Verified algorithms via proofs

Nils Köpp and Helmut Schwichtenberg

Mathematisches Institut, LMU, München

Gothenburg, 26. August 2022

- Proofs: on constructive real arithmetic (Bishop).
- Algorithms: operate on reals represented by streams of digits.
- Algorithms via proofs: extract computational content. Needs $P(x)$ and $P^{\mathbf{r}}(x,t)$ "t realizes $P(x)$" for predicates P.
- Verification: formal proof that the extracted term realizes A.

A theory of computable functionals (TCF)

Similar to HA^ω, but

- add inductively and coinductively defined predicates,
- distinguish computationally relevant (c.r.) and non-computational (n.c.) predicates,
- add realizability predicates (internal "meta"-step),
- allow partial functionals, defined by equations (possibly non-terminating, like corecursion),
- minimal logic: only $\to$, $\forall$ primitive. $\vee$, $\exists$ inductively defined

Minlog implements TCF.

- Assume M is an **r**-free proof of A, i.e., M contains no realizability predicates $P^{\mathbf{r}}$ or ${}^{\mathrm{co}}P^{\mathbf{r}}$.
- We can define its extracted term $\mathrm{et}(M)$, of type $\tau(A)$, with the aim to express M's computational content.
- $\mathrm{et}(M)$ is a term built from variables, constructors, destructors and (co)recursion operators by λ-abstraction and application.

Theorem (Soundness)

Let M be an **r***-free derivation of a formula A from assumptions $u_i : C_i$ $(i < n)$. Then we can derive*

$$\begin{cases} \mathrm{et}(M)\ \mathbf{r}\ A & \text{if } A \text{ is c.r.} \\ A & \text{if } A \text{ is n.c.} \end{cases}$$

from assumptions

$$\begin{cases} z_{u_i}\ \mathbf{r}\ C_i & \text{if } C_i \text{ is c.r.} \\ C_i & \text{if } C_i \text{ is n.c.} \end{cases}$$

We express

- Kolmogorov's view of "formulas as problems"[1]
- Feferman's dictum "to assert is to realize"[2]

by invariance axioms:

For **r**-free c.r. formulas A we require as axioms

$$\mathrm{InvAll}_A\colon \forall_z(z\ \mathbf{r}\ A \to A).$$
$$\mathrm{InvEx}_A\colon A \to \exists_z(z\ \mathbf{r}\ A).$$

[1] Zur Deutung der intuitionistischen Logik, Math. Zeitschr., 1932

[2] Constructive theories of functions and classes, Logic Colloquium 78, p.208

Invariance axioms used in the proof of soundness (1):

Case $(\lambda_{u^A} M^B)^{A\to B}$ with B n.c. We need a derivation of $A \to B$.
Subcase A c.r. By IH we have a derivation of B from $z\ \mathbf{r}\ A$. Using the invariance axiom $A \to \exists_z(z\ \mathbf{r}\ A)$ we get the required derivation of B from A:

$$\dfrac{\dfrac{A \to \exists_z(z\ \mathbf{r}\ A) \qquad A}{\exists_z(z\ \mathbf{r}\ A)} \qquad \begin{array}{c}[z\ \mathbf{r}\ A] \\ \mid \mathrm{IH} \\ B\end{array}}{B}\ \exists^-$$

Invariance axioms used in the proof of soundness (2):

Case $(M^{A\to B}N^A)^B$ with B n.c. Goal: find a derivation of B.
Subcase A c.r. By IH we have derivations of $A \to B$ and of $\mathrm{et}(N)\ \mathbf{r}\ A$. From the invariance axiom $\forall_z(z\ \mathbf{r}\ A \to A)$ we obtain the required derivation of B by $\to^-$ from the derivation of $A \to B$ and

$$\dfrac{\dfrac{\forall_z(z\ \mathbf{r}\ A \to A) \qquad \mathrm{et}(N)}{\mathrm{et}(N)\ \mathbf{r}\ A \to A} \qquad \begin{array}{c|}\ \\ \mathrm{IH} \\ \mathrm{et}(N)\ \mathbf{r}\ A\end{array}}{A}$$

Define inductively I (unary, on reals in $[-1,1]$) by the clauses

$$0 \in I,$$

$$d \in \mathrm{Sd} \to x \in I \to \frac{x+d}{2} \in I \quad (\mathrm{Sd}\text{: signed digit } -1,0,1).$$

The dual ${}^{\mathrm{co}}I$ of I is defined by its closure axiom

$$x \in {}^{\mathrm{co}}I \to x = 0 \vee \exists_{d,x'}\left(d \in \mathrm{Sd} \wedge x' \in {}^{\mathrm{co}}I \wedge x = \frac{x'+d}{2}\right)$$

Coinduction (or gfp-axiom): every "competitor" satisfying the same closure axiom is contained in ${}^{\mathrm{co}}I$.

