Small variants of Bishop/Bridges' development of constructive analysis.

Idea: use separability to avoid high type levels.

Reals

A real number x is a pair $((a_n)_{n\in\mathbb{N}}, \alpha)$ with $a_n \in \mathbb{Q}$ and $\alpha\colon \mathbb{N} \to \mathbb{N}$ such that $(a_n)_n$ is a Cauchy sequence with modulus α, that is

$$\forall_{k,n,m} \left(\alpha(k) \leq n, m \to |a_n - a_m| \leq 2^{-k}\right),$$

and α is weakly increasing.

Two reals $x := ((a_n)_n, \alpha)$, $y := ((b_n)_n, \beta)$ are equivalent (written $x = y$), if

$$\forall_k |a_{\alpha(k+1)} - b_{\beta(k+1)}| \leq 2^{-k}.$$

Nonnegative and positive reals

A real $x := ((a_n)_n, \alpha)$ is nonnegative (written $x \in \mathbb{R}^{0+}$) if

$$\forall_k \; -2^{-k} \leq a_{\alpha(k)}.$$

It is k-positive (written $x \in_k \mathbb{R}^+$) if

$$2^{-k} \leq a_{\alpha(k+1)}.$$

$x \in \mathbb{R}^{0+}$ and $x \in_k \mathbb{R}^+$ are compatible with equivalence.
Can define $x \mapsto k_x$ such that $a_n \leq 2^{k_x}$ for all n.
However, $x \mapsto k_x$ is not compatible with equivalence.

Arithmetical functions

Given $x := ((a_n)_n, \alpha)$ and $y := ((b_n)_n, \beta)$, define

z	c_n	$\gamma(k)$
$x + y$	$a_n + b_n$	$\max(\alpha(k+1), \beta(k+1))$
$-x$	$-a_n$	$\alpha(k)$
$\lvert x \rvert$	$\lvert a_n \rvert$	$\alpha(k)$
$x \cdot y$	$a_n \cdot b_n$	$\max(\alpha(k+1+k_{\lvert y \rvert}), \beta(k+1+k_{\lvert x \rvert}))$
$\frac{1}{x}$ for $\lvert x \rvert \in_I \mathbb{R}^+$	$\begin{cases} \frac{1}{a_n} & \text{if } a_n \neq 0 \\ 0 & \text{if } a_n = 0 \end{cases}$	$\alpha(2(I+1)+k)$

Cleaning up a real

After some computations involving reals, rationals in the Cauchy sequences may become complex. Hence: clean up a real, as follows.

Lemma

For every real $x = ((a_n)_n, \alpha)$ we can construct an equivalent real $y = ((b_n)_n, \beta)$ where the rationals b_n are of the form $c_n/2^n$ with integers c_n, and with modulus $\beta(k) = k + 2$.

Proof.

$c_n := \lfloor a_{\alpha(n)} \cdot 2^n \rfloor$. □

Redundant dyadic representation of reals

The existence of the usual b-adic representation of reals cannot be proved constructively ($1.000\ldots$ vs $.999\ldots$). Cure: in addition to $0, \ldots, b-1$ also admit -1 as a numeral. For $b = 2$:

Lemma

Every real x can be represented in the form

$$\sum_{n=-k}^{\infty} a_n 2^{-n} \quad \textit{with } a_n \in \{-1, 0, 1\}.$$

Notice: uniqueness is lost (this is not a problem).

Comparison of reals

Write $x \leq y$ for $y - x \in \mathbb{R}^{0+}$ and $x < y$ for $y - x \in \mathbb{R}^{+}$.

$$x \leq y \leftrightarrow \forall_k \exists_p \forall_{n \geq p}\, a_n \leq b_n + 2^{-k}$$

$$x < y \leftrightarrow \exists_{k,q} \forall_{n \geq q}\, a_n + 2^{-k} \leq b_n$$

Write $x <_{k,q} y$ (or simply $x <_k y$ if q is not needed) when we want to call these witnesses. Notice:

$$x \leq y \leftrightarrow y \not< x.$$

Continuous functions

A continuous function $f : I \to \mathbb{R}$ on a compact interval I with rational end points is given by

- an approximating map $h_f : (I \cap \mathbb{Q}) \to \mathbb{N} \to \mathbb{Q}$,
- a (uniform) modulus map $\alpha_f : \mathbb{N} \to \mathbb{N}$ such that $(h_f(c, n))_n$ is a real with modulus α_f;
- $\omega_f : \mathbb{N} \to \mathbb{N}$ (uniform) modulus of continuity:

$$|a - b| \leq 2^{-\omega_f(k)+1} \to |h_f(a, n) - h_f(b, n)| \leq 2^{-k}$$

for $n \geq \alpha_f(k)$. α_f, ω_f required to be weakly increasing.

Notice: h_f, α_f, ω_f are of type level 1 only.

Application of a continuous function to a real

Definition

Given a continuous function f (by h_f, α_f, ω_f) and a real $x := ((a_n)_n, \alpha)$, application $f(x)$ is defined to be

$$(h_f(a_n, n))_n$$

with modulus $k \mapsto \max(\alpha_f(k+2), \alpha(\omega_f(k+1)-1))$.

Lemma

$$x = y \to f(x) = f(y),$$
$$|x - y| \leq 2^{-\omega_f(k)} \to |f(x) - f(y)| \leq 2^{-k}.$$

Intermediate value theorem

Let $a < b$ be rationals. If $f : [a, b] \to \mathbb{R}$ is continuous with $f(a) \leq 0 \leq f(b)$, and with a uniform lower bound on its slope, then we can find $x \in [a, b]$ such that $f(x) = 0$.

Proof sketch.

1. Approximate Splitting Principle. Let x, y, z be given with $x < y$. Then either $z \leq y$ or $x \leq z$.
2. IVTAux. Assume $a \leq c < d \leq b$, say $2^{-n} < d - c$, and $f(c) \leq 0 \leq f(d)$. Construct c_1, d_1 with $d_1 - c_1 = \frac{2}{3}(d - c)$, such that $a \leq c \leq c_1 < d_1 \leq d \leq b$ and $f(c_1) \leq 0 \leq f(d_1)$.
3. IVTcds. Iterate the step $c, d \mapsto c_1, d_1$ in IVTAux.

Let $x = (c_n)_n$ and $y = (d_n)_n$ with the obvious modulus. As f is continuous, $f(x) = 0 = f(y)$ for the real number $x = y$. □

Inverse functions

Theorem

Let $f : [a, b] \to \mathbb{R}$ be continuous with a uniform lower bound on its slope. Let $f(a) \leq a' < b' \leq f(b)$. We can find a continuous $g : [a', b'] \to \mathbb{R}$ such that $f(g(y)) = y$ for every $y \in [a', b']$ and $g(f(x)) = x$ for every $x \in [a, b]$ such that $a' \leq f(x) \leq b'$.

Proof sketch.

Let $f(a) \leq a' < b' \leq f(b)$. Construct a continuous $g : [a', b'] \to \mathbb{R}$ by the Intermediate Value Theorem. $\square$

Example: squaring $f : [1, 2] \to [1, 4]$

Given by

- the approximating map $h_f(a, n) := a^2$,
- the uniform Cauchy modulus $\alpha_f(k) := 1$, and
- the modulus $k \mapsto k + 1$ of uniform continuity.

The lower bound on its slope is $l := 0$, because for all $c, d \in [1, 2]$

$$2^{-m} \leq d - c \to c^2 <_m d^2.$$

Then $h_g(u, n) := c_n^{(u)}$, as constructed in the IVT for $x^2 - u$, iterating IVTAux. The Cauchy modulus α_g is such that $(2/3)^n \leq 2^{-k+3}$ for $n \geq \alpha_g(k)$, and the modulus of uniform continuity is $\omega_f(k) := k + 2$.

Program extraction

Formalization: many details. Important: representation of data. Here: direct approach, by explicitely building the required number systems (natural numbers in binary, rationals, reals as Cauchy sequences of rationals with a modulus, continuous functions in the sense of the type-1 representation described above, etc.)

Method of program extraction based on modified realizability

Related work on program extraction

- Luis Cruz-Filipe: Thesis in Nijmegen 2004 (Geuvers), on C-CoRN.
- Stefan Berghofer: "Proofs, Programs and Executable Specifications in Higher Order Logic", 2003 (Nipkow).
- Monika Seisenberger: "On the Constructive Content of Proofs", 2003.

C-CoRN: Constructive Coq Repository at Nijmegen

Grew out of the FTA project.

- Strong extensionality required: $\forall_{x,y}(f(x)\#f(y) \rightarrow x\#y)$. Missing witness harmful for program extraction.
- The Set, Prop distinction in Coq was found to be insufficient. Introduced CProp in addition.

Alternative here: use modified realizability interpretation for (internal) program extraction. Soundness proof can be machine generated.

Some issues:

Animation

Suppose a proof of a theorem uses a lemma.

- Then the proof term contains the name of the lemma, say `L`.
- In the term extracted from this proof we want to preserve the structure of the original proof. So we use a new constant `cL` at places where the computational content of the lemma is needed.
- When we want to execute the program, we have to replace the constant `cL` corresponding to a lemma `L` by the extracted program of its proof. This can be achieved by adding computation rules for `cL`.
- We can be rather flexible here and enable/block rewriting by using `animate`/`deanimate` as desired.

Let

It often happens that a subterm has many occurrences in a term, which leads to unwanted recomputations when evaluating it.

- Cure: "optimize" the term after extraction, and replace for instance $M[x := N]$ with many occurrences of x in M by $(\lambda x M)N$ (or a corresponding "let"-expression).
- This can already be done at the proof level: When an object (value of a variable or realizer of a premise) is used more than once, make sure (if necessary by a cut) that the goal has the form $A \to B$ or $\forall_x A$.
- Now use the "identity lemma" $\mathtt{Id}\colon \hat{P} \to \hat{P}$, with a predicate variable $\hat{P}$. Its realizer then has the form $\lambda f, x.fx$.
- If $\mathtt{cId}$ is not animated, the extracted term has the form $\mathtt{cId}(\lambda x M)N$, which is printed as $[\mathtt{let}\ x\ N\ M]$.

Quantifiers without computational content

Besides the usual quantifiers, $\forall$ and $\exists$, Minlog has so-called non-computational quantifiers, $\forall^{\mathrm{nc}}$ and $\exists^{\mathrm{nc}}$, which allow for the extraction of simpler programs.

- The nc-quantifiers, which were first introduced by Berger (1993), can be viewed as a refinement of the Set/Prop distinction in constructive type systems like Coq or Agda.
- Intuitively, a proof of $\forall^{\mathrm{nc}}_x A(x)$ ($A(x)$ non-Harrop) represents a procedure that assigns to every x a proof $M(x)$ of $A(x)$ where $M(x)$ does not make "computational use" of x, i.e., the extracted program $[\![M(x)]\!]$ does not depend on x.
- Dually, a proof of $\exists^{\mathrm{nc}}_x A(x)$ is a proof of $M(x)$ for some x where the witness x is "hidden", that is, not available for computational use.

Ordinary differential equations

Let $f : D \to \mathbb{R}$ be continuous, $D \subseteq \mathbb{R}^2$. A solution of

$$y' = f(x, y), \tag{1}$$

on an interval I is a continuous function $\varphi : I \to \mathbb{R}$ with a continuous derivative φ' such that $(x, \varphi(x)) \in D$ and

$$\varphi'(x) = f(x, \varphi(x)) \qquad (x \in I)$$

Uniqueness

Theorem

Let $f : D \to \mathbb{R}$ be continuous. Assume that f satisfies a Lipschitz condition w.r.t. its 2nd argument:

$$|f(x, y_1) - f(x, y_2)| \leq L|y_1 - y_2|$$

with $L > 0$. Let $\varphi, \psi : I \to \mathbb{R}$ be two solutions of (1). If $\varphi(a) = \psi(a)$ for some $a \in I$, then $\varphi(x) = \psi(x)$ for all $x \in I$.

