

Herbrand's theorem and extractive proof theory

Herbrand Centenary, ENS

Ulrich Kohlenbach
Department of Mathematics
Technische Universität Darmstadt

February 15, 2008

Extractive Proof Theory: New results by logical analysis of proofs

Input: Ineffective proof P of C

Goal: Additional information on C :

- effective bounds,
- algorithms,
- continuous dependency or full independence from certain parameters,
- generalizations of proofs: weakening of premises.

E.g. Let $C \equiv \forall x \in \mathbb{N} \exists y \in \mathbb{N} F(x, y)$

Naive Attempt: try to extract an explicit computable function realizing (or bounding) ' $\exists y$ ': $\forall x \in \mathbb{N} F(x, f(x))$.

Naive attempt fails

Proposition

There exist a sentence $A \equiv \forall x \exists y \forall z A_{qf}(x, y, z)$ in the language of arithmetic (A_{qf} quantifier-free and hence decidable), such

- *A is **logical valid**,*

Naive attempt fails

Proposition

There exist a sentence $A \equiv \forall x \exists y \forall z A_{qf}(x, y, z)$ in the language of arithmetic (A_{qf} quantifier-free and hence decidable), such

- A is *logical valid*,
- there is *no recursive bound* f s.t. $\forall x \exists y \leq f(x) \forall z A_{qf}(x, y, z)$.

Naive attempt fails

Proposition

There exist a sentence $A \equiv \forall x \exists y \forall z A_{qf}(x, y, z)$ in the language of arithmetic (A_{qf} quantifier-free and hence decidable), such

- A is *logical valid*,
- there is *no recursive bound* f s.t. $\forall x \exists y \leq f(x) \forall z A_{qf}(x, y, z)$.

Proof: Take

$$A := \forall x \exists y \forall z (T(x, x, y) \vee \neg T(x, x, z)),$$

where T is the (primitive recursive) Kleene-T-predicate.

Any bound g on ‘ $\exists y$ ’, i.e. no computable g such that

$$\forall x \exists y \leq g(x) \forall z (T(x, x, y) \vee \neg T(x, x, z))$$

since this would solve the halting problem!

However, one can obtain such **witness candidates** and bounds (and even realizing function(al)s) for a **weakened version** A^H of A :

Definition

$A \equiv \exists x_1 \forall y_1 \exists x_2 \forall y_2 A_{qf}(x_1, y_1, x_2, y_2)$. Then the **Herbrand normal form** of A is defined as

$$A^H := \exists x_1, x_2 A_{qf}(x_1, f(x_1), x_2, g(x_1, x_2)),$$

where f, g are new function symbols, called index functions.

A and A^H are equivalent with respect to logical validity, i.e.

$$\models A \Leftrightarrow \models A^H,$$

but are not logically equivalent since in general

$$\text{PL} \not\vdash A^H \rightarrow A.$$

However the converse implication holds

$$\text{PL} \vdash A \rightarrow A^H.$$

Let PL^2 be the extension of PL by the addition of **function quantifiers**.

Let furthermore AC denote the **schema of choice**

$$\text{AC} : \forall \underline{x} \exists y A(\underline{x}, y) \rightarrow \exists f \forall \underline{x} A(\underline{x}, f(\underline{x})) \quad (\underline{x} = x_1 \dots x_n),$$

then it is an easy exercise to show that

$$\text{PL}^2 + \text{AC} \vdash A \leftrightarrow A^H.$$

We now consider again the sentence

$$A \equiv \forall x \exists y \forall z (P(x, y) \vee \neg P(x, z)),$$

In contrast to A , the **Herbrand normal form** A^H of A

$$A^H \equiv \exists y (P(x, y) \vee \neg P(x, g(y)))$$

allows to construct a list of candidates (uniformly in x, g) for ' $\exists y$ ', namely $(c, g(c))$ (and also $(x, g(x))$) for any constant c :

$$A^{H,D} := (P(x, c) \vee \neg P(x, g(c))) \vee (P(x, g(c)) \vee \neg P(x, g(g(c))))$$