Theorem (Av)

$$x, y \in {}^{\mathrm{co}}I \to \frac{x+y}{2} \in {}^{\mathrm{co}}I$$

Proof (Berger & Seisenberger, 2010). Let

$$P := \left\{ \frac{x+y}{2} \,\middle|\, x, y \in {}^{\mathrm{co}}I \right\}, \quad Q := \left\{ \frac{x+y+i}{2} \,\middle|\, x, y \in {}^{\mathrm{co}}I, i \in \mathrm{Sd}_2 \right\}$$

with $\mathrm{Sd}_2 := \{-2, -1, 0, 1, 2\}$. Then prove

$$P \subseteq Q \subseteq {}^{\mathrm{co}}I.$$

$P \subseteq Q$ is easy. For $Q \subseteq {}^{\mathrm{co}}I$ show that Q also satisfies the closure axiom for ${}^{\mathrm{co}}I$. Coinduction gives the claim.

The computational content of this proof is an algorithm operating on stream representations of reals.

- Realizers of $x \in {}^{\mathrm{co}}I$: finite or infinite lists of signed digits $\mathbb{L}(\mathbb{D})$.
- $\mathbb{L}(\mathbb{D})$ has constructors U (empty list) and $d :: u$.
- Define inductively $I^{\mathbf{r}}$ (n.c., binary) by

$$I^{\mathbf{r}}(0, U),$$
$$I^{\mathbf{r}}(x, u) \to I^{\mathbf{r}}\Big(\frac{x+d}{2}, d :: u\Big).$$

- The dual ${}^{\mathrm{co}}I^{\mathbf{r}}$ of $I^{\mathbf{r}}$ is defined by its closure axiom

$${}^{\mathrm{co}}I^{\mathbf{r}}(x, u) \to (x = 0 \land u = U) \lor$$
$$\exists_{d,x',u'}\Big({}^{\mathrm{co}}I^{\mathbf{r}}(x', u') \land x = \frac{x'+d}{2} \land u = d :: u'\Big).$$

Lemma (Init)

$$x, y \in {}^{\mathrm{co}}I \to \exists_{i,x',y'}\Big(i \in \mathrm{Sd}_2 \land x', y' \in {}^{\mathrm{co}}I \land \frac{x+y}{2} = \frac{x'+y'+i}{4}\Big).$$

Extracted term

$$\mathrm{Av}_{\mathrm{init}} : \mathbb{L}(\mathbb{D}) \to \mathbb{L}(\mathbb{D}) \to \mathbb{D}_2 \times \mathbb{L}(\mathbb{D}) \times \mathbb{L}(\mathbb{D})$$
$$(d :: u, e :: v) \mapsto \langle d + e, u, v \rangle.$$

Lemma (Step)

$$i \in \mathrm{Sd}_2 \to x, y \in {}^{\mathrm{co}}I \to \exists_{d \in \mathrm{Sd}} \exists_{j \in \mathrm{Sd}_2} \exists_{x', y' \in {}^{\mathrm{co}}I} \left(\frac{x+y+i}{4} = \frac{\frac{x'+y'+j}{4} + d}{2} \right).$$

Extracted term

$$\begin{aligned} &\mathrm{Av}_{\mathrm{step}} : \mathbb{D}_2 \to \mathbb{L}(\mathbb{D}) \to \mathbb{L}(\mathbb{D}) \to \mathbb{D} \times \mathbb{D}_2 \times \mathbb{L}(\mathbb{D}) \times \mathbb{L}(\mathbb{D}) \\ &(i, d :: u, e :: v) \mapsto \langle K(d+e+2i), J(d+e+2i), u, v \rangle. \end{aligned}$$

with $J, K : \mathbb{Z} \to \mathbb{Z}$ such that

$$\begin{aligned} &i = J(i) + 4K(i), \\ &|J(i)| \leq 2, \\ &|i| \leq 6 \to |K(i)| \leq 1. \end{aligned}$$

