Program Extraction in Constructive Analysis

Helmut Schwichtenberg

Mathematisches Institut, Universität München

Motivation

Bishop: “mathematics as a numerical language”.

Extract programs from proofs, for exact real numbers.

Special emphasis on low type level witnesses (making use of separability).

Here: approximate solutions of ODEs.

Ordinary differential equations

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

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

on an interval I is a continuous function $\varphi\colon I \longrightarrow \mathbb{R}$ 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 $f \colon D \to \mathbb{R}$ 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 $\varphi, \psi \colon I \to \mathbb{R}$ be two solutions of (1). If $\varphi(a) = \psi(a)$ for some $a \in I$, then $\varphi(x) = \psi(x)$ for all $x \in I$.

The example $y' = y^{1/3}$ 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: cannot expect to be able to prove it constructively.

Picard's existence theorem for ODEs

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

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

Then $(\varphi_n)_{n \in \mathbb{N}}$ 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 $R\colon |x - a_0| \le h, |y - b_0| \le Mh$, 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 $f\colon D \to \mathbb{R}$ 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

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

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

The Cauchy-Euler existence theorem for ODEs

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

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

(φ_n) converges uniformly and absolutely to a soln 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 $\varphi \le \varphi_n$ or $\varphi_n \le \varphi$. Then there is a constant c independent of n such that $|\varphi(x) - \varphi_n(x)| \le 2^{-n}c$ for $|x - a_0| \le h$.

Proof. By the Fundamental Inequality

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

Tools

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

Idea: use separability to avoid high type levels. Where?

- "Order located" instead of "totally bounded".
- Continuity in $\mathbb{R}$, and $\mathbb{R}^2$.
- Uniformly convergent sequences of functions.

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

$$\forall k, n, m.\ \alpha(k) \le n, m \to |a_n - a_m| \le 2^{-k},$$

and α is weakly increasing.

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

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

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 $\lvert x\rvert \in_l \mathbb{R}^+$	$\begin{cases} \frac{1}{a_n} & \text{if } a_n \neq 0 \\ 0 & \text{if } a_n = 0 \end{cases}$	$\alpha(2(l+1)+k)$

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

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

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 $y - x \in \mathbb{R}^{+}$.

$$x \leq y \leftrightarrow \forall k \exists p \forall n. p \leq n \rightarrow a_n \leq b_n + 2^{-k}$$

$$x < y \leftrightarrow \exists k, q \forall n.\ q \leq n \rightarrow a_n + 2^{-k} \leq b_n$$

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

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

Approximate Splitting Principle. Let x, y, z be given and assume $x < y$. Then we can find k, q such that either $z <_{k,q} y$ or $x <_{k,q} z$.

Proof. Let $x := ((a_n)_n, \alpha)$, $y := ((b_n)_n, \beta)$, $z := ((c_n)_n, \gamma)$. From $x < y$ obtain p, k such that with $\varepsilon := 2^{-k}$

$$\forall n.\, p \leq n \to a_n + 3\varepsilon \leq b_n - 3\varepsilon.$$

Let $q := \max\bigl(\alpha(k), \beta(k), \gamma(k), p\bigr)$. Cases: $c_q \leq b_q - 3\varepsilon$ or $b_q - 3\varepsilon < c_q$. $\square$

$z < y$ or $x < z$ depends on the representation of x, y, z.

Suprema

Let S be a set of reals. A real y is an upper bound of S if $x \leq y$ for all $x \in S$. A real y is a supremum of S if y is an upper bound of S, and for every rational $a < y$ there is a real $x \in S$ such that $a \leq x$.

A set S of reals is order located above if for every $a < b$, either $x \leq b$ for all $x \in S$ or else $a \leq x$ for some $x \in S$.

Least-Upper-Bound Principle. Let S be an inhabited set of reals that is bounded above. Then S has a supremum iff it is order located above.

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

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

- $\omega_f\colon \mathbb{N} \to \mathbb{N}$ (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 $f(x)$ of a continuous f (given by h_f, α_f, ω_f) to a real $x := ((a_n)_n, \alpha)$ is defined to be

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

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

Can show:

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

Composition of continuous functions

Let $f\colon I \to \mathbb{R}$ and $g\colon J \to \mathbb{R}$ be continuous. Assume that $h_f[(I \cap \mathbb{Q}) \times \mathbb{N}] \subseteq J$. Then $g \circ f\colon I \to \mathbb{R}$ is defined by

$$h_{g\circ f}(a,n) := h_g(h_f(a,n),n)$$

$$\alpha_{g\circ f}(k) := \max\big(\alpha_g(k+2), \alpha_f(\omega_g(k+1)-1)\big)$$

$$\omega_{g\circ f}(k) := \omega_f(\omega_g(k)-1)+1$$

Bound for the range of f

Let $f\colon [a,b] \to \mathbb{R}$ be continuous, given by h_f, α_f and ω_f. Then for all $n \geq n_0 := \alpha_f(0)$ and rationals $c \in I$

$$|h_f(c,n)| \leq M := |h_f(a,n_0)| + N + 1,$$

where $(c-a)2^{\omega_f(0)-1} \leq N \in \mathbb{N}$.