We now consider again the sentence

$$A \equiv \forall x \exists y \forall z (P(x, y) \vee \neg P(x, z)),$$

In contrast to A , the **Herbrand normal form** A^H of A

$$A^H \equiv \exists y (P(x, y) \vee \neg P(x, g(y)))$$

allows to construct a list of candidates (uniformly in x, g) for ' $\exists y$ ', namely $(c, g(c))$ (and also $(x, g(x))$) for any constant c :

$$A^{H,D} := (P(x, c) \vee \neg P(x, g(c))) \vee (P(x, g(c)) \vee \neg P(x, g(g(c))))$$


∈TAUT

We now consider again the sentence

$$A \equiv \forall x \exists y \forall z (P(x, y) \vee \neg P(x, z)),$$

In contrast to A , the **Herbrand normal form** A^H of A

$$A^H \equiv \exists y (P(x, y) \vee \neg P(x, g(y)))$$

allows to construct a list of candidates (uniformly in x, g) for ' $\exists y$ ', namely $(c, g(c))$ (and also $(x, g(x))$) for any constant c :

$$A^{H,D} := (P(x, c) \vee \neg P(x, g(c))) \vee (P(x, g(c)) \vee \neg P(x, g(g(c))))$$


∈TAUT

is a tautology.

A tautology remains a tautology if we replace all occurrences of a term s by a variable y : Replace $g(g(c))$ by z and $g(c)$ by y :

$$A^D \equiv (P(x, c) \vee \neg P(x, y)) \vee (P(x, y) \vee \neg P(x, z)),$$

which still is a tautology.

A tautology remains a tautology if we replace all occurrences of a term s by a variable y : Replace $g(g(c))$ by z and $g(c)$ by y :

$$A^D := (P(x, c) \vee \neg P(x, y)) \vee (P(x, y) \vee \neg P(x, z)),$$

which still is a tautology.

From A^D we can derive A by a so-called **direct proof**:

$$P(x, c) \vee \neg P(x, y) \vee P(x, y) \vee \neg P(x, z)$$

$$\begin{aligned} &P(x, c) \vee \neg P(x, y) \vee P(x, y) \vee \neg P(x, z) \\ &\quad \Downarrow (\forall\text{-introduction}) \\ &P(x, c) \vee \neg P(x, y) \vee \forall z (P(x, y) \vee \neg P(x, z)) \end{aligned}$$

$$P(x, c) \vee \neg P(x, y) \vee P(x, y) \vee \neg P(x, z)$$

↓ (\forall -introduction)

$$P(x, c) \vee \neg P(x, y) \vee \forall z (P(x, y) \vee \neg P(x, z))$$

↓ (\exists -introduction)

$$P(x, c) \vee \neg P(x, y) \vee \exists y \forall z (P(x, y) \vee \neg P(x, z))$$

$$\begin{aligned}
& P(x, c) \vee \neg P(x, y) \vee P(x, y) \vee \neg P(x, z) \\
& \quad \Downarrow (\forall\text{-introduction}) \\
& P(x, c) \vee \neg P(x, y) \vee \forall z (P(x, y) \vee \neg P(x, z)) \\
& \quad \Downarrow (\exists\text{-introduction}) \\
& P(x, c) \vee \neg P(x, y) \vee \exists y \forall z (P(x, y) \vee \neg P(x, z)) \\
& \quad \Downarrow (\forall\text{-introduction}) \\
& \forall y (P(x, c) \vee \neg P(x, y)) \vee \exists y \forall z (P(x, y) \vee \neg P(x, z))
\end{aligned}$$

$$\begin{aligned}
& P(x, c) \vee \neg P(x, y) \vee P(x, y) \vee \neg P(x, z) \\
& \quad \Downarrow (\forall\text{-introduction}) \\
& P(x, c) \vee \neg P(x, y) \vee \forall z (P(x, y) \vee \neg P(x, z)) \\
& \quad \Downarrow (\exists\text{-introduction}) \\
& P(x, c) \vee \neg P(x, y) \vee \exists y \forall z (P(x, y) \vee \neg P(x, z)) \\
& \quad \Downarrow (\forall\text{-introduction}) \\
& \forall y (P(x, c) \vee \neg P(x, y)) \vee \exists y \forall z (P(x, y) \vee \neg P(x, z)) \\
& \quad \Downarrow (\exists\text{-introduction}) \\
& \exists u \forall y (P(x, u) \vee \neg P(x, y)) \vee \exists y \forall z (P(x, y) \vee \neg P(x, z))
\end{aligned}$$