Example

The equation $y' = y^{1/3}$ with $y(0) = 0$ shows that the Lipschitz condition is necessary for uniqueness: we have two solutions $\varphi(x) = 0$ und $\varphi(x) = (\frac{2}{3}x)^{3/2}$.

Uniqueness

Theorem

Let $f : D \to \mathbb{R}$ be continuous. Assume that f satisfies a Lipschitz condition w.r.t. its 2nd argument:

$$|f(x, y_1) - f(x, y_2)| \leq L|y_1 - y_2|$$

with $L > 0$. Let $\varphi, \psi : I \to \mathbb{R}$ be two solutions of (1). If $\varphi(a) = \psi(a)$ for some $a \in I$, then $\varphi(x) = \psi(x)$ for all $x \in I$.

Example

The equation $y' = y^{1/3}$ with $y(0) = 0$ shows that the Lipschitz condition is necessary for uniqueness: we have two solutions $\varphi(x) = 0$ und $\varphi(x) = (\frac{2}{3}x)^{3/2}$.

Peano's existence theorem for ODEs

- . . . does not require a Lipschitz condition.
- But: Peano's existence theorem entails that for every real x we can decide whether $x \geq 0$ or $x \leq 0$ (Aberth 1970).
- Hence: We cannot expect to be able to prove it constructively.

Picard's existence theorem for ODEs

Theorem

On $R\colon |x - a_0| \le h, |y - b_0| \le Mh$, *let* f *be continuous and bounded by* M. *Assume that* f *satisfies a Lipschitz condition w.r.t. its 2nd argument. Let* $\varphi_0(x) := b_0$,

$$\varphi_{n+1}(x) := b_0 + \int_{a_0}^{x} f(t, \varphi_n(t))\, dt, \quad |x - a_0| \le h.$$

$(\varphi_n)_{n\in\mathbb{N}}$ *converges uniformly and absolutely to a solution of* (1).

Algorithmic problem: For $\varphi_{n+1}(x)$ one needs φ_n on $[a_0, x]$.

The Cauchy-Euler method

Simple idea: polygons ($\Rightarrow$ possibly adaptive). What is an "approximate solution"?

(a) It satisfies (1) up to ε.

(b) It differs from the exact solution by at most ε.

We aim for (b), but initially only get (a):

Theorem

On $R\colon |x - a_0| \leq h, |y - b_0| \leq Mh$, let f be continuous and bounded by M. We can construct an approximate solution (a polygon) $\varphi_n\colon [a_0 - h, a_0 + h] \to \mathbb{R}$ of (1) up to the error 2^{-n} such that $\varphi_n(a_0) = b_0$.

The fundamental inequality

Let $f : D \to \mathbb{R}$ be continuous, and satisfy a Lipschitz condition w.r.t. its second argument. Let

$$\varphi, \psi \colon [a, b] \to \mathbb{R}$$

be solutions up to $2^{-k}, 2^{-l}$ of (1). Assume $\varphi \leq \psi$ on $[a, b]$, or else that φ and ψ are rational polygons. Then

$$\big|\psi(x) - \varphi(x)\big| \leq e^{L(x-a)}\big|\psi(a) - \varphi(a)\big| + \frac{2^{-k} + 2^{-l}}{L}\big(e^{L(x-a)} - 1\big)$$

for all $x \in [a, b]$.

The Cauchy-Euler existence theorem for ODEs

Theorem

On $R\colon |x - a_0| \le h, |y - b_0| \le Mh$, *let* f *be continuous and bounded by* M. *Assume that* f *satisfies a Lipschitz condition w.r.t. its 2nd argument. Let* φ_n *be the rational polygon, which is an approximate solution of* (1) *up to the error* 2^{-n}*:*

$$|\varphi_n'(x) - f(x, \varphi_n(x))| \le 2^{-n} \text{ for } x \in I \text{ with } \varphi_n'(x) \text{ defined.}$$

(φ_n) *converges uniformly and absolutely to a solution of* (1).

Algorithmic note: φ_n is not defined recursively.

Approximate and exact solutions

Theorem

Assume the hypotheses of the Cauchy-Euler theorem. Let $\varphi\colon [a_0 - h, a_0 + h] \to \mathbb{R}$ *be an exact solution of* (1) *such that* $\varphi(a_0) = b_0$, φ_n *be an approximate solution up to the error* 2^{-n} *such that* $\varphi_n(a_0) = b_0$, *and* $\varphi \leq \varphi_n$ *or* $\varphi_n \leq \varphi$. *Then there is a constant* c *independent of* n *such that* $|\varphi(x) - \varphi_n(x)| \leq 2^{-n}c$ *for* $|x - a_0| \leq h$.

Proof.

By the fundamental inequality

$$|\varphi(x) - \varphi_n(x)| \leq 2^{-n} \cdot \underbrace{\frac{1}{L}(e^{Lh} - 1)}_{c}$$

Moore's Kth-order method

Reference: R.E. Moore, Interval Analysis, Prentice-Hall 1966.

It is convenient to consider autonomous systems of first-order differential equations

$$\frac{dy(x)}{dx} = f(y(x)). \tag{2}$$

$f : A_1 \times \cdots \times A_n (\subseteq \mathbb{R}^n) \to \mathbb{R}^n$ is a vector-valued function.

Assume that there is a positive rational L such that for all points u, v in A, f satisfies the Lipschitz condition

$$|f(u) - f(v)| \leq L|u - v|. \tag{3}$$

Moore's first-order method

$\sim$ Cauchy-Euler existence proof: We have a constructive proof of

$$\forall_{y_0} \exists_{a_0>0} \exists_{y(x):\,[-a_0,a_0]\to\mathbb{R}^n} \left(\frac{dy(x)}{dx} = f(y(x)) \wedge y(0) = y_0 \right).$$

By the Lipschitz condition we can find an a_0 which works uniformly for all y_0 in some $[a, b]$, hence

$$\forall_{y_0\in[a,b]} \exists_{y(x):\,[-a_0,a_0]\to\mathbb{R}^n} \left(\frac{dy(x)}{dx} = f(y(x)) \wedge y(0) = y_0 \right).$$

Recall that $y(x)$ is computed as a continuous function, which includes a Cauchy sequence representation of its values together with a (uniform) Cauchy modulus (depending on y_0 only). So without additional effort we have some enclosure information.

Quality? At least it allows us to compute for any initial value y_0 in $[a, b]$ and any given point x in $[-a_0, a_0]$ an interval containing the value of the solution for y_0 at the point x.

Moore's first-order method

$\sim$ Cauchy-Euler existence proof: We have a constructive proof of

$$\forall_{y_0} \exists_{a_0>0} \exists_{y(x)\colon [-a_0,a_0]\to\mathbb{R}^n} \left(\frac{dy(x)}{dx} = f(y(x)) \wedge y(0) = y_0 \right).$$

By the Lipschitz condition we can find an a_0 which works uniformly for all y_0 in some $[a, b]$, hence

$$\forall_{y_0\in[a,b]} \exists_{y(x)\colon [-a_0,a_0]\to\mathbb{R}^n} \left(\frac{dy(x)}{dx} = f(y(x)) \wedge y(0) = y_0 \right).$$

Recall that $y(x)$ is computed as a continuous function, which includes a Cauchy sequence representation of its values together with a (uniform) Cauchy modulus (depending on y_0 only). So without additional effort we have some enclosure information.

Quality? At least it allows us to compute for any initial value y_0 in $[a, b]$ and any given point x in $[-a_0, a_0]$ an interval containing the value of the solution for y_0 at the point x.

Moore's first-order method

$\sim$ Cauchy-Euler existence proof: We have a constructive proof of

$$\forall_{y_0} \exists_{a_0>0} \exists_{y(x)\colon [-a_0,a_0]\to\mathbb{R}^n} \left(\frac{dy(x)}{dx} = f(y(x)) \wedge y(0) = y_0 \right).$$

By the Lipschitz condition we can find an a_0 which works uniformly for all y_0 in some $[a, b]$, hence

$$\forall_{y_0\in[a,b]} \exists_{y(x)\colon [-a_0,a_0]\to\mathbb{R}^n} \left(\frac{dy(x)}{dx} = f(y(x)) \wedge y(0) = y_0 \right).$$

Recall that $y(x)$ is computed as a continuous function, which includes a Cauchy sequence representation of its values together with a (uniform) Cauchy modulus (depending on y_0 only). So without additional effort we have some enclosure information.

Quality? At least it allows us to compute for any initial value y_0 in $[a, b]$ and any given point x in $[-a_0, a_0]$ an interval containing the value of the solution for y_0 at the point x.

Moore's first-order method

$\sim$ Cauchy-Euler existence proof: We have a constructive proof of

$$\forall_{y_0} \exists_{a_0>0} \exists_{y(x)\colon [-a_0,a_0]\to\mathbb{R}^n} \left(\frac{dy(x)}{dx} = f(y(x)) \wedge y(0) = y_0 \right).$$

By the Lipschitz condition we can find an a_0 which works uniformly for all y_0 in some $[a, b]$, hence

$$\forall_{y_0\in[a,b]} \exists_{y(x)\colon [-a_0,a_0]\to\mathbb{R}^n} \left(\frac{dy(x)}{dx} = f(y(x)) \wedge y(0) = y_0 \right).$$

Recall that $y(x)$ is computed as a continuous function, which includes a Cauchy sequence representation of its values together with a (uniform) Cauchy modulus (depending on y_0 only). So without additional effort we have some enclosure information.

Quality? At least it allows us to compute for any initial value y_0 in $[a, b]$ and any given point x in $[-a_0, a_0]$ an interval containing the value of the solution for y_0 at the point x.

Moore's Kth-order method

Write equation (2) in component form, and assume that the $f_i(y) = f_i(y_1, \ldots, y_n)$ have continuous total derivatives. Then

$$\frac{dy(x)}{dx} = f(y(x)) =: f^{(0)}(y(x)),$$

$$\frac{d^2y(x)}{dx^2} = \frac{d}{dx} f^{(0)}(y(x)) = \sum_{m=1}^{n} \frac{\partial f_i^{(0)}}{\partial y_m} f_m^{(0)} =: f^{(1)}(y(x)),$$

$$\frac{d^3y(x)}{dx^2} = \frac{d}{dx} f^{(1)}(y(x)) = \sum_{m=1}^{n} \frac{\partial f_i^{(1)}}{\partial y_m} f_m^{(0)} =: f^{(2)}(y(x))$$

etc., where $f_i^{(j)}$ stands for $f_i^{(j)}(y_1(x), \ldots, y_n(x))$.

Taylor's theorem

Generally, for $j = 1, 2, \ldots,$

$$\frac{d^j y(x)}{dx^j} = f_i^{(j-1)}(y_1(x), \ldots, y_n(x)).$$

Now apply Taylor's theorem with Lagrangian rest. Then for sufficiently small h,

$$y_i(x) = y_i(0) + \sum_{j=1}^{K-1} \frac{f_i^{(j-1)}(y_1(0), \ldots, y_n(0))}{j!} x^j + \frac{f_i^{(K-1)}(y_1(\vartheta), \ldots, y_n(\vartheta))}{K!} x^K \tag{4}$$

for some $\vartheta \in [0, h]$.

Idea of Moore's Kth-order method

(4) used as a basis for enclosure arguments. Feature of (4): the unknown ϑ only appears in a product with a factor x^K, which is small if x is small and K is big. Given an approximation A of the initial value $y(0)$ so that $Y([0,h]) \subset A$, some interval estimates become possible. By clever heuristics Moore chooses h and A so that $w(Y(x))$ is "as small as possible for as long as possible".