Hence: range of f is bounded above by M.

Supremum $\|f\|_I$ of $f\colon I \to \mathbb{R}$

...can be shown to exist constructively. Bishop's proof uses "totally bounded sets", a type level 2 concept:

A k-net for a set S of reals is given by a finite list y_i $(i < n_k)$ of reals in S, and a map $\mathsf{sel}_k\colon S \to \{0, \dots, n_k - 1\}$ (of type level 2): $|y_i - x| \le 2^{-k}$, with $i := \mathsf{sel}_k(x)$.

S is totally bounded if for every k we have a k-net for S.

We prove instead that the range is order located above, which entails that is has a supremum:

Lemma. Let $f\colon I \to \mathbb{R}$ be continuous. Then the range of f is order located above. $(\Rightarrow \|f\|_I$ exists).

Proof. Given $a < b$, fix k such that $2^{-k} \leq \frac{1}{3}(b-a)$. Take a partition $a_0, \dots, a_l$ of I of mesh $\leq 2^{-\omega_f(k)+2}$. Then for every $c \in I$ there is an i such that $|c - a_i| \leq 2^{-\omega_f(k)+1}$. Let $n_k := \alpha_f(k)$ and consider all finitely many

$$h(a_i, n_k) \quad \text{for } i = 0, \dots, l.$$

Let $h(a_j, n_k)$ be the maximum of all those.
If $h(a_j, n_k) \leq a + \frac{1}{3}(b-a)$, then $f(x) \leq b$ for all x.
If $a + \frac{1}{3}(b-a) < h(a_j, n_k)$, then $a \leq f(a_j)$. □

Approximate intermediate value theorem

For every continuous $f\colon [a,b] \to \mathbb{R}$ with $f(a) \leq 0 \leq f(b)$, and every k, we can find $c \in [a,b]$ such that $|f(c)| \leq 2^{-k}$.

Problem: need to partition $[a,b]$ into as many pieces as the modulus of the continuous function requires.

Reason: f may be flat.

Cure: use more knowledge on f.

$f\colon [a,b] \to \mathbb{R}$ is locally nonconstant whenever if $a \le a' < b' \le b$ and c is arbitrary, then $f(x) \ne c$ for some $x \in [a',b']$.

Intermediate Value Theorem. If $f\colon [a,b] \to \mathbb{R}$ is continuous with $f(a) < 0 < f(b)$, and locally nonconstant, then we can find $x \in [a,b]$ with $f(x) = 0$.

Proof. Construct $(c_n)_n$ and $(d_n)_n$ such that for all n

$$a = c_0 \le c_1 \le \cdots \le c_n < d_n \le \cdots \le d_1 \le d_0 = b,$$
$$f(c_n) < 0 < f(d_n),$$
$$d_n - c_n \le \left(\frac{2}{3}\right)^n (b-a).$$

Example: $f\colon [1,2] \to \mathbb{R}$ mapping $x \mapsto x^2 - 2$, given by

- the approximating map $h_f(a,n) := a^2 - 2$,
- the uniform Cauchy modulus $\alpha_f(k) := 0$, and
- the modulus $k \mapsto k + p - 1$ of uniform continuity, where $p := 2$ is such that $|a+b| \le 2^p$ for $a, b \in [1,2]$, because

$$|a - b| \le 2^{-k-p} \to |a^2 - b^2| = |(a-b)(a+b)| \le 2^{-k}.$$

Clearly $f(1) < 0 < f(2)$, and f is strictly monotic. Hence: proof of $\exists x{\in}[1,2](f(x) = 0)$ contains algorithm for $\sqrt{2}$. (Implemented in Coq with Pierre Letouzey; very fast).