$$\begin{aligned}
& P(x, c) \vee \neg P(x, y) \vee P(x, y) \vee \neg P(x, z) \\
& \quad \Downarrow (\forall\text{-introduction}) \\
& P(x, c) \vee \neg P(x, y) \vee \forall z (P(x, y) \vee \neg P(x, z)) \\
& \quad \Downarrow (\exists\text{-introduction}) \\
& P(x, c) \vee \neg P(x, y) \vee \exists y \forall z (P(x, y) \vee \neg P(x, z)) \\
& \quad \Downarrow (\forall\text{-introduction}) \\
& \forall y (P(x, c) \vee \neg P(x, y)) \vee \exists y \forall z (P(x, y) \vee \neg P(x, z)) \\
& \quad \Downarrow (\exists\text{-introduction}) \\
& \exists u \forall y (P(x, u) \vee \neg P(x, y)) \vee \exists y \forall z (P(x, y) \vee \neg P(x, z)) \\
& \quad \Downarrow (\text{contraction}) \\
& \exists y \forall z (P(x, y) \vee \neg P(x, z))
\end{aligned}$$

$$\begin{aligned}
& P(x, c) \vee \neg P(x, y) \vee P(x, y) \vee \neg P(x, z) \\
& \quad \Downarrow (\forall\text{-introduction}) \\
& P(x, c) \vee \neg P(x, y) \vee \forall z (P(x, y) \vee \neg P(x, z)) \\
& \quad \Downarrow (\exists\text{-introduction}) \\
& P(x, c) \vee \neg P(x, y) \vee \exists y \forall z (P(x, y) \vee \neg P(x, z)) \\
& \quad \Downarrow (\forall\text{-introduction}) \\
& \forall y (P(x, c) \vee \neg P(x, y)) \vee \exists y \forall z (P(x, y) \vee \neg P(x, z)) \\
& \quad \Downarrow (\exists\text{-introduction}) \\
& \exists u \forall y (P(x, u) \vee \neg P(x, y)) \vee \exists y \forall z (P(x, y) \vee \neg P(x, z)) \\
& \quad \Downarrow (\text{contraction}) \\
& \quad \exists y \forall z (P(x, y) \vee \neg P(x, z)) \\
& \quad \quad \Downarrow (\forall\text{-introduction}) \\
& \forall x \exists y \forall z (P(x, y) \vee \neg P(x, z))
\end{aligned}$$

J. Herbrand's Theorem ('Théorème fondamental', 1930)

Theorem

Let $A \equiv \exists x_1 \forall y_1 \exists x_2 \forall y_2 A_{qf}(x_1, y_1, x_2, y_2)$. Then:

$\text{PL} \vdash A$ iff there are terms $s_1, \dots, s_k, t_1, \dots, t_n$ (built up out of the constants and variables of A and the **index functions** used for the formation of A^H) such that

$$A^{H,D} := \bigvee_{i=1}^k \bigvee_{j=1}^n A_{qf}(s_i, f(s_i), t_j, g(s_i, t_j))$$

is a tautology. $A^{H,D}$ is called **Herbrand Disjunction**.

Note that the length of this disjunction is fixed: $k \cdot n$.

The terms s_i, t_j can be extracted from a given PL-proof of A .

Herbrand's Theorem continued

Replacing in $A^{H,D}$ all terms ' $g(s_i, t_j)$ ', ' $f(s_i)$ ', by new variables (treating larger terms first) results in another tautological disjunction A^D s.t. A can be inferred from A by a **direct proof**.

Corollary

A very restricted **cut-free** set of rules already is complete for PL.

An example

(Ulrich Berger) Consider the open first-order theory \mathcal{T} in the language of first-order logic with equality and a constant 0 and two unary function symbols S, f . The only non-logical axiom of \mathcal{T} is $\forall x(S(x) \neq 0)$.

Proposition

$\mathcal{T} \vdash \exists x(f(S(f(x))) \neq x)$.