One has to assume (among other things) that the interval versions $F^{(j)}$ of $f^{(j)}$ satisfy a Lipschitz condition

$$w(F^{(j)}(Y)) \leq L_j w(Y), \tag{5}$$

where $F^{(j)}(Y)$ denotes the interval-vector-valued function with components $F^{(j)}(Y) = (F_1^{(j)}(Y), \ldots, F_n^{(j)}(Y))$, and $w(I)$ is the width of the interval I.

Alternative

It is possible to use the constructive proof of (4) directly as a source for estimates. This avoids

- the bad effect of interval computations with many occurrences of the same expression ($y(0)$ in our case), and
- makes the somewhat stiff assumption (5) superfluous.

Conclusion

- Constructive analysis with witnesses of low type level. Type level 1 representation of continuous functions.
- Extraction of reasonable programs is possible.
- The Cauchy-Euler construction of approximate solutions to ODEs as a type level 1 process.
- Moore's ideas for enclosure of solutions can be transferred to the present setting, with Cauchy sequences instead of intervals.

Extraction from classical proofs

- Minimal and intuitionistic arithmetic
- Realizability
- The Dragalin-Friedman A-translation
- Example: *hsh* is not the identity
- Example: Wellfoundedness of $\mathbb{N}$
- Example: Dickson's lemma

Goal

- A refined method of extracting reasonable and sometimes unexpected programs from classical proofs.
- In $\forall_x \tilde{\exists}_y B(x, y)$, the kernel $B(x, y)$ need not be quantifier-free, but only has to belong to the strictly larger class of goal formulas.
- Allow unproven lemmas D in the proof of $\forall_x \tilde{\exists}_y B(x, y)$, where D is a definite formula.

Idea

Transform a proof of $\tilde{\exists}_y G$ (G quantifier-free) into a proof of $\exists_y G$.

Simple idea: replace $\bot$ anywhere in the proof by $\exists_y G$. Then $\forall_y(G \to \bot) \to \bot$ is turned into $\forall_y(G \to \exists_y G) \to \exists_y G$, and since the premise is trivially provable, we have the claim.

Unfortunately, this simple argument is not quite correct.

- G may contain $\bot$, and hence is changed under the substitution $\bot \mapsto \exists_y G$.
- We may have used axioms or lemmata involving $\bot$ (e.g., $\bot \to P$), which need not be derivable after the substitution.

However, the simple idea can be turned into something useful.

Lemmata

Given a derivation (in minimal logic) of $\tilde{\exists}_y G$ from $\vec{D}$ and axioms

$$\begin{aligned}
&\mathrm{Ind}_{n,A}: && A(0) \to \forall_n(A(n) \to A(n+1)) \to \forall_n A(n),\\
&\mathrm{Ind}_{p,A}: && P(\mathrm{tt}) \to P(\mathrm{ff}) \to \forall p A(p),\\
&\mathrm{Ax}_{\mathrm{tt}}: && \mathrm{atom}(\mathrm{tt}),\\
&\mathrm{Efq}_A: && \mathrm{atom}(\mathrm{ff}) \to A.
\end{aligned}$$

atom is a unary predicate symbol taking one argument of the type **B** of booleans. The intended interpretation of atom is the set $\{\mathrm{tt}\}$; hence "$\mathrm{atom}(r)$" means "$r = \mathrm{tt}$". Assume that the lemmata $\vec{D}$ and the goal formula G are such that (with $B^A := B[\bot := A]$)

$$\vdash_i \vec{D} \to D_i^{\exists_y G}, \tag{6}$$

$$\vdash_i G^{\exists_y G} \to \exists_y G; \tag{7}$$

here $\vdash_i$ means derivability in intuitionistic arithmetic, i.e., with the additional axioms $\mathrm{Efq}\text{-}\mathrm{Log}_A: \bot \to A$.

Lemmata (continued)

The substitution $\bot \mapsto \exists_y G$ turns the axioms (except $\mathrm{Efq\text{-}Log}_A$) into new instances of the same scheme. Hence from our given derivation (in minimal logic) of $\vec{D} \to \forall_y(G \to \bot) \to \bot$ we obtain

$$\vdash_i \vec{D}^{\exists_y G} \to \forall_y(G^{\exists_y G} \to \exists_y G) \to \exists_y G.$$

Now (6) allows to drop the substitution in $\vec{D}$, and by (7) the second premise is derivable. Hence

$$\vdash_i \vec{D} \to \exists_y G.$$

We identify classes of formulas – to be called definite and goal formulas – such that slight generalizations of (6) and (7) hold.

Obtain an explicit representation of the term extracted (via realizability) from the derivation M of $\vec{D} \to \exists_y G$ just constructed, of the form $pt_1 \dots t_n s$, where p is extracted from M and $t_1, \dots, t_n, s$ are determined by the formulas $\vec{D}$ and G only.

Terms and formulas

Constants: $\varepsilon^{\mathbf{U}},\ \mathrm{tt}^{\mathbf{B}},\ \mathrm{ff}^{\mathbf{B}},\ \mathrm{nil}_\rho,\ \mathrm{cons}_\rho,\ \mathcal{R}_\tau^{\mathbf{L}(\rho)},\ \mathbf{C}_\tau^{\mathbf{L}(\rho)}$.

Terms: $c^\rho,\ x^\rho,\ (\lambda x^\rho r^\sigma)^{\rho\to\sigma},\ (r^{\rho\to\sigma} s^\rho)^\sigma$.

Formulas: $\bot,\ X,\ \mathrm{atom}(t^{\mathbf{B}}),\ A \to B,\ \forall_{x^\rho} A$.

Abbreviations: $\neg A := A \to \bot$, $\tilde{\exists}_x A := \neg\forall_x \neg A$.
Further predicate symbols P might be added.

Axioms for Z_0^X (minimal arithmetic)

Induction axiom $\mathrm{Ind}_{l,A}$:

$$A(\mathrm{nil}_\rho) \to \forall_{x^\rho, l^{\mathsf{L}(\rho)}}(A(l) \to A(\mathrm{cons}_\rho(x, l))) \to \forall_{l^{\mathsf{L}(\rho)}} A(l).$$

Cases axiom $\mathrm{Cases}_{l,A}$:

$$A(\mathrm{nil}_\rho) \to \forall_{x^\rho, l^{\mathsf{L}(\rho)}} A(\mathrm{cons}_\rho(x, l)) \to \forall_{l^{\mathsf{L}(\rho)}} A(l),$$
$$A(\mathrm{tt}) \to A(\mathrm{ff}) \to \forall_{p^{\mathsf{B}}} A(p).$$

Truth and falsity axioms:

$$\mathrm{Ax}_{\mathrm{tt}}: \quad \mathrm{atom}(\mathrm{tt}),$$
$$\mathrm{Ax}_{\mathrm{ff},A}: \quad \mathrm{atom}(\mathrm{ff}) \to A.$$

Axioms for Z^X (intuitionistic arithmetic)

In addition ex-falso-quodlibet:

$$\text{Efq-Log}_A\colon \quad \bot \to A$$

Notice that every instance $\bot \to A$ of ex-falso-quodlibet is derivable from

$$\bot \to X,$$
$$\bot \to \mathrm{atom}(\mathsf{ff}).$$

Derivations in minimal logic

$$
\begin{array}{l}
u^B \quad \text{(assumptions)} \mid \text{axioms} \mid \\
(\lambda u^A M^B)^{A\to B} \mid (M^{A\to B} N^A)^B \mid \\
(\lambda x M^A)^{\forall_x A} \mid (M^{\forall_x A(x)}\, t)^{A(t)}
\end{array}
$$

where in $\lambda x M^A$, $x \notin \mathrm{FV}(B)$ for all B with $u^B \in \mathrm{FA}(M)$.

$\vdash$ denotes derivability in minimal logic.

Computational content $\tau(A)$ (type or ε)

$$
\begin{aligned}
\tau(X) &:= \nu, \\
\tau(P(\vec{s})) &:= \varepsilon \quad (\text{in particular, } \tau(\bot) = \varepsilon), \\
\tau(\forall_{x^\rho} A) &:= \begin{cases} \varepsilon & \text{if } \tau(A) = \varepsilon \\ \rho \to \tau(A) & \text{otherwise,} \end{cases} \\
\tau(A \to B) &:= \begin{cases} \tau(B) & \text{if } \tau(A) = \varepsilon \\ \varepsilon & \text{if } \tau(B) = \varepsilon \\ \tau(A) \to \tau(B) & \text{otherwise.} \end{cases}
\end{aligned}
$$

Extracted term $[\![M]\!]$: $\tau(A)$ for $\tau(A) \neq \varepsilon$

$$
\begin{aligned}
[\![u^A]\!] &:= x_u^{\tau(A)} \quad (x_u^{\tau(A)} \text{ associated with } u^A) \\
[\![\lambda u^A M]\!] &:= \begin{cases} [\![M]\!] & \text{if } \tau(A) = \varepsilon \\ \lambda x_u^{\tau(A)} [\![M]\!] & \text{otherwise} \end{cases} \\
[\![M^{A \to B} N]\!] &:= \begin{cases} [\![M]\!] & \text{if } \tau(A) = \varepsilon \\ [\![M]\!][\![N]\!] & \text{otherwise} \end{cases} \\
[\![\lambda x^\rho M]\!] &:= \lambda x^\rho [\![M]\!] \\
[\![Mt]\!] &:= [\![M]\!] t
\end{aligned}
$$

Extracted terms for axioms

For the induction axiom $\mathrm{Ind}_{I,A}$:

$$A(\mathrm{nil}_\rho) \to \forall_{x^\rho, I^{\mathbf{L}(\rho)}}(A(I) \to A(\mathrm{cons}_\rho(x, I))) \to \forall_{I^{\mathbf{L}(\rho)}} A(I).$$

$[\![\mathrm{Ind}_{I,A}]\!] := \mathcal{R}_\tau^{\mathbf{L}(\rho)} : \tau \to (\rho \to \mathbf{L}(\rho) \to \tau \to \tau) \to \mathbf{L}(\rho) \to \tau$
Similarly: $[\![\mathrm{Cases}_{I,A}]\!] := \mathbf{c}_\tau^{\mathbf{L}(\rho)}$ (with $\tau := \tau(A) \neq \varepsilon$).

Let $[\![\text{Efq-Log}_X]\!] := a_0^\nu$.

For M^A with $\tau(A) = \varepsilon$ (i.e. A without computational content) define $[\![M]\!] := \varepsilon$ (some new symbol).

Realizability

Fix a comprehension term $\mathcal{A} = \{\, y \mid A_0 \,\}$, A_0 without X.

$$
\begin{aligned}
r \ \mathrm{mr}_{\mathcal{A}}\ X &:= \mathcal{A}(r) \\
r \ \mathrm{mr}_{\mathcal{A}}\ P(\vec{s}) &:= P(\vec{s}) \\
r \ \mathrm{mr}_{\mathcal{A}}\ \forall_x A &:= \begin{cases} \forall_x(\varepsilon \ \mathrm{mr}_{\mathcal{A}}\ A) & \text{if } \tau(A) = \varepsilon \\ \forall_x(rx \ \mathrm{mr}_{\mathcal{A}}\ A) & \text{otherwise} \end{cases} \\
r \ \mathrm{mr}_{\mathcal{A}}\ A{\to}B &:= \begin{cases} \varepsilon \ \mathrm{mr}_{\mathcal{A}}\ A \ \to \ r \ \mathrm{mr}_{\mathcal{A}}\ B & \text{if } \tau(A){=}\varepsilon \\ \forall_x(x \ \mathrm{mr}_{\mathcal{A}}\ A \ \to \ \varepsilon \ \mathrm{mr}_{\mathcal{A}}\ B) & \text{if } \tau(A){\neq}\varepsilon{=}\tau(B) \\ \forall_x(x \ \mathrm{mr}_{\mathcal{A}}\ A \ \to \ rx \ \mathrm{mr}_{\mathcal{A}}\ B) & \text{otherwise} \end{cases}
\end{aligned}
$$

Notice: $\tau(A) = \varepsilon$ and $\varepsilon \ \mathrm{mr}_{\mathcal{A}}\ A = A$, for A without X.