Differentiation

Let $f, g\colon I \to \mathbb{R}$ be continuous. g is called derivative of f with modulus $\delta_f\colon \mathbb{N} \to \mathbb{N}$ if for $x, y \in I$ with $x < y$,

$$y \le x + 2^{-\delta_f(k)} \to \left| f(y) - f(x) - g(x)(y-x) \right| \le 2^{-k}(y-x).$$

A bound on f' serves as a Lipschitz constant for f:

Lemma. Let $f\colon I \to \mathbb{R}$ be continuous with derivative f'. Let f' be bounded by M. Then for $x, y \in I$ with $x < y$,

$$\left| f(y) - f(x) \right| \le M(y-x).$$

Lemma (Rolle). Let $f\colon [a,b] \to \mathbb{R}$ be continuous with derivative f', and assume $f(a) = f(b)$. Then for every $k \in \mathbb{N}$ we can find $c \in [a,b]$ such that $|f'(c)| \leq 2^{-k}$.

Mean Value Theorem. Let $f\colon [a,b] \to \mathbb{R}$ be continuous with derivative f'. Then for every $k \in \mathbb{N}$ we can find $c \in [a,b]$ such that

$$\left| f(b) - f(a) - f'(c)(b-a) \right| \leq 2^{-k}(b-a).$$

Integration

Assume that $f\colon [a,b] \to \mathbb{R}$ is continuous with modulus ω_f.

$$S(f,a,b,n) := \frac{b-a}{n} \sum_{i=0}^{n-1} h_f(a_i,n) \quad \text{with } a_i := a + \tfrac{i}{n}(b-a)$$

Then $(S(f,a,b,n))_{n\in\mathbb{N}}$ is a Cauchy sequence of rationals with modulus $\alpha(p) = 2^{\omega_f(p+q+1)}(b-a)$, where q is such that $b-a \leq 2^q$; we denote this real by

$$\int_a^b f(x)\,dx.$$

Fundamental theorem of calculus

Given a continuous $f\colon [a,b] \to \mathbb{R}$ and $c \in [a,b]$, we can establish

$$F(x) := \int_c^x f(t)\,dt$$

as a continuous function, via

$$\begin{aligned} h_F(a,n) &:= S(f,c,a,n), \\ \alpha_F(k) &:= \max\bigl(\alpha_f(0), 2^{\omega_f(k+1)}\bigr), \\ \omega_F(k) &:= \max\bigl(p+k, \omega_f(k+1)\bigr), \end{aligned}$$

where p is such that $h_f(b_i,n) \le 2^p$, for $n \ge \alpha_f(0)$.

Fundamental theorem of calculus (ctd.)

Theorem. Let $f\colon [a,b] \to \mathbb{R}$ be continuous, $c \in [a,b]$ and

$$F(x) := \int_c^x f(t)\,dt.$$

Then F has f as derivative, with modulus ω_f. If G is any differentiable function on $[a,b]$ with $G' = f$, then the difference $F - G$ is a constant function.

Corollary. Let $f\colon I \to \mathbb{R}$ be continuous and $F\colon I \to \mathbb{R}$ such that $F' = f$. Then for all $a, b \in I$

$$\int_a^b f(x)\,dx = F(b) - F(a).$$

Related work on exact real numbers

- Redundant b-adic notation (Wiedmer '80, Boehm & Cartwright '90, Ciaffaglione & Di Gianantonio '99)
- Continued fractions (Gosper '90, Vuillemin '90)
- Möbius transformation as a unifying approach to real computation (Edalat & Potts '97)
- PCF + real number data type (Di Gianantonio '93, '96, Escardó '96)
- ODEs via domain theory (Edalat & Pattinson '03)

Related work on program extraction

1. Luis Cruz-Filipe: Thesis in Nijmegen 2004 (Geuvers), on C-CoRN.

2. Stefan Berghofer: "Proofs, Programs and Executable Specifications in Higher Order Logic", 2003 (Nipkow).

3. Monika Seisenberger: "On the Constructive Content of Proofs", 2003.

C-CoRN: Constructive Coq Repository at Nijmegen

Lecture by Herman Geuvers on friday. Grew out of the FTA project. Comments:

- Strong extensionality required: $\forall x, y. f(x) \# f(y) \to x \# y$. Missing witness harmful for program extraction.

- The **Set**, **Prop** distinction in Coq was found to be insufficient. Introduced **CProp** in addition.

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

Conclusion

- Constructive analysis with witnesses of low type level. Type level 1 representation of continuous functions.

- The Cauchy-Euler construction of approximate solutions to ODEs as a type level 1 process.

Future work

1. Case studies for program extraction. (Kneser's proof of the fundamental theorem of algebra, cf. Geuvers et al. in Nijmegen and Letouzey in Paris).

2. Resource sensitivity. Gödel's T can be restricted (using ramification and linearity) such that the definable functions are the poly-time ones [BNS '00, Hofmann]. Work with corresponding arithmetical system.