By coinduction from the Step Lemma we obtain

$$\exists_{i,x,y}\left(i \in \mathrm{Sd}_2 \land x, y \in {}^{\mathrm{co}}I \land z = \frac{x+y+i}{4}\right) \to z \in {}^{\mathrm{co}}I$$

Theorem (Av)

$$x, y \in {}^{\mathrm{co}}I \to \frac{x+y}{2} \in {}^{\mathrm{co}}I$$

Extracted term of type

$$\mathbb{L}(\mathbb{D}) \to \mathbb{L}(\mathbb{D}) \to \mathbb{L}(\mathbb{D})$$

defined using corecursion:

- From $u, v \in \mathbb{L}(\mathbb{D})$ form an initial triple $\mathrm{Av}_{\mathrm{init}}(u, v) \in \mathbb{D}_2 \times \mathbb{L}(\mathbb{D}) \times \mathbb{L}(\mathbb{D})$.
- Iterate $\mathrm{Av}_{\mathrm{step}} \colon \mathbb{D}_2 \to \mathbb{L}(\mathbb{D}) \to \mathbb{L}(\mathbb{D}) \to \mathbb{D} \times \mathbb{D}_2 \times \mathbb{L}(\mathbb{D}) \times \mathbb{L}(\mathbb{D})$ starting with $\mathrm{Av}_{\mathrm{init}}(u, v)$.
- Return the stream of generated $d \in \mathbb{D}$.

Initially we had

$$\mathrm{Av}_{\mathrm{init}}(d :: u, e :: v) := \langle \underbrace{d + e}_{\in \mathrm{Sd}_2}, u, v \rangle$$

and then

$$\mathrm{Av}_{\mathrm{step}}(i, d :: u, e :: v) := \langle \underbrace{K(d + e + 2i)}_{\in \mathrm{Sd}}, \underbrace{J(d + e + 2i), u, v}_{\text{triple for rec. call}} \rangle.$$

Hence the lookahead is $n + 1$. This can be expressed as

$$u|_{n+1} = u'|_{n+1} \to v|_{n+1} = v'|_{n+1} \to \mathrm{cAv}(u, v)|_n = \mathrm{cAv}(u', v')|_n$$

and proved by induction on n.

Idea (Nils Köpp): insert the expected lookahead into the goal. Replace the unary coinductive predicate ${}^{\mathrm{co}}I$ on reals by a binary inductive predicate L with the property that

> *a realizer of $x \in L_n$ is a list of length n of signed digits approximating x with error bound $\frac{1}{2^n}$.*

The intended meaning of L is expressed by its two clauses

$$|x| \le 1 \to x \in L_0, \tag{1}$$

$$d \in \mathrm{Sd} \to x \in L_n \to y = \frac{x+d}{2} \to y \in L_{n+1}. \tag{2}$$

(2) says that if we know n digits of a representation of x and $y = \frac{x+d}{2}$, then we know $n+1$ digits of a representation of y.

We proceed as before, but with L instead of ${}^{\mathrm{co}}I$:

Lemma (LInit)

$$x, y \in L_{n+1} \to \exists_{i,x',y'}\left(i \in \mathrm{Sd}_2 \land x', y' \in L_n \land \frac{x+y}{2} = \frac{x'+y'+i}{4}\right).$$

Extracted term

$$\begin{aligned}&\mathrm{LAv}_{\mathrm{init}} : \mathbb{L}(\mathbb{D}) \to \mathbb{L}(\mathbb{D}) \to \mathbb{D}_2 \times \mathbb{L}(\mathbb{D}) \times \mathbb{L}(\mathbb{D})\\ &(d :: u, e :: v) \mapsto \langle d + e, u, v\rangle.\end{aligned}$$

Same as for ${}^{\mathrm{co}}I$.