Soundness

Let $x_u^{\tau(A)} := \varepsilon$ in case $\tau(A) = \varepsilon$.

Theorem

Assume that M is a Z^X-derivation of B. Then there is a Z-derivation of

$$[\![M]\!]\,\mathrm{mr}_{\mathcal{A}}\,B$$

from the assumptions $\{\, x_u^{\tau(C)}\ \mathrm{mr}_{\mathcal{A}}\ C \mid u^C \in \mathrm{FA}(M)\,\}$.

Proof.

Induction on M. □

Definite and goal formulas

A formula C is relevant if it "ends" with $\bot$:

$$C := \bot \mid B \to C \mid \forall_x C$$

Define definite and goal formulas D, G simultaneously by

$$\begin{aligned} D &:= P \mid G \to D && \text{if } G \text{ relevant} \Rightarrow D \text{ relevant} \\ &\quad \mid \forall_x D \\ G &:= P \mid D \to G && \text{if } D \text{ qfree or relevant} \\ &\quad \mid \forall_x G && \text{if } G \text{ irrelevant} \end{aligned}$$

Write B^A for $B[\bot := A]$.

Translation

Lemma
$Z^X \vdash D \to D^X$ *and* $Z^X \vdash (\vec{G} \to X) \to \vec{G}^X \to X$.

Theorem (Translation)
If $Z_0 \vdash \vec{D} \to \forall_{\vec{y}}(\vec{G} \to \bot) \to \bot$, *then* $Z^X \vdash \vec{D} \to \forall_{\vec{y}}(\vec{G} \to X) \to X$.
Hence $(X := \exists_{\vec{y}}\vec{G})$

$$Z \vdash \vec{D} \to \exists_{\vec{y}}\vec{G}.$$

Proof.
$Z_0^X \vdash \vec{D}^X \to \forall_{\vec{y}}(\vec{G}^X \to X) \to X$ (replace $\bot$ by X). Now use the Lemma. □

Open Problem. Characterize $\{ A \mid Z^X \vdash A \to A^X \}$.

A non-example (Kreisel)

$\vdash \forall_x \tilde{\exists}_y A$ generally does not yield a program to compute y from x:
Clearly $\vdash \forall_x(\forall_y Txxy \to \forall_z Txxz)$, hence

$$\vdash \forall_x \tilde{\exists}_y(Txxy \to \forall_z Txxz).$$

However, there is no computable f satisfying

$$Txx(fx) \to \forall_z Txxz,$$

for then $Txx(fx) \leftrightarrow \forall_z Txxz$.

Definite & goal formulas: examples

$$\vdash \forall_z(\underbrace{\neg\neg Txxz \to Txxz}_{\text{not definite}}) \to \forall_y\big((\underbrace{Txxy \to \forall_z Txxz}_{\text{goal}}) \to \bot\big) \to \bot.$$

Replace T by $\neg S$:

$$\vdash \big(\forall_y(\neg Sxxy \to \underbrace{\forall_z \neg Sxxz}_{\text{not goal}}) \to \bot\big) \to \bot.$$

The Dragalin-Friedman A-translation

- Insert $\neg\neg$ at every atom ($\neq \bot$): all formulas are relevant.
- Then all qfree formulas are definite and goal formulas; all Π^0_1-formulas are definite.
- Problem (cf. Murthy's case study for Higman's lemma): too many negations.

Example 1: *hsh* is not the identity

We show $\tilde{\exists}_n\, h(s(hn)) \neq n$ and extract an (unexpected) program from it (due to U.Berger).

Surjectivity Lemma. $g \circ f$ surjective implies g surjective.

Injectivity Lemma. $g \circ f$ injective implies f injective.

Surjectivity-Injectivity Lemma.
$g \circ f$ surjective and g injective implies f surjective.

Example 1: *hsh* is not the identity (ctd.)

***hsh*-Theorem.** $\forall_n\, s(n) \neq 0 \rightarrow \neg\forall_n\, h(s(h(n))) = n$.

Proof.

Assume $h \circ s \circ h$ is the identity.

h injective	by the Injectivity Lemma
$s \circ h$ surjective	by the Surjectivity-Injectivity Lemma
s surjective	by the Surjectivity Lemma

Contradiction. □

Example 1: *hsh* is not the identity (ctd.)

hsh-**Theorem-dn.** $\forall_n\, s(n) \neq 0 \to \tilde{\exists}_n\, h(s(h(n))) \neq n.$

$\forall_n\, s(n) \neq 0$ is a definite formula.

General theory applies: the proof contains an algorithm.
Which one?

Extracted term

```
[s,h][if (h(s(h(h 0)))=h 0)
         [if (h(s(h(s(h(h 0)))))=s(h(h 0)))
             0
             (s(h(h 0)))]
         (h 0)]
```

If $h(s(h(h0))) \neq h0$, take $h0$. Assume $h(s(h(h0))) = h0$. If $h(s(h(s(h(h0))))) = s(h(h0))$, then also $h(s(h0)) = s(h(h0))$; take 0 (using our assumption on s). Assume $h(s(h(s(h(h0))))) \neq s(h(h0))$. Take $s(h(h0))$.

Example 2: Wellfoundedness of $\mathbb{N}$

$$\forall_{f^{\mathbf{N}\to\mathbf{N}}}\tilde{\exists}_k(f(k+1) < f(k) \to \bot).$$

Classical proof: "choose k such that $f(k)$ is minimal".

But: this k cannot possibly be computed.

So what is the extracted algorithm?

Minimum principle

$$\tilde{\exists}_k\, R(k) \to \tilde{\exists}_k(R(k) \land \forall_{l<k}(R(l) \to \bot)).$$

This is logically equivalent to

$$\forall_k\big(R(k) \to \forall_{l<k}(R(l) \to \bot) \to \bot\big) \to \forall_k(R(k) \to \bot).$$

Premise: progressiveness of $R(k) \to \bot$ w.r.t. $<$

$$\mathrm{Prog} := \forall_k\big(\forall_{l<k}(R(l) \to \bot) \to R(k) \to \bot\big).$$

Prove this by zero-successor-induction on n w.r.t.

$$B(n) := \forall_{k<n}(R(k) \to \bot).$$

Extracted program

```
[f][if (f 1<f 0)
       ((Rec nat=>nat=>nat)
        ([n1]0)
        ([n1,f2,n3][if (f(Succ n3)<f n3)
                       (f2(Succ n3))
                       n3])
        (f 0)
        1)
       0]
```

More readable description of the algorithm

Discussion: `Rec` defines a function h of type $\mathbf{N} \to \mathbf{N} \to \mathbf{N}$. After an initial check whether $f(1) < f(0)$, in the positive case, apply h to $f(0)$ and 1.

```
Point-of-increase(f) := [if (f 1<f 0) (h(f 0)1) 0]
where

  h 0     := [m] 0
  h(n+1) := [m] [if (f(m+1)<f m) (h n(m+1)) m]
```

Example 3: Dickson's lemma

For every k, l,

$$\forall_{f_1,\ldots,f_k} \tilde{\exists}_{i_0,\ldots,i_l} \bigwedge\!\!\!\!\bigwedge_{\lambda<l} \big(i_\lambda < i_{\lambda+1} \wedge \bigwedge\!\!\!\!\bigwedge_{\kappa=1}^{k} f_\kappa(i_\lambda) \leq f_\kappa(i_{\lambda+1})\big).$$

Applications in algebra, combinatorics, Petri net theory.

Proof uses the minimum principle for undecidable sets.
No obvious computational content.

Proof of Dickson's lemma

$$\forall_{f_1,\dots,f_k} \tilde{\exists}_{i_0,\dots,i_l} \bigwedge_{\lambda<l} \big(i_\lambda < i_{\lambda+1} \wedge \bigwedge_{\kappa=1}^{k} f_\kappa(i_\lambda) \le f_\kappa(i_{\lambda+1})\big).$$

Proof from minimum principle w.r.t. a measure function.
Call $Q \subseteq \mathbb{N}$ unbounded if $\forall_x \tilde{\exists}_{y>x} Q(y)$.

Lemma 1. Let Q be unbounded and $f : \bar{Q} \supseteq Q \to \mathbb{N}$. Then the set Q_f of left f-minima w.r.t. Q is unbounded

$$Q_f(x) := Q(x) \wedge \forall_{y>x}(Q(y) \to f(x) \le f(y)).$$

Proof of Lemma 1. Given x, we must find $y > x$ with $Q_f(y)$. Minimum principle for $\{\, y > x \mid Q(y) \,\}$ with measure f:

$$\tilde{\exists}_{y>x} Q(y) \to \tilde{\exists}_{y>x}(Q(y) \wedge \forall_{z>x}(Q(z) \to f(y) \leq f(z))).$$

Since Q is unbounded, the premise is true. We show that the y provided by the conclusion satisfies $Q_f(y)$, i.e.,

$$Q(y) \wedge \forall_{z>y}(Q(z) \to f(y) \leq f(z)).$$

Let $z > y$ with $Q(z)$ be given. From $y > x$ obtain $z > x$, hence $f(y) \leq f(z)$.

Let Q be unbounded and $f_0, f_1 \ldots$ be functions from a superset of Q to $\mathbb{N}$. Then for every k there is an unbounded subset Q_k of Q such that $f_0, \ldots, f_{k-1}$ increase on Q_k w.r.t. Q.

Lemma 2.

$$\forall_x \tilde{\exists}_{y>x} Q(y) \rightarrow \forall_k \tilde{\exists}_{Q_k \subseteq Q}$$
$$\big(\forall_x \tilde{\exists}_{y>x} Q_k(y) \wedge \forall_{i<k} \forall_{x,y;x<y} (Q_k(x) \rightarrow Q(y) \rightarrow f_i(x) \leq f_i(y))\big).$$

Proof. Induction on k. *Base* Let $Q_0 := Q$. *Step*. Consider $(Q_k)_{f_k}$. By IH, $f_0, \ldots, f_{k-1}$ increase on Q_k w.r.t. Q, and therefore also on its subset $(Q_k)_{f_k}$. Moreover, by construction also f_k increases on $(Q_k)_{f_k}$ w.r.t. Q.

Dickson's lemma: Extracted term

Note: Recursion parameters serve as upper bounds only. Reason: induction via minimum principle. Hence $\mathcal{R} \mapsto$

$$\mathcal{R}^{\mathrm{gen}}_{\tau} : (\mathbf{N} \to \mathbf{N} \to \tau) \to \mathbf{N} \to \tau, \quad \mathcal{R}^{\mathrm{gen}} hx = hx(\mathcal{R}^{\mathrm{gen}} h).$$

Extracted program: $\varphi(0)$, where

$$\begin{aligned} \varphi(i) &= \psi(i, \varphi) \\ \psi(i, h) &= \xi_{i,h}(i+1) \\ \xi_{i,h}(j) &= \begin{cases} \psi(j, \xi_{i,h}) & \text{if } g(j) < g(i) \\ h(j) & \text{if not \& } f(j) < f(i) \\ (i, j) & \text{otherwise} \end{cases} \end{aligned}$$

Dialectica interpretation

- Gödel translation, soundness
- A unified view of realizability and Dialectica interpretation
- Extraction of uniform bounds
- The negative fragment: classical arithmetic
- Extraction of uniform bounds from classical proofs with extensionality

Gödel translation $A \mapsto \exists_x \forall_y |A|^x_y$

In his original functional interpretation, Gödel (1958) assigned to every formula A a new one $\exists_x A_1(x)$ with $A_1(x)$ a universal formula, i.e., of the form $\forall_y |A|^x_y$ with $|A|^x_y$ quantifier-free.