Proof: Suppose that

$$\forall x(f(S(f(x))) = x),$$

then f is injective, but also (since $S(x) \neq 0$) surjective on $\{x : x \neq 0\}$ and hence non-injective. Contradiction! □

Analyzing the above proof yields the following Herbrand terms:

$$\text{PL} \vdash (S(s) \neq 0) \rightarrow \bigvee_{j=1}^3 (f(S(f(t_j))) \neq t_j),$$

where

$$t_1 := 0, t_2 := f(0), t_3 := S(f(f(0))), s := f(f(0)).$$

□

Remark

- For sentences $A \equiv \forall x \exists y \forall z A_{qf}(x, y, z)$, A^D can always be written in the form

$$A_{qf}(x, t_1, b_1) \vee A_{qf}(x, t_2, b_2) \vee \dots \vee A_{qf}(x, t_k, b_k),$$

where the b_i are new variables and t_i does not contain any b_j with $i \leq j$ (used by Luckhardt's analysis of Roth's theorem, see below).

Remark

- For sentences $A \equiv \forall x \exists y \forall z A_{qf}(x, y, z)$, A^D can always be written in the form

$$A_{qf}(x, t_1, b_1) \vee A_{qf}(x, t_2, b_2) \vee \dots \vee A_{qf}(x, t_k, b_k),$$

where the b_i are new variables and t_i does not contain any b_j with $i \leq j$ (used by Luckhardt's analysis of Roth's theorem, see below).

- Herbrand's theorem immediately extends to so-called open theories, i.e. first-order theories \mathcal{T} whose non-logical axioms G_1, \dots, G_n are all purely universal.

Theorem (Roth 1955)

An algebraic irrational number α has only finitely many exceptionally good rational approximations, i.e. for $\varepsilon > 0$ there are only finitely many $q \in \mathbb{N}$ such that

$$R(q) :\equiv q > 1 \wedge \exists! p \in \mathbb{Z} : (p, q) = 1 \wedge |\alpha - pq^{-1}| < q^{-2-\varepsilon}.$$

Theorem (Roth 1955)

An algebraic irrational number α has only finitely many exceptionally good rational approximations, i.e. for $\varepsilon > 0$ there are only finitely many $q \in \mathbb{N}$ such that

$$R(q) := q > 1 \wedge \exists! p \in \mathbb{Z} : (p, q) = 1 \wedge |\alpha - pq^{-1}| < q^{-2-\varepsilon}.$$

Theorem (Luckhardt 1985/89)

The following upper bound on $\#\{q : R(q)\}$ holds:

$$\#\{q : R(q)\} < \frac{7}{3}\varepsilon^{-1} \log N_\alpha + 6 \cdot 10^3 \varepsilon^{-5} \log^2 d \cdot \log(50\varepsilon^{-2} \log d),$$

where $N_\alpha < \max(21 \log 2h(\alpha), 2 \log(1 + |\alpha|))$ and h is the logarithmic absolute homogeneous height and $d = \deg(\alpha)$.

Independently: Bombieri and van der Poorten 1988.



Towards generalizations of Herbrand's theorem

Allow **functionals** $\Phi(x, f)$ instead of just Herbrand terms: Let's consider again the example

$$A \equiv \forall x \exists y \forall z (T(x, x, y) \vee \neg T(x, x, z)).$$

Towards generalizations of Herbrand's theorem

Allow **functionals** $\Phi(x, f)$ instead of just Herbrand terms: Let's consider again the example

$$A \equiv \forall x \exists y \forall z (T(x, x, y) \vee \neg T(x, x, z)).$$

A^H can be realized by a computable functional of type level 2 which is defined by cases:

$$\Phi(x, g) := \begin{cases} x & \text{if } \neg T(x, x, g(x)) \\ g(x) & \text{otherwise.} \end{cases}$$

Towards generalizations of Herbrand's theorem

Allow **functionals** $\Phi(x, f)$ instead of just Herbrand terms: Let's consider again the example

$$A \equiv \forall x \exists y \forall z (T(x, x, y) \vee \neg T(x, x, z)).$$

A^H can be realized by a computable functional of type level 2 which is defined by cases:

$$\Phi(x, g) := \begin{cases} x & \text{if } \neg T(x, x, g(x)) \\ g(x) & \text{otherwise.} \end{cases}$$

From this definition it easily follows that

$$\forall x, g (T(x, x, \Phi(x, g)) \vee \neg T(x, x, g(\Phi(x, g)))).$$

If A is not provable in PL but e.g. in PA more **complicated functionals** are needed (Kreisel 1951):

Let (a_n) be a nonincreasing sequence in $[0, 1]$. Then, clearly, (a_n) is convergent and so a Cauchy sequence which we write as:

$$(1) \forall k \in \mathbb{N} \exists n \in \mathbb{N} \forall m \in \mathbb{N} \forall i, j \in [n; n + m] (|a_i - a_j| \leq 2^{-k}),$$

where $[n; n + m] := \{n, n + 1, \dots, n + m\}$.