Lemma (LStep)

$$i \in \mathrm{Sd}_2 \to x, y \in L_{n+1} \to$$
$$\exists_{d \in \mathrm{Sd}} \exists_{j \in \mathrm{Sd}_2} \exists_{x', y' \in L_n} \Big(\frac{x+y+i}{4} = \frac{\frac{x'+y'+j}{4} + d}{2} \Big).$$

Extracted term

$$\mathrm{LAv}_{\mathrm{step}} \colon \mathbb{D}_2 \to \mathbb{L}(\mathbb{D}) \to \mathbb{L}(\mathbb{D}) \to \mathbb{D} \times \mathbb{D}_2 \times \mathbb{L}(\mathbb{D}) \times \mathbb{L}(\mathbb{D})$$
$$(i, d :: u, e :: v) \mapsto \langle K(d+e+2i), J(d+e+2i), u, v \rangle.$$

Same as for ${}^{\mathrm{co}}I$.

By induction on n from the Step Lemma we obtain

$$i \in \mathtt{Sd}_2 \to x, y \in L_n \to \frac{x + y + i}{4} \in L_n.$$

Extracted term:
Contains the recursion operator.

Theorem (LAv)

$$x, y \in L_{n+1} \to \frac{x+y}{2} \in L_n.$$

Proof. Induction on n.
Extracted term of type

$$\mathbb{N} \to \mathbb{L}(\mathbb{D}) \to \mathbb{L}(\mathbb{D}) \to \mathbb{L}(\mathbb{D})$$

defined using recursion:

- From $u, v \in \mathbb{L}(\mathbb{D})$ form an initial triple $\mathrm{LAv}_{\mathrm{init}}(u, v) \in \mathbb{D}_2 \times \mathbb{L}(\mathbb{D}) \times \mathbb{L}(\mathbb{D})$.
- Iterate $\mathrm{LAv}_{\mathrm{step}} \colon \mathbb{D}_2 \to \mathbb{L}(\mathbb{D}) \to \mathbb{L}(\mathbb{D}) \to \mathbb{D} \times \mathbb{D}_2 \times \mathbb{L}(\mathbb{D}) \times \mathbb{L}(\mathbb{D})$ n times starting with $\mathrm{LAv}_{\mathrm{init}}(u, v)$.
- Return the stream of generated $d \in \mathbb{D}$.

Same as for ${}^{\mathrm{co}}I$.

The lookahead analysis is built into the goal, i.e., the specification. There is no need for a lookahead analysis after the proof. Then why ${}^{\mathrm{co}}I$-proofs?

- ${}^{\mathrm{co}}I$-proofs can be done without knowledge of a lookahead.
- From them we may guess a lookahead and then do a similar L-proof, replacing coinduction by induction.
- For multiplication this worked as well, with lookahead $3n + 3$.

- L-proofs can also be done without a previous ${}^{\mathrm{co}}I$-proof.
- One then needs a direct intuition about the lookahead.

Example: multiplication.

- There the L-proof is based on the decomposition

$$\frac{\frac{x'+d'}{2}+d}{2}\cdot\frac{\frac{y'+e'}{2}+e}{2}=\frac{\frac{\frac{x'y'+d'e'}{2}+de'}{2}+de}{2}+\frac{\frac{\frac{e'x'+d'y'}{2}+d'e}{2}+\frac{dy'+ex'}{2}}{2}}{2}$$

  (easy calculation).
- The r.h.s. uses average 4 times.
- We obtain an L-proof with lookahead $n+6$.

Goal: formal proof (in TCF) that the extracted term is a realizer.
Minlog implements

- generation of a soundness proof: `add-sound`
- proof checking: `check-proof`, `check-and-display-proof`

```
(add-sound "LAverage")

ok, LAverageSound has been added as a new theorem:

allnc x,y,n,u^(
 LMR x(Succ n)u^ ->
 allnc u^0(LMR y(Succ n)u^0 ->
           LMR((1#2)*(x+y))n(cLAverage n u^ u^0)))

with computation rule

cLAverage eqd([n,u,u0]cLAvcToL n(cLAvToAvc u u0))
```

```
(cp (theorem-name-to-proof "LAverageSound"))

Ok, proof is correct.
No free assumption variables.
No global assumptions used.
No implicit global assumptions (unproven rewrite rules).
Theorems used: (LCompatSound RealEqSym LAvcToLSound
                LAvToAvcSound LToReal)
```

Future work

- Study of extracted terms (i.e., implicit computational content) for proofs in constructive analysis (integration, ode's).
- Compare the Soundness Theorem with MetaCoq.
- Exact real arithmetic in Agda, "MetaAgda"?