To determine the types of x and y, we assign to every formula A types $\tau^+(A)$, $\tau^-(A)$. $\tau^+(A)$ is intended to be the type of a (Dialectica-)realizer to be extracted from a proof of A, and $\tau^-(A)$ the type of a challenge for the claim that this term realizes A.

Rather than including amongst the types a special "nulltype" object ε and case distinctions – as we did for realizability –, it is more convenient here to use the unit type $\mathbf{U}$ instead and so avoid case distinctions. Using some obvious isomorphisms (like $(\rho \to \mathbf{U}) \cong \mathbf{U}$ and $(\mathbf{U} \to \rho) \cong \rho$) we can later "clean" such types.

Positive and negative types

$$\begin{aligned}
\tau^+(P(\vec{s})) &:= \mathbf{U}, & \tau^-(P(\vec{s})) &:= \mathbf{U},\\
\tau^+(\forall_{x^\rho} A) &:= \rho \to \tau^+(A), & \tau^-(\forall_{x^\rho} A) &:= \rho \times \tau^-(A),\\
\tau^+(\exists_{x^\rho} A) &:= \rho \times \tau^+(A), & \tau^-(\exists_{x^\rho} A) &:= \tau^-(A).
\end{aligned}$$

and for implication

$$\begin{aligned}
\tau^+(A \to B) &:= \big(\tau^+(A) \to \tau^+(B)\big) \times \big(\tau^+(A) \to \tau^-(B) \to \tau^-(A)\big),\\
\tau^-(A \to B) &:= \tau^+(A) \times \tau^-(B).
\end{aligned}$$

Example

$$\begin{aligned}
\tau^+(\forall_n \exists_k G(n,k)) &:= \mathbf{N} \to \mathbf{N} \times \mathbf{U} \mapsto_{\text{clean}} \mathbf{N} \to \mathbf{N},\\
\tau^-(\forall_n \exists_k G(n,k)) &:= \mathbf{N} \times \mathbf{U} \mapsto_{\text{clean}} \mathbf{N}.
\end{aligned}$$

Gödel translation

$$
\begin{aligned}
|P(\vec{s}\,)|^r_s &:= P(\vec{s}\,),\\
|\forall_x A(x)|^r_s &:= |A(s0)|^{r(s0)}_{s1},\\
|\exists_x A(x)|^r_s &:= |A(r0)|^{r1}_{s},\\
|A \to B|^r_s &:= |A|^{s0}_{r1(s0)(s1)} \to |B|^{r0(s0)}_{s1}.
\end{aligned}
$$

For readability we sometimes write terms of a pair type in pair form. Then

$$
\begin{aligned}
|\forall_z A|^x_{z,y} &:= |A|^{xz}_{y},\\
|\exists_z A|^{z,x}_{y} &:= |A|^{x}_{y},\\
|A \to B|^{f,g}_{x,u} &:= |A|^{x}_{gxu} \to |B|^{fx}_{u}.
\end{aligned}
$$

Markov principle (M^ω), for higher type variables and quantifier-free formulas A_0, B_0:

$$(\forall_y A_0 \to B_0) \to \exists_y (A_0 \to B_0) \quad (y \notin \mathrm{FV}(B_0)).$$

Theorem (Soundness)

Let M be a derivation in

$$\mathrm{WE\text{-}HA}^\omega + \mathrm{AC} + \mathrm{IP}^\omega_\forall + \mathrm{M}^\omega + \mathrm{Ax}_\forall \text{ of a formula } A$$

from assumptions $u_i \colon C_i$ $(i = 1, \ldots, n)$. Let x_i of type $\tau^+(C_i)$ be variables for realizers of the assumptions, and y be a variable of type $\tau^-(A)$ for a challenge of the goal. Then we can find terms $[\![M]\!]^+ =: t$ of type $\tau^+(A)$ with $y \notin \mathrm{FV}(t)$ and $[\![M]\!]^-_i =: r_i$ of type $\tau^-(C_i)$, and a derivation $\mu(M)$ in

$$\mathrm{WE\text{-}HA}^\omega + \mathrm{Ax}_\forall \text{ of the formula } |A|^t_y$$

from assumptions $\bar{u}_i \colon |C_i|^{x_i}_{r_i}$.

Theorem (Characterization)

$$\mathrm{AC} + \mathrm{IP}^{\omega}_{\forall} + \mathrm{M}^{\omega} \vdash A \leftrightarrow \exists_x \forall_y |A|^x_y.$$

Proof.

Induction on A; we only treat one case.

$$\begin{aligned}
(A \to B) &\leftrightarrow (\exists_x \forall_y |A|^x_y \to \exists_v \forall_u |B|^v_u) && \text{by IH} \\
&\leftrightarrow \forall_x (\forall_y |A|^x_y \to \exists_v \forall_u |B|^v_u) && \text{by } \mathrm{ML}^{\omega} \\
&\leftrightarrow \forall_x \exists_v (\forall_y |A|^x_y \to \forall_u |B|^v_u) && \text{by } (\mathrm{IP}^{\omega}_{\forall}) \\
&\leftrightarrow \forall_x \exists_v \forall_u (\forall_y |A|^x_y \to |B|^v_u) && \text{by } \mathrm{ML}^{\omega} \\
&\leftrightarrow \forall_x \exists_v \forall_u \exists_y (|A|^x_y \to |B|^v_u) && \text{by } (\mathrm{M}^{\omega}) \\
&\leftrightarrow \exists_f \forall_x \forall_u \exists_y (|A|^x_y \to |B|^{fx}_u) && \text{by } (\mathrm{AC}) \\
&\leftrightarrow \exists_{f,g} \forall_{x,u} (|A|^x_{gxu} \to |B|^{fx}_u) && \text{by } (\mathrm{AC}) \\
&\leftrightarrow \exists_{f,g} \forall_{x,u} |A \to B|^{f,g}_{x,u} && \text{by definition.}
\end{aligned}$$

A unified view of realizability and the Dialectica interpretation (Oliva 2006)

Modified realizability can be treated in such a way that similarities with the Dialectica interpretation become visible. Change the definitions of $\tau^+(A)$, $\tau^-(A)$ and $|A|^x_y$ in the implicational case:

$$\begin{aligned}\boldsymbol{\tau}^+(A \to B) &:= \boldsymbol{\tau}^+(A) \to \boldsymbol{\tau}^+(B),\\ \boldsymbol{\tau}^-(A \to B) &:= \boldsymbol{\tau}^+(A) \times \boldsymbol{\tau}^-(B),\end{aligned} \qquad |A \to B|^f_{x,u} := \forall_y |A|^x_y \to |B|^{fx}_u.$$

Then mr can be expressed in terms of the (new) $|A|^x_y$:

$$\vdash r \ \mathrm{mr}\ A \leftrightarrow \forall_y |A|^r_y.$$

Proved by induction on A. *Case* $A \to B$ ($\boldsymbol{\tau}^+(A) \neq \varepsilon$, $\boldsymbol{\tau}^-(A) \neq \varepsilon$):

$$\begin{aligned} r \ \mathrm{mr}\ (A \to B) &\leftrightarrow \forall_x (x \ \mathrm{mr}\ A \to rx \ \mathrm{mr}\ B) && \text{by definition}\\ &\leftrightarrow \forall_x (\forall_y |A|^x_y \to \forall_u |B|^{rx}_u) && \text{by IH}\\ &\leftrightarrow \forall_{x,u} (\forall_y |A|^x_y \to |B|^{rx}_u) && \text{by ML}^\omega\\ &= \forall_{\mathbf{z}} |A \to B|^r_{x,u} && \text{by definition.}\end{aligned}$$

Theorem (Extraction)

Assume

$$\mathsf{WE\text{-}HA}^\omega + \mathrm{AC} + \mathrm{IP}^\omega_\forall + \mathrm{M}^\omega + \mathrm{Ax}_\forall \vdash \forall_x \big(\forall_u A_0(x,u) \to \exists_y B(x,y)\big)$$

with A_0 quantifier-free, and all formulas have at most the displayed variables free. Then we can find a closed HA^ω-term t such that

$$\mathsf{WE\text{-}HA}^\omega + \mathrm{AC} + \mathrm{IP}^\omega_\forall + \mathrm{M}^\omega + \mathrm{Ax}_\forall \vdash \forall_x \big(\forall_u A_0(x,u) \to B(x,tx)\big).$$

Majorization and the Dialectica interpretation

The Dialectica interpretation produces complex extracted terms, as opposed to the realizability interpretation. This is partially due to contraction (necessary in the $\rightarrow^{-}$-rule). Therefore

- consider derivations from lemmata, and
- try to simplify extracted terms by only looking for majorants.

This has led Kohlenbach to develop his "monotone Dialectica interpretation", where one only looks for bounds of realizers rather than exact realizers.
Essential: one can then deal with additional axioms $\mathrm{Ax}_{\forall\exists_{\leq}\forall}$:

$$\forall_{x^\rho}\exists_{y\leq_\sigma rx}\forall_{z^\tau}A_0(x,y,z) \qquad (A_0 \text{ quantifier-free}),$$

with r a closed term of type $\rho\rightarrow\sigma$. Need to consider strenghtened versions $\mathrm{Ax}'_{\forall\exists_{\leq}\forall}$ of these assumptions:

$$\exists_{Y\leq_{\rho\rightarrow\sigma}r}\forall_{x^\rho,z^\tau}A_0(x,Yx,z).$$

Theorem (Soundness with majorants)

Let M be a derivation in

$$\mathrm{WE\text{-}HA}^\omega + \mathrm{AC} + \mathrm{IP}^\omega_\forall + \mathrm{M}^\omega + \mathrm{Ax}_{\forall\exists_\leq\forall} \text{ of a formula } A$$

from assumptions $u_i\colon C_i$ $(i = 1, \ldots, n)$. Let x_i of type $\tau^+(C_i)$ be variables for realizers of the assumptions, and y of type $\tau^-(A)$ be a variable for a challenge of the goal. Let $\vec{z}$ of type $\vec{\rho}$ be the variables free in M. Then we can find closed terms $[\![\lambda\vec{z},\vec{u}M]\!]^{*+} =: T^*$ of type $\tau^+(C_1) \to \ldots \to \tau^+(C_n) \to \vec{\rho} \to \tau^+(A)$ and $[\![\lambda\vec{z},\vec{u}M]\!]^{*-}_i =: R^*_i$ of type $\tau^+(C_1) \to \ldots \to \tau^+(C_n) \to \vec{\rho} \to \tau^-(A) \to \tau^-(C_i)$, and a derivation $\mu(M)$ in

$$\mathrm{WE\text{-}HA}^\omega + \mathrm{Ax}'_{\forall\exists_\leq\forall}$$

of the formula

$$\exists_{T,R_1,\ldots,R_n}\big(T^* \operatorname{maj} T \wedge R^*_1 \operatorname{maj} R_1 \wedge \cdots \wedge R^*_n \operatorname{maj} R_n \wedge \forall_{\vec{x},\vec{z},y}(|C_1|^{x_1}_{R_1\vec{x}\vec{z}y} \to \cdots \to |C_n|^{x_n}_{R_n\vec{x}\vec{z}y} \to |A|^{T\vec{x}\vec{z}}_y)\big).$$

Theorem (Extraction of uniform bounds)

Let s be a closed HA^ω*-term, $A(x,y,z)$ a formula with at most the displayed variables free, and τ a type of level ≤ 2. Assume that*

$$\mathsf{WE\text{-}HA}^\omega + \mathrm{AC} + \mathrm{IP}^\omega_\forall + \mathrm{M}^\omega + \mathrm{Ax}_{\forall\exists_{\leq}\forall} \vdash \forall_{x^1}\forall_{y\leq_\rho sx}\exists_{z^\tau} A(x,y,z).$$

Then we can find a closed HA^ω*-term t such that*

$$\mathsf{WE\text{-}HA}^\omega + \mathrm{AC} + \mathrm{IP}^\omega_\forall + \mathrm{M}^\omega + \mathrm{Ax}_{\forall\exists_{\leq}\forall} \vdash \forall_{x^1}\forall_{y\leq_\rho sx}\exists_{z\leq_\tau tx} A(x,y,z).$$

Moreover, if A contains $\forall\exists$-premises only, then the conclusion can already be derived in $\mathsf{WE\text{-}HA}^\omega + \mathrm{Ax}'_{\forall\exists_{\leq}\forall}$.