If A is not provable in PL but e.g. in PA more **complicated functionals** are needed (Kreisel 1951):

Let (a_n) be a nonincreasing sequence in $[0, 1]$. Then, clearly, (a_n) is convergent and so a Cauchy sequence which we write as:

$$(1) \forall k \in \mathbb{N} \exists n \in \mathbb{N} \forall m \in \mathbb{N} \forall i, j \in [n; n+m] (|a_i - a_j| \leq 2^{-k}),$$

where $[n; n+m] := \{n, n+1, \dots, n+m\}$.

Then the (partial) Herbrand normal form of this statement is

$$(2) \forall k \in \mathbb{N} \forall g \in \mathbb{N}^{\mathbb{N}} \exists n \in \mathbb{N} \forall i, j \in [n; n+g(n)] (|a_i - a_j| \leq 2^{-k}).$$

By E. Specker ('Specker sequences') there exist computable such sequences (a_n) even in $\mathbb{Q} \cap [0, 1]$ without computable bound on ' $\exists n$ ' in (1). By contrast, there is a simple (primitive recursive) bound $\Phi^*(g, k)$ on (2) (also referred to as 'metastability' by Tao):

Proposition

Let (a_n) be any nonincreasing sequence in $[0, 1]$ then

$$\forall k \in \mathbb{N} \forall g \in \mathbb{N}^{\mathbb{N}} \exists n \leq \Phi^*(g, k) \forall i, j \in [n; n + g(n)] (|a_i - a_j| \leq 2^{-k}),$$

where

$$\Phi^*(g, k) := \tilde{g}^{(2^k)}(0) \text{ with } \tilde{g}(n) := n + g(n).$$

Moreover, there exists an $i < 2^k$ such that n can be taken as $\tilde{g}^{(i)}(0)$.

Remark

The previous result can be viewed as a polished form of a **Herbrand disjunction** of **variable (in k) length**:

$$\bigvee_{i=0}^{2^k-1} (|a_{\tilde{g}^{(i)}(0)} - a_{\tilde{g}(\tilde{g}^{(i)}(0))}| \leq 2^{-k}).$$

Remark

The previous result can be viewed as a polished form of a **Herbrand disjunction** of **variable (in k) length**:

$$\bigvee_{i=0}^{2^k-1} (|a_{\tilde{g}^{(i)}(0)} - a_{\tilde{g}(\tilde{g}^{(i)}(0))}| \leq 2^{-k}).$$

Corollary

(T. Tao's finite convergence principle)

$$\forall k \in \mathbb{N}, g : \mathbb{N} \rightarrow \mathbb{N} \exists M \in \mathbb{N} \forall 0 \leq a_0 \leq \dots \leq a_M \leq 1 \exists N \in \mathbb{N} \\ (N + g(N) \leq M \wedge \forall n, m \in [N, N + g(N)] (|a_n - a_m| \leq 2^{-k})).$$

Kreisel's no-counterexample interpretation

Definition

Let $A \equiv \exists x_1 \forall y_1 \dots \exists x_n \forall y_n A_{qf}(x_1, y_1, \dots, x_n, y_n)$. If a tuple of functionals Φ_1, \dots, Φ_n realizes the Herbrand normal form A^H of A , i.e. if

$$\forall \underline{f} A_{qf}(\Phi_1(\underline{f}), f_1(\Phi_1(\underline{f})), \dots, \Phi_n(\underline{f}), f_n(\Phi_1(\underline{f}), \dots, \Phi_n(\underline{f})))$$

is true (where $\underline{f} = f_1, \dots, f_n$), then we say that $\underline{\Phi} (= \Phi_1, \dots, \Phi_n)$ satisfies the **no-counterexample interpretation** of A .

Motivation for the name 'no-counterexample interpretation':

Let A be as above. Then $\neg A$ is equivalent to

$$\forall x_1 \exists y_1 \dots \forall x_n \exists y_n \neg A_{qf}(x_1, y_1, \dots, x_n, y_n).$$

So a counterexample to A is given by functions f_1, \dots, f_n such that

$$(+) \forall \underline{x} \neg A_{qf}(x_1, f_1(x_1), \dots, x_n, f_n(x_1, \dots, x_n))$$

holds. Hence functionals $\underline{\Phi}$ satisfying the n.c.i. of A produce a **counterexample** to (+) i.e. to the existence of **counterexample functions** f_1, \dots, f_n .