The weak Lemma of König as a $\forall\exists_{\leq}\forall$-Axiom

WKL says that every infinite binary tree has an infinite path. When we try to directly formalize it in our (functional) language, it does not quite have the required form, since the assumption that the given tree is infinite needs an additional $\forall$ in the premise.

However, one can easily find an equivalent statement of the required form. To this end, we define the "infinite extension" of a given tree, and let WKL' say that for every t, the infinite extension $I(\hat{t})$ of its "associated tree" $\hat{t}$ has an infinite path. It then is easy to see that WKL and WKL' are equivalent.

Consider hereditary extensional equality, defined as follows

$$x_1 \approx_\mu x_2 \quad := x_1 =_\mu x_2,$$
$$x_1 \approx_{\rho\to\sigma} x_2 := \forall_{y_1,y_2}(y_1 \approx_\rho y_2 \to x_1 y_1 \approx_\sigma x_2 y_2).$$

By definition, $x_1 \approx_1 x_2$ is the same as $x_1 =_1 x_2$.

Lemma

$\vdash x_1 =_\rho x_1' \to x_2 =_\rho x_2' \to x_1 \approx_\rho x_2 \to x_1' \approx_\rho x_2'$.

Lemma

For every HA^ω*-term* t,

$$\mathsf{HA}^\omega \vdash \vec{x}_1 \approx \vec{x}_2 \to t(\vec{x}_1) \approx_\rho t(\vec{x}_2).$$

Corollary

For every closed HA^ω*-term* t, $\mathsf{HA}^\omega \vdash t \approx_\rho t$.

Application: uniform moduli of continuity

Theorem

For every closed HA^ω*-term* t *of type* 2*, we can find another closed* HA^ω*-term* $\bar{t}$ *of* HA^ω *also of type* 2 *such that*

$$\mathsf{HA}^\omega \vdash \forall_{k,y}\forall_{x,x'\leq_1 y}\big(\forall_{i<\bar{t}ky} xi = x'i \to \forall_{j<k} txj = tx'j\big).$$

Proof.

Because of the remark above, from $\mathsf{HA}^\omega \vdash t \approx_2 t$ we obtain

$$\begin{aligned}
&\mathsf{HA}^\omega && \vdash \forall_{x,x'}(\forall_i xi = x'i \to \forall_k \forall_{j<k} txj = tx'j),\\
&\mathsf{HA}^\omega + \mathrm{M}^\omega && \vdash \forall_{k,x,x'}\exists_i(xi = x'i \to \forall_{j<k} txj = tx'j),\\
&\mathsf{HA}^\omega + \mathrm{M}^\omega && \vdash \forall_{k,y}\forall_{x,x'\leq_1 y}\exists_i(xi = x'i \to \forall_{j<k} txj = tx'j).
\end{aligned}$$

Extraction of uniform bounds gives a closed HA^ω-term $\bar{t}$:

$$\mathsf{HA}^\omega \vdash \forall_{k,y}\forall_{x,x'\leq_1 y}\exists_{i\leq\bar{t}ky}(xi = x'i \to \forall_{j<k} txj = tx'j). \qquad \square$$

The negative fragment: classical arithmetic

Classically, we understand an existential formula "there is an x such that $A(x)$" as an abbreviation for "it is not true that for all x, $A(x)$ is false". We propose to make this distinction explicit and use both $\exists_x A$ and $\tilde{\exists}_x A$, where the latter is an abbreviation for $\neg\forall_x\neg A$. Then in a classical context we only deal with $\tilde{\exists}_x A$, and hence need to work with $\to\forall\wedge\bot$-formulas only.
Recall that in arithmetic atomic formulas have the form $\mathrm{atom}(r^{\mathbf{B}})$. There is no need for (logical) falsity $\bot$, since we can take the atomic formula $F := \mathrm{atom}(\mathrm{ff})$ – called *arithmetical falsity* – built from the boolean constant ff instead.
In particular, stability $\neg\neg A \to A$ holds for atomic formulas, and therefore every atomic formula is equivalent to a negated formula. Hence it suffices in classical arithmetic PA^ω to work with $\to\forall\wedge$-formulas only. This implies stability for all formulas.

IP, M and AC with classical existence

We study what happens to the Independence of Premise axiom (IP^ω) and Markov's Principle (M^ω) – both of which involve $\exists$ – under the "negative interpretation" of the existential quantifier, that is, replacement of $\exists$ by $\tilde{\exists}$. It turns out that both become derivable.

Lemma

(a) ($\tilde{\mathrm{IP}}^\omega$) *is derivable from* $F \to A$*:*

$$\vdash (F \to A) \to (A \to \tilde{\exists}_{x^\rho} B) \to \tilde{\exists}_{x^\rho}(A \to B) \quad (x \notin \mathrm{FV}(A)).$$

(b) ($\tilde{\mathrm{M}}^\omega$) *is derivable from* $\forall_{x^\rho}(\neg\neg A \to A)$*:*

$$\vdash \forall_{x^\rho}(\neg\neg A \to A) \to (\forall_{x^\rho} A \to B) \to \tilde{\exists}_{x^\rho}(A \to B) \quad (x \notin \mathrm{FV}(B)).$$

AC with classical existence

Lemma

$(\mathrm{QF\text{-}\tilde{AC}})$ *is derivable from* $(\mathrm{QF\text{-}AC})$ *plus Markov's Principle* (M^ω) *for quantifier-free formulas.*

Proof.

We argue informally. Assume $(\mathrm{QF\text{-}AC})$

$$\forall_{x^\rho} \exists_{y^\sigma} A_0(x, y) \to \exists_{f^{\rho\to\sigma}} \forall_{x^\rho} A_0(x, f(x))$$

with A_0 quantifier-free. Then

$$\begin{array}{ll}
\forall_x \tilde{\exists}_y A_0(x, y) & \\
\forall_x (\forall_y \neg A_0(x, y) \to F) & \\
\forall_x \exists_y (\neg A_0(x, y) \to F) & \text{by } (\mathrm{M}^\omega) \\
\forall_x \exists_y A_0(x, y) & \text{by stability } \neg\neg A_0 \to A_0 \\
\exists_f \forall_x A_0(x, f(x)) & \text{by } (\mathrm{QF\text{-}AC}) \\
\tilde{\exists}_f \forall_x A_0(x, f(x)), &
\end{array}$$

Extraction from classical proofs

Assume

$$\text{WE-PA}^\omega + \text{QF-}\tilde{\text{AC}} + \text{Ax}_\forall \vdash \forall_x \tilde{\exists}_y A_0(x, y),$$

$A_0(x, y)$ a quantifier-free formula with at most the displayed variables free. Then we can find a closed HA^ω-term t such that

$$\text{WE-HA}^\omega + \text{Ax}_\forall \vdash \forall_x A_0(x, tx).$$

Proof:

$$\begin{aligned}
&\text{WE-PA}^\omega + \text{QF-}\tilde{\text{AC}} + \text{Ax}_\forall \vdash \forall_x \tilde{\exists}_y A_0(x, y)\\
&\text{WE-HA}^\omega + \text{QF-AC} + \text{M}^\omega + \text{Ax}_\forall \vdash \forall_x \tilde{\exists}_y A_0(x, y)\\
&\text{WE-HA}^\omega + \text{QF-AC} + \text{M}^\omega + \text{Ax}_\forall \vdash \forall_x \exists_y A_0(x, y) \quad \text{by } (\text{M}^\omega)\\
&\text{WE-HA}^\omega + \text{Ax}_\forall \vdash |\forall_x \exists_y A_0(x, y)|_x^t
\end{aligned}$$

for some closed HA^ω-term t, where in the last step we have used the Soundness Theorem. But

$$|\forall_x \exists_y A_0(x, y)|_x^t = |\exists_y A_0(x, y)|_\varepsilon^{tx} = |A_0(x, tx)|_\varepsilon^\varepsilon = A_0(x, tx).$$

Extraction of uniform bounds from classical proofs

Let s be a closed HA^ω-term, and τ, γ types of level ≤ 2. Assume

$$\mathrm{WE\text{-}PA}^\omega + \mathrm{QF\text{-}}\tilde{\mathrm{AC}} \vdash \forall_{a^\delta} \tilde{\exists}_{b \leq_\sigma ra} \forall_{c^\gamma} B_0(a,b,c) \to \forall_{x^1} \forall_{y \leq_\rho sx} \tilde{\exists}_{z^\tau} A_0(x,y,z).$$

Then we can find a closed HA^ω-term t such that

$$\mathrm{WE\text{-}HA}^\omega \vdash \forall_c \exists_{B \leq_{\delta\to\sigma} r} \forall_{a^\delta} \forall_{c' \leq_\gamma c} B_0(a, Ba, c') \to \forall_{x^1} \forall_{y \leq_\rho sx} \exists_{z \leq_\tau tx} A_0(x,y,z).$$

Why interesting? ε-weakening of the Skolem normal form of the $\forall\exists\forall$-form WKL' of WKL is derivable:

$$\mathrm{WKL}' := \forall_t \exists_{f^{\mathbf{N}\to\mathbf{B}}} \forall_n \bar{f}(n) \in I(\hat{t}).$$

The ε-weakening of its Skolem normal form is

$$\forall_n \exists_F \forall_t \forall_{n' \leq n} \overline{Ft}(n') \in I(\hat{t}).$$

But this is easy to derive (in HA^ω): Given n, let $F_n t$ pick from the infinite tree $I(\hat{t})$ a path of length n.

Elimination of extensionality (Gandy 1956)

$$
\begin{aligned}
E_\mu x &:= \mathrm{tt},\\
(x_1 =^e_\mu x_2) &:= (x_1 =_\mu x_2),\\
E_{\rho\to\sigma} x &:= \forall_{y_1,y_2}(y_1 =^e_\rho y_2 \to xy_1 =^e_\sigma xy_2),\\
(x_1 =^e_{\rho\to\sigma} x_2) &:= E_{\rho\to\sigma}x_1 \wedge E_{\rho\to\sigma}x_2 \wedge \forall_y(E_\rho y \to x_1 y =^e_\sigma x_2 y).
\end{aligned}
$$

Properties (A^E from A by relativizing all quantifiers to E).

- $E_\rho x \to x =^e x$.
- $E_{\rho\to\sigma}x \to E_\rho y \to E_\sigma(xy)$.
- $x_1 =^e_{\rho\to\sigma} x_2 \leftrightarrow \forall_{y_1,y_2}(y_1 =^e_\rho y_2 \to x_1y_1 =^e_\sigma x_2y_2)$.
- $(E_\rho x)^E \leftrightarrow E_\rho x$ and $(x_1 =^e_{\rho\to\sigma} x_2)^E \leftrightarrow (x_1 =^e_{\rho\to\sigma} x_2)$.
- $\vec{x}_1 =^e \vec{x}_2 \to r(\vec{x}_1) =^e r(\vec{x}_2)$.
- $E_\rho x_1 \wedge E_\rho x_2 \wedge (x_1 =_\rho x_2)^E \leftrightarrow x_1 =^e_\rho x_2$.

Theorem

E-HA$^\omega \vdash A(\vec{x})$ *implies* HA$^\omega \vdash E(\vec{x}) \to A^E(\vec{x})$.

Extraction of uniform bounds from classical proofs with extensionality

Theorem (Kohlenbach)

Let Δ be a set of axioms from $\mathrm{Ax}_{\forall\exists_{\leq}\forall}$, and Δ_ε consist of their ε-weakenings. Assume that the types of the existential variables are all ≤ 1 and of the final $\forall$-variables are ≤ 2. Let s be a closed HA^ω-term, $A_0(x,y,z)$ a quantifier-free formula with at most the displayed variables free, and τ a type of level ≤ 2. Assume that

$$\mathsf{E\text{-}PA}^\omega + \mathrm{QF\text{-}\tilde{AC}}^{0,1} + \tilde{\Delta} + \mathrm{W\tilde{K}L} \vdash \forall_{x^1}\forall_{y\leq_1 sx}\tilde{\exists}_{z^\tau} A_0(x,y,z).$$

Then we can find a closed HA^ω-term t such that

$$\mathsf{HA}^\omega + \Delta_\varepsilon \vdash \forall_{x^1}\forall_{y\leq_1 sx}\exists_{z\leq_\tau tx} A_0(x,y,z).$$

Comments

- Obtain constructive existence from classical existence
- Useful for parameter independence
- Terms found by the Dialectica interpretation are complex

Best L_1 approximation

Goal: A classical proof, when analyzed logically, can give quantitative and hence constructive information.

- CSM-spaces, compactness
- Cheney's proof
- Logical analysis

Best L_1-approximation

Theorem

Let $f \in C[0,1]$, given with a modulus ω_f of uniform continuity, and $p_1, p_2 \in P_n$ ($:=$ the set of all polynomials of degree $\leq n$). One can find a modulus of uniqueness $\Psi(\omega_f, n, k)$ for best L_1-approximation, that is,

$$\bigwedge_{i=1,2} \big(\|f - p_i\|_1 - d_1(f, P_n) \leq \Psi(\omega_f, n, \varepsilon)\big) \rightarrow \|p_1 - p_2\|_1 \leq \varepsilon,$$

where $d_1(f, P_n) := \inf_{p \in P_n} \|f - p\|_1$.

Moduli of continuity for polynomials

Lemma (Markov inequality)

Consider $q \in P_n$ *as* $\in C[0,1]$. *Then*

$$\|q'\|_\infty \leq 2 \cdot n^2 \|q\|_\infty.$$

Lemma (Estimate of $\|q\|_\infty$ by $\|q\|_1$)

Consider $q \in P_n$ *as* $\in C[0,1]$. *Then* $\|q\|_\infty \leq 2(n+1)^2\|q\|_1$.

Lemma (Lipschitz constant for $q \in P_n$ in terms of $\|q\|_1$)

Consider $q \in P_n$ *as* $\in C[0,1]$. *Then* $4n^2(n+1)^2\|q\|_1$ *is a Lipschitz constant for* q.

Lemma (Estimate of the coefficients)

Consider $q(x) = a_n x^n + \cdots + a_1 x + a_0$ *as* $\in C[0,1]$. *Then* $\max_i |a_i| \leq n^{2n}\|q\|_\infty$.

CSM-spaces

Let X be a complete separable metric space (CSM-space, or Polish space), given with a countable subset $Q \subseteq X$ and a metric $d\colon Q \to Q \to \mathbb{R}$ with the property that Q is dense in X.

Examples

(a) The set $\mathbb{R}$ of reals is a CSM-space, with the rationals $\mathbb{Q}$ as countable dense subset.

(b) The Baire space $\mathbf{N}^{\mathbf{N}}$ is a CSM-space.

(c) $C[0,1]$, where each function comes with a modulus of uniform continuity. The metric is defined from a norm; for instance $\|f\|_1$ or $\|f\|_\infty$. As the required countable dense subset we can take either the polynomials with rational coefficients (which is dense by a theorem of Weierstraß), or else the set of all rational polygons (i.e., polygons with rational corners).

Compactness

$K \subseteq X$ is compact if it is closed and totally bounded. The latter notion means that for every $k \in \mathbf{N}$ we have a k-net in $Q \cap K$, that is, a finite list $a_{k,0}, \ldots a_{k,h(k)-1}$ of elements of $Q \cap X$ such that for every $a \in Q \cap K$ there is an $i < h(k)$ such that $d(a, a_{k,i}) \leq 2^{-k}$.

Examples

(a) $[0,1]$ is a compact subset of the reals.

(b) The set $\{ f \in \mathbf{N}^{\mathbf{N}} \mid \forall_n f(n) \leq M \}$ and in particular the Cantor space $\{0,1\}^{\mathbf{N}}$ is a compact subset of the Baire space $\mathbf{N}^{\mathbf{N}}$.

(c) For fixed M, n, the set $K_{M,n}$ of all polynomials $p \in P_n$ with $\|p\|_1 \leq M$ is a compact subset of $C[0,1]$. To see this, recall that the coefficients of $p \in P_n$ can be estimated by $n^{2n}\|p\|_\infty$, so (Markov's inequality) by $n^{2n}(n+1)^2\|p\|_1 \leq n^{2n}(n+1)^2 M$.

(d) The set of all $f \in C[0,1]$ with $f(0) = 0$ and a fixed modulus ω of uniform continuity is a compact subset of $C[0,1]$. It is denoted by $C_\omega[0,1]$.

Reduction to a compact subspace of P_n

It suffices to prove the theorem for p_1, p_2 in the compact set $K_{M,n} := \{\, p \in P_n \mid \|p\|_1 \leq M \,\}$, for some $M \geq \frac{5}{2}\|f\|_1$.

Lemma

Assume $\Psi(\omega_f, n, \varepsilon) \leq \frac{\varepsilon}{8}$. *Then*

$$\forall_{p_1,p_2 \in K_{M,n}} \Big(\bigwedge_{i=1,2} \big(\|f - p_i\|_1 \leq d_1(f, P_n) + \Psi(\omega_f, n, \varepsilon) \big) \rightarrow \|p_1 - p_2\|_1 \leq \varepsilon \Big)$$

implies

$$\forall_{p_1,p_2 \in P_n} \Big(\bigwedge_{i=1,2} \big(\|f - p_i\|_1 \leq d_1(f, P_n) + \Psi(\omega_f, n, \varepsilon) \big) \rightarrow \|p_1 - p_2\|_1 \leq \varepsilon \Big).$$

Uniform continuity of g

We want to show that

$$g(x) := |f(x) - p(x)| - \frac{1}{2}\big(|f(x) - p_1(x)| + |f(x) - p_2(x)|\big).$$

is continuous:

$$\forall_\omega \forall_{f \in C_\omega[0,1]; f(0)=0} \forall_{p_1,p_2 \in K_{M,n}} \forall_\varepsilon \forall_{a,b \in [0,1]} \tilde{\exists}_\delta (a <_\delta b \to |g(a) - g(b)| < \varepsilon),$$

with $M := \lceil \frac{1}{\omega(1)} \rceil$. The logical form is

$$\forall_\omega \forall_{f \in C_\omega[0,1]; f(0)=0} \forall_{p_1,p_2 \in K_{M,n}} \forall_\varepsilon \forall_{a,b \in [0,1]} \tilde{\exists}_\delta \tilde{\exists} B_0.$$

Recall that the set of all $f \in C_\omega[0,1]$ with $f(0) = 0$ is compact. Therefore if we would have a derivation, then by the metatheorem we could extract a Dialectica realizer of δ depending on n, ω and ε only.

$$\forall_\omega \forall_{f \in C_\omega[0,1]; f(0)=0} \forall_{p_1,p_2 \in K_{M,n}} \forall_\varepsilon \forall_{a,b \in [0,1]}$$
$$\big(|b - a| < 2^{\Phi_g(n,\omega,\varepsilon)} \to |g(a) - g(b)| < \varepsilon\big).$$

If g is uniformly continuous, ≤ 0 and $\|g\|_1 = 0$, then $g = 0$

$$\forall_\omega \forall_{g \in C_\omega[0,1]} \big(\forall_{a,b \in [0,1]} \forall_{\eta \in \mathbb{Q}^+} (|b-a| \leq \omega(\eta) \to |g(b) - g(a)| \leq \eta) \to \forall_{a \in [0,1]} g(a) \leq 0 \to \int |g| = 0 \to \forall_{x \in [0,1]} g(x) = 0 \big).$$

We make some of the hidden quantifiers explicit:

$$\forall_\omega \forall_{g \in C_\omega[0,1]} \forall_{\varepsilon \in \mathbb{Q}^+} \tilde{\exists}_{\delta \in \mathbb{Q}^+} \big(\forall_{a,b \in [0,1]} \forall_{\eta \in \mathbb{Q}^+} (|b-a| \leq \omega(\eta) \to |g(b) - g(a)| \leq \eta) \to \forall_{a \in [0,1]} g(a) \leq 0 \to \int |g| \leq \delta \to \|g\|_\infty \leq \varepsilon \big).$$

Observe that the purely universal premises do no harm to the applicability of the metatheorem: we can always pull their universal quantifiers as weak existential ones to the front, and disregard bounds provided for them.

If g is uniformly continuous, ≤ 0 and $\|g\|_1 = 0$, then $g = 0$ (continued)

So by the metatheorem we have Φ_∞ such that

$$\begin{aligned}&\forall_\omega \forall_{g \in C_\omega[0,1]} \forall_{\varepsilon \in \mathbb{Q}^+} \big(\\ &\quad \forall_{a,b \in [0,1]} \forall_{\eta \in \mathbb{Q}^+} (|b-a| \leq \omega(\eta) \to |g(b) - g(a)| \leq \eta) \to \\ &\quad \forall_{a \in [0,1]} g(a) \leq 0 \to \int |g| \leq \Phi_\infty(\omega, \varepsilon) \to \|g\|_\infty \leq \varepsilon \big).\end{aligned}$$

If $q \in P_n$ has $n+1$ zeros, then $q = 0$

We write this fact in a form where q is restricted to the compact subset $K_{f,n}$ of P_n, and the hidden quantifiers are made explicit, in order to know that a modulus can be extracted from a proof.

$$\forall_{n\in\mathbb{N}}\forall_{q\in K_{f,n}}\forall_{x_0,\dots,x_n\in[0,1]}\forall_{r,\varepsilon}\tilde{\exists}_{\delta\in\mathbb{Q}^+}$$
$$\big(\forall_{i<n}x_i + r \le x_{i+1} \to \forall_{i\le n}|q(x_i)| \le \delta \to \|q\|_\infty \le \varepsilon\big).$$

By the metatheorem we have Φ_{many} such that

$$\forall_{n\in\mathbb{N}}\forall_{q\in K_{f,n}}\forall_{x_0,\dots,x_n\in[0,1]}\forall_{r,\varepsilon}$$
$$\big(\forall_{i<n}x_i + r \le x_{i+1} \to \forall_{i\le n}|q(x_i)| \le \Phi_{\mathrm{many}}(n,r,\varepsilon) \to \|q\|_\infty \le \varepsilon\big).$$

Using Cheney's Lemma 1 to guarantee $n+1$ roots

Lemma (Cheney)

$$\forall_{f\in C[0,1]}\forall_{n\in\mathbb{N}}\forall_{p_1,p_2\in K_{f,n}}\forall_{x_1\leq\cdots\leq x_n\in[0,1]}\forall_{M\in\mathbb{N}}\forall_{h\in P_n;\|h\|_\infty\leq M}\forall_{\varepsilon,r\in\mathbb{Q}^+}\tilde{\exists}_{l\in\mathbb{Q}^+}\Big($$

$$\forall_{y\in A}|f_0(y)|>\varepsilon\rightarrow\sum_{i=1}^{n+1}\sigma_i\int_{A_i}h>\int_B|h|+1\rightarrow$$

$$\tilde{\exists}_{\lambda\in\mathbb{R}}(\|f_0-\lambda h\|_1+l<\|f_0\|_1)\Big),$$

where

$$\sigma_i:=f_0(\frac{x_{i-1}+x_i}{2})\quad \textit{with } x_0:=0 \textit{ and } x_{n+1}:=1,$$

$$B:=\bigcup_{i=1}^{n}(x_i-\frac{r}{2},x_i+\frac{r}{2}),\quad ,A:=[0,1]\setminus B.$$

Here ε and l have been (vacuously) added in order to obtain useful moduli.