Problems of the no-counterexample interpretation

For principles $F \in \exists\forall\exists$ n.c.i. no longer 'correct'.

Direct example: Infinitary Pigeonhole Principle (IPP):

$$\forall n \in \mathbb{N} \forall f : \mathbb{N} \rightarrow C_n \exists i \leq n \forall k \in \mathbb{N} \exists m \geq k (f(m) = i),$$

where $C_n := \{0, 1, \dots, n\}$.

IPP implies \exists -induction and can cause arbitrary **primitive recursive complexity**, but it has a trivial n.c.i.:

$$\begin{aligned} (\text{IPP})^H &\equiv \\ \forall n \in \mathbb{N} \forall f : \mathbb{N} \rightarrow C_n \forall F : C_n \rightarrow \mathbb{N} \exists i \leq n \exists m \geq F(i) (f(m) = i). \end{aligned}$$

Trivial n.c.i.-solution:

$$M(n, f, F) := \max\{F(i) : i \leq n\} \text{ and } I(n, f, F) := f(M(n, f, F))$$

are realizers for ' $\exists m$ ' and ' $\exists i$ ' in $(\text{IPP})^H$.

M, I **do not reflect** true complexity of IPP!

Gödel Functional Interpretation G (Gödel)

$$\text{QF-AC} : \forall x \exists y A_{qf}(x, y) \rightarrow \exists Y \forall x A_{qf}(x, Y(x)).$$

$$(\text{IPP}) \stackrel{\text{QF-AC}}{\Leftrightarrow}$$

$$\forall f : \mathbb{N} \rightarrow C_n \exists i \leq n \exists g : \mathbb{N}^{\mathbb{N}} \forall k \in \mathbb{N} (g(k) \geq k \wedge f(g(k)) = i)$$

$$\stackrel{\text{QF-AC}}{\Leftrightarrow}$$

$$(\text{IPP})^G \equiv \left\{ \begin{array}{l} \forall f : \mathbb{N} \rightarrow C_n \forall K : C_n \times \mathbb{N}^{\mathbb{N}} \rightarrow \mathbb{N} \exists i \leq n \exists g : \mathbb{N}^{\mathbb{N}} \\ (g(K(i, g)) \geq K(i, g) \wedge f(g(K(i, g)))) = i. \end{array} \right.$$

Solution for (IPP)^G (P. Oliva 2006)

Let $B_{fin} : (\mathbb{N} \times \mathbb{N}^{\mathbb{N}} \rightarrow \mathbb{N}) \times \mathbb{N} \times \mathbb{N} \rightarrow \mathbb{N}$ satisfy

$$B_{fin}(K, n, s) = \begin{cases} \langle \rangle, & \text{if } lth(s) \geq n + 1 \\ \langle c_s \rangle * B_{fin}(K, n, s * \langle c_s \rangle), & \text{otherwise,} \end{cases}$$

where

$$c_s := K(lth(s), h_s) \wedge h_s := \lambda x. M(s * \langle x \rangle * B_{fin}(K, n, s * \langle x \rangle)),$$

with $M(k) := \max\{(k)_0, \dots, (k)_{lth(k)-1}\}$.

Define $\langle x_0, \dots, x_n \rangle := B_{fin}(n, \langle \rangle)$, $g_i := h_{\langle x_0, \dots, x_{i-1} \rangle}$ and, finally,

$$I(n, f, K) := i := f(g_0(K(0, g_0))) \wedge G(n, f, K) := g := g_i.$$

General facts on Gödel's functional interpretation

- Extracts **primitive recursive functionals** of higher order from proofs in Peano Arithmetic PA^ω in all types ρ plus the axiom of choice for quantifier-free formulas (Gödel, Kreisel, Yasugi):

$$\text{QF-AC} : \forall x^\rho \exists y^\tau A_{qf}(x, y) \rightarrow \exists Y^{\rho \rightarrow \tau} \forall x^\rho A_{qf}(x, Y(x)).$$

General facts on Gödel's functional interpretation

- Extracts **primitive recursive functionals** of higher order from proofs in Peano Arithmetic PA^ω in all types ρ plus the axiom of choice for quantifier-free formulas (Gödel, Kreisel, Yasugi):

$$\text{QF-AC} : \forall x^\rho \exists y^\tau A_{qf}(x, y) \rightarrow \exists Y^{\rho \rightarrow \tau} \forall x^\rho A_{qf}(x, Y(x)).$$

- Extracts **bar recursive functionals** (C. Spector) from proofs augmented by full dependent choice (full classical analysis \mathcal{A}^ω).