Proof of Cheney's lemma

Notice that A consists of $n+1$ possibly degenerated closed intervals A_i. We can assume that $\mathrm{sg} f$ is defined on A (arbitrarily on degenerated closed intervals A_i). Hence

$$\int_A h \cdot \mathrm{sg} f = \sum_{i=1}^{n+1} \sigma_i \int_{A_i} h > 1 + \int_B |h|.$$

Choose $I := \lambda := \varepsilon / M$. Then on nondegenerated intervals A_i we have $|\lambda h| < \varepsilon < |f|$; hence $\mathrm{sg} f = \mathrm{sg}(f - \lambda h)$. Therefore

Proof of Cheney's lemma (continued)

$$
\begin{aligned}
\int |f-\lambda h| &= \int_B |f-\lambda h| + \int_A |f-\lambda h| \\
&= \int_B |f-\lambda h| + \int_A (f-\lambda h)\mathrm{sg} f \\
&= \int_B |f-\lambda h| + \int_A |f| - \lambda \int_A h\cdot \mathrm{sg} f \\
&= \int_B |f-\lambda h| - \int_B |f| + \int |f| - \lambda \int_A h\cdot \mathrm{sg} f \\
&\leq \lambda \int_B |h| - \lambda \int_A h\cdot \mathrm{sg} f + \int |f| \\
&< -\lambda + \int |f|.
\end{aligned}
$$

Using Cheney's Lemma 1 to guarantee $n+1$ roots

By the metatheorem we have $\Phi_C(\omega, n, \varepsilon, r, M)$ such that

$$\forall_\omega \forall_{f \in C_\omega[0,1]} \forall_{n \in \mathbb{N}} \forall_{p_1,p_2 \in K_{f,n}} \forall_{x_1 \le \dots \le x_n \in [0,1]} \forall_{M \in \mathbb{N}} \forall_{h \in P_n; \|h\|_\infty \le M} \forall_{\varepsilon, r \in \mathbb{Q}^+} \Big($$
$$\forall_{y \in A} |f_0(y)| > \varepsilon \to \sum_{i=1}^{n+1} \sigma_i \int_{A_i} h > \int_B |h| + 1 \to$$
$$\tilde{\exists}_{\lambda \in \mathbb{R}} (\|f_0 - \lambda h\|_1 + \Phi_C(\omega, n, \varepsilon, r, M) < \|f_0\|_1)\Big). \tag{8}$$

We now aim at removing the dependency of Φ_C on r and M, by working with a particular polynomial h. We claim that

$$\forall_{n \in \mathbb{N}} \forall_{x_1 \le \dots \le x_n \in [0,1]} \forall_{\sigma_1 \le \dots \le \sigma_{n+1} \in [-1,1]} \tilde{\exists}_{h \in P_n} \tilde{\exists}_{r \in \mathbb{Q}^+} \Big(\sum_{i=1}^{n+1} \sigma_i \int_{A_i} h > \int_B |h|\Big).$$

To see this, let $y_1, \dots, y_m$ consist of those x_i with $\sigma_i \neq \sigma_{i+1}$, and $h := \pm(x - y_1) \dots (x - y_m)$. Now choose r sufficiently small.

Using Cheney's Lemma 1 to guarantee $n+1$ roots

Using the hidden η in the inequality and $h/\eta \in P_n$ we get

$$\forall_{n\in\mathbb{N}}\forall_{x_1\le\cdots\le x_n\in[0,1]}\forall_{\sigma_1\le\cdots\le\sigma_{n+1}\in[-1,1]}\tilde{\exists}_{h\in P_n}\tilde{\exists}_{r\in\mathbb{Q}^+}$$
$$\Big(\sum_{i=1}^{n+1}\sigma_i\int_{A_i}h > \int_B |h|+1\Big).$$

We can also add the existence of a $k \ge \|h\|_\infty$. Now the metatheorem provides us with $\Phi_{\mathrm{dist}}(n)$ and $\Phi_{\mathrm{bound}}(n)$ such that

$$\forall_{n\in\mathbb{N}}\forall_{x_1\le\cdots\le x_n\in[0,1]}\forall_{\sigma_1\le\cdots\le\sigma_{n+1}\in[-1,1]}\tilde{\exists}_{h\in P_n}\tilde{\exists}_{r\ge\Phi_{\mathrm{dist}}(n)}\Big($$
$$\sum_{i=1}^{n+1}\sigma_i\int_{A_i}h > \int_B |h|+1 \wedge \Phi_{\mathrm{bound}}(n)\ge\|h\|_\infty\Big). \tag{9}$$

Here we can leave out $\tilde{\exists}_{r\ge\Phi_{\mathrm{dist}}(n)}$ and replace r in the kernel by $r := \Phi_{\mathrm{dist}}(n)$ (recall that A_i and B depend on r), because h is such that $\sum_i \sigma_i \int_{A_i} h = \int_A |h|$, and hence is monotone in r.

Using Cheney's Lemma 1 to guarantee $n+1$ roots

Now let $f \in C[0,1]$, $n \in \mathbb{N}$, $p_1, p_2 \in K_{k,n}$ and $x_1 \leq \cdots \leq x_n \in [0,1]$ be fixed. Let $\hat{h}$ be the h provided by (9), where $\sigma_i := f_0(\frac{x_{i-1}+x_i}{2})$ with $x_0 := 0$ and $x_{n+1} := 1$. Therefore from (8) we get (by contraposition)

$$\forall_\omega \forall_{f \in C_\omega[0,1]} \forall_{n \in \mathbb{N}} \forall_{p_1,p_2 \in K_{f,n}} \forall_{x_1 \leq \cdots \leq x_n \in [0,1]} \forall_{\varepsilon \in \mathbb{Q}^+}$$
$$\big(\forall_{\lambda \in \mathbb{R}}(\|f_0 - \lambda h\|_1 + \Phi'_C(\omega, n, \varepsilon) \geq \|f_0\|_1) \to \tilde{\exists}_{y \in A}|f_0(y)| \leq \varepsilon\big)$$

with $\Phi'_C(\omega, n, \varepsilon) := \Phi_C(\omega, n, \varepsilon, \Phi_{\mathrm{dist}}(n), \Phi_{\mathrm{bound}}(n))$. It follows that

$$\forall_\omega \forall_{f \in C_\omega[0,1]} \forall_{n \in \mathbb{N}} \forall_{p_1,p_2 \in K_{f,n}} \forall_{\varepsilon \in \mathbb{Q}^+} \big($$
$$\forall_{h \in P_n}(\|f_0 - h\|_1 + \Phi'_C(\omega, n, \varepsilon) \geq \|f_0\|_1) \to$$
$$\forall_{x_1 \leq \cdots \leq x_n \in [0,1]} \tilde{\exists}_{y \in A}|f_0(y)| \leq \varepsilon\big).$$

Using Cheney's Lemma 1 to guarantee $n+1$ roots

This (classical) existence proof of $y \in A$ can be repeated until we have $n+1$ roots which are $\Phi_{\mathrm{dist}}(n)$ apart:

$$\begin{aligned}
&\forall_{\omega}\forall_{f\in C_{\omega}[0,1]}\forall_{n\in\mathbb{N}}\forall_{p_1,p_2\in K_{f,n}}\forall_{\varepsilon\in\mathbb{Q}^+}\big(\\
&\quad \forall_{h\in P_n}(\|f_0-h\|_1+\Phi'_C(\omega,n,\varepsilon)\geq\|f_0\|_1)\to\\
&\quad \tilde{\exists}_{x_0,\dots,x_n\in[0,1]}(\forall_{i<n}|f_0(x_i)|\leq\varepsilon\wedge\forall_{i<n}x_i+\Phi_{\mathrm{dist}}(n)\leq x_{i+1})\big).
\end{aligned}$$

Putting the parts together

Fix ω, $f \in C_\omega[0,1]$, $n \in \mathbb{N}$, $p_1, p_2 \in K_{f,n}$ and $\varepsilon \in \mathbb{Q}^+$. Assume (for $i = 1, 2$)

$$\|f - p_i\|_1 - d_1(f, P_n) < \Psi(\omega, n, \varepsilon) :=$$
$$\min\{\Phi_\infty(\Phi_g(n, \omega, \varepsilon), \Phi_{\text{many}}(n, \Phi_{\text{dist}}(n), \varepsilon)),$$
$$\Phi'_C(\omega, n, \Phi_{\text{many}}(n, \Phi_{\text{dist}}(n), \varepsilon)), \frac{\varepsilon}{8}\}.$$

Using the previous arguments, this implies $\|p_1 - p_2\|_\infty \leq \varepsilon$, as required.

Conclusion (Kohlenbach and Oliva, APAL 2003)

By a logical analysis of a Cheney's proof of uniqueness of best L_1-approximations

$$\mathrm{E\text{-}PA}^{\omega} + \mathrm{QF\text{-}\tilde{AC}}^{0,1} + \mathrm{W\tilde{K}L} \vdash \forall_{\omega}\forall_{f\in C_{\omega}[0,1]}\forall_{n\in\mathbb{N}}\forall_{p_1,p_2\in P_n}\forall_{\varepsilon}\tilde{\exists}_{\delta}$$
$$\Big(\bigwedge_{i=1,2} (\|f-p_i\|_1 - d_1(f,P_n) \leq \delta) \to \|p_1-p_2\|_1 \leq \varepsilon \Big)$$

we have seen that there exists a modulus of uniqueness depending on ω, n and ε only:

$$\mathrm{HA}^{\omega} \vdash \forall_{\omega}\forall_{f\in C_{\omega}[0,1]}\forall_{n\in\mathbb{N}}\forall_{p_1,p_2\in P_n}\forall_{\varepsilon}$$
$$\Big(\bigwedge_{i=1,2} (\|f-p_i\|_1 - d_1(f,P_n) < \Psi(\omega,n,\varepsilon)) \to \|p_1-p_2\|_1 \leq \varepsilon \Big).$$

Future work

- Get more experience in unwinding classical proofs. Compare Gödel's Dialectica interpretation and its variants (Kohlenbach, Ferreira/Oliva) with refinements of the Dragalin-Friedman A-translation.
- Existence proofs for ODE's by the Cauchy-Euler method. Compare estimates based on the "fundamental inequality", with the ones obtained by Moore's first-order and Kth-order interval method.
- Type theory with approximations: a constructive theory of formal neighborhoods approximating continuous functionals, based on Scott's "information systems", representing domains. Computable functionals appear as r.e. limits or "ideals".