General facts on Gödel's functional interpretation

- Extracts **primitive recursive functionals** of higher order from proofs in Peano Arithmetic PA^ω in all types ρ plus the axiom of choice for quantifier-free formulas (Gödel, Kreisel, Yasugi):

$$\text{QF-AC} : \forall x^\rho \exists y^\tau A_{qf}(x, y) \rightarrow \exists Y^{\rho \rightarrow \tau} \forall x^\rho A_{qf}(x, Y(x)).$$

- Extracts **bar recursive functionals** (C. Spector) from proofs augmented by full dependent choice (full classical analysis \mathcal{A}^ω).
- Can be combined with a majorization technique to extract highly **uniform bounds** (monotone functional interpretation, Kohlenbach).

General facts on Gödel's functional interpretation

- Extracts **primitive recursive functionals** of higher order from proofs in Peano Arithmetic PA^ω in all types ρ plus the axiom of choice for quantifier-free formulas (Gödel, Kreisel, Yasugi):

$$\text{QF-AC} : \forall x^\rho \exists y^\tau A_{qf}(x, y) \rightarrow \exists Y^{\rho \rightarrow \tau} \forall x^\rho A_{qf}(x, Y(x)).$$

- Extracts **bar recursive functionals** (C. Spector) from proofs augmented by full dependent choice (full classical analysis \mathcal{A}^ω).
- Can be combined with a majorization technique to extract highly **uniform bounds** (monotone functional interpretation, Kohlenbach).
- Applies to extensions of \mathcal{A}^ω by **abstract structures** such as arbitrary metric, hyperbolic, CAT(0), normed, Hilbert spaces (Kohlenbach).

An application in Metric Fixed Point Theory

- (X, d, W) is a **hyperbolic space** (e.g. convex subset of a normed space).
- $f : X \rightarrow X$ is a **nonexpansive mapping**: $d(f(x), f(y)) \leq d(x, y)$.
- (λ_n) is a sequence in $[0, 1]$ that is **bounded away from 1** and **divergent in sum**.
- $x_{n+1} = (1 - \lambda_n)x_n \oplus \lambda_n f(x_n)$ (**Krasnoselski-Mann iter.**).

An application in Metric Fixed Point Theory

- (X, d, W) is a **hyperbolic space** (e.g. convex subset of a normed space).
- $f : X \rightarrow X$ is a **nonexpansive mapping**: $d(f(x), f(y)) \leq d(x, y)$.
- (λ_n) is a sequence in $[0, 1]$ that is **bounded away from 1** and **divergent in sum**.
- $x_{n+1} = (1 - \lambda_n)x_n \oplus \lambda_n f(x_n)$ (**Krasnoselski-Mann iter.**).

Theorem (Ishikawa 1976, Goebel/Kirk 1983)

If (x_n) is bounded, then $d(x_n, f(x_n)) \rightarrow 0$.

Logical analysis of the proof of Ishikawa's theorem

Let $K \in \mathbb{N}$ and $\alpha : \mathbb{N} \rightarrow \mathbb{N}$ be such that

$$(\lambda_n)_{n \in \mathbb{N}} \in [0, 1 - \frac{1}{K}]^{\mathbb{N}} \text{ and } \forall n \in \mathbb{N} (n \leq \sum_{i=0}^{\alpha(n)} \lambda_i).$$

Logical metatheorem applied to proof of Ishikawa's theorem yields computable Ψ, Φ s.t. for all $l \in \mathbb{N}$ and n.e. f

$$\forall i, j \leq \Psi(K, \alpha, b, \tilde{b}, l) (d(x, f(x)) \leq \tilde{b} \wedge d(x_i, x_j) \leq b) \rightarrow \\ \forall m \geq \Phi(K, \alpha, b, \tilde{b}, l) (d(x_m, f(x_m)) < 2^{-l}).$$

holds in **any (nonempty) hyperbolic space** (X, d, W) .

Theorem (K.2007)

$(X, d, W), (\lambda_n), K$ as above, $f : X \rightarrow X$ nonexpansive the following holds for all $\varepsilon, b, \tilde{b} > 0$:

If $d(x, f(x)) \leq \tilde{b}$ and $\forall i \leq \Phi \forall j \leq \alpha(\Phi, M)$ ($d(x_i, x_{i+j}) \leq b$)
then $\forall n \geq \Phi$ ($d(x_n, f(x_n)) \leq \varepsilon$),

where

$$\Phi := \Phi(K, \alpha, b, \tilde{b}, \varepsilon) := \hat{\alpha} \left(\left\lceil \frac{\tilde{b} \cdot \exp\left(K \cdot \left(\frac{3\tilde{b}+b}{\varepsilon} + 1\right)\right)}{\varepsilon} \right\rceil - 1, M \right),$$

$$M := \left\lceil \frac{3\tilde{b}+b}{\varepsilon} \right\rceil,$$

$$\hat{\alpha}(0, n) := \tilde{\alpha}(0, n), \quad \hat{\alpha}(i+1, n) := \tilde{\alpha}(\hat{\alpha}(i, n), n) \text{ with}$$

$$\tilde{\alpha}(i, n) := i + \alpha(i, n) \quad (i, n \in \mathbb{N})$$

with α s.t.

$$\forall i, n \in \mathbb{N} ((\alpha(i, n) \leq \alpha(i+1, n)) \wedge (n \leq \sum_{s=i}^{i+\alpha(i,n)-1} \lambda_s)).$$

Remark

If (λ_n) in $[\frac{1}{K}, 1 - \frac{1}{K}]$, then we may take $\alpha(i, n) := K \cdot n$.

Corollary (K.2007)

Let (λ_n) in $[a, b] \subset (0, 1)$.

If $\lim_{n \rightarrow \infty} \frac{c(n)}{n} \rightarrow 0$, where $c(n) := \max\{d(x, x_j) : j \leq n\}$,

then

$$\lim_{n \rightarrow \infty} d(x_n, f(x_n)) = 0.$$

Result optimal: $c(n) \leq C \cdot n$ for some $C > 0$ not sufficient!

An application in Ergodic Theory

Let X be a **Hilbert space**, $f : X \rightarrow X$ **linear and nonexpansive**.

$$A_n(x) := \frac{1}{n+1} S_n(x), \text{ where } S_n(x) := \sum_{i=0}^n f^i(x).$$

Theorem (von Neuman Mean Ergodic Theorem)

For every $x \in X$, the sequence $(A_n(x))_n$ converges.

Based on the logical metatheorems discussed above:

Theorem (Avigad-Gerhardy-Towsner 2007)

$$\forall g : \mathbb{N} \rightarrow \mathbb{N} \forall \varepsilon > 0 \forall x \in X \exists n \leq \Phi \forall i, j \in [n; n + g(n)] \\ (\|A_i(x) - A_j(x)\| \leq \varepsilon),$$

where $\Phi = h^{(k)}(0)$ with

$$\rho := \lceil \frac{\|x\|}{\varepsilon} \rceil, \quad k := 2^9 \rho^2, \quad h(n) := n + 2^{13} \rho^4 \tilde{g}((n+1)\tilde{g}(2n\rho)\rho^2) \\ \text{and } \tilde{g}(n) := \max\{i + g(i) : i \leq n\}.$$

Result is used in a recent paper of T. Tao!

Announcement

As part of the

Colloquium Logicum 2008
of the
German Association for Logic (DVMLG)
TU Darmstadt, September 10-12, 2008
<http://www.mathematik.tu-darmstadt.de/fbereiche/logik/events/collogicum/>

there will be a

Herbrand Centenary Lecture
by
Georg Kreisel (Salzburg, FRS)

Further Reading

- 1 Kohlenbach, U., Proof Interpretations and the Computational Content of Proofs in Mathematics. Bulletin EATCS **93**, pp. 143-173 (2007).
- 2 Kohlenbach, U., Applied Proof Theory: Proof Interpretations and their Use in Mathematics. Springer Monographs in Mathematics, xx+536pp. (2008).