Proof mining and families of mappings

Andrei Sipoș

Institute of Mathematics of the Romanian Academy

May 11, 2016
PhDs in Logic VIII
Darmstadt, Germany

Proof mining

Proof mining (introduced and developed by U. Kohlenbach) aims to obtain quantitative information from proofs of theorems (from various areas of mathematics) of a nature which is not (fully) constructive. A comprehensive reference is:

U. Kohlenbach, *Applied proof theory: Proof interpretations and their use in mathematics*, Springer, Berlin/Heidelberg, 2008.

We will first introduce the logical framework and theorems and then focus on an illustrative example in nonlinear analysis.

The logical systems

We generally use systems of arithmetic in all finite types, intuitionistic or classical, augmented by restricted non-constructive rules (such as choice principles) and by types referring to (metric/normed/Hilbert) spaces and functionals involving them.

Two such systems are denoted by $\mathcal{A}_i^\omega[X, \langle\cdot,\cdot\rangle, C]$ (intuitionistic) and $\mathcal{A}^\omega[X, \langle\cdot,\cdot\rangle, C]$ (classical).
One typically uses proof interpretations to extract the necessary quantitative information. Two sample metatheorems guaranteeing this fact were developed by Gerhardy and Kohlenbach in the 2000s. We will present simplified versions of them.

The main intuitionistic metatheorem for proof mining

The following theorem uses Kreisel's "modified realizability" interpretation, in its "monotone" variant introduced by Kohlenbach.

Theorem (Gerhardy and Kohlenbach, 2006)

Let σ, ρ, τ be types (subject to some restrictions). Let $s^{\sigma\to\rho}$ be a closed term of $\mathcal{A}_i^\omega[X, \langle\cdot,\cdot\rangle]$ and let A (resp. B) be an arbitrary formulas with only x, y, z, n (resp. x, y, z) free. If

$$\mathcal{A}_i^\omega[X, \langle\cdot,\cdot\rangle] \vdash \forall x^\sigma \forall y \leq_\rho s(x) \forall z^\tau (\neg B \to \exists n^0 A),$$

then there exists an extractable Gödel primitive recursive functional Φ such that for all $b \in \mathbb{N}$

$$\forall x^\sigma \forall y \leq_\rho s(x) \forall z^\tau \exists n \leq \Phi(x, b)(\neg B \to A)$$

holds in every model for the such-augmented logical system where C is b-bounded.

The main classical metatheorem for proof mining

The following theorem uses Gödel's functional interpretation, also in its "monotone" variant introduced by Kohlenbach, combined with the negative translation.

Theorem (Gerhardy and Kohlenbach, 2008)

Let σ, ρ, τ be types (again, subject to some restrictions). Let $s^{\sigma \to \rho}$ be a closed term of $\mathcal{A}^\omega[X, \langle\cdot,\cdot\rangle]$ and let $B_\forall(x^\sigma, y^\rho, z^\tau, u^0)$ (resp. $C_\exists(x^\sigma, y^\rho, z^\tau, v^0)$) be a $\forall$-formula with only x, y, z, u free (resp. an $\exists$-formula with only x, y, z, v free). If

$$\mathcal{A}^\omega[X, \langle\cdot,\cdot\rangle] \vdash \forall x^\sigma \forall y \leq_\rho s(x) \forall z^\tau (\forall u^0 B_\forall \to \exists v^0 C_\exists),$$

then there exists an extractable computable functional Φ such that for all $b \in \mathbb{N}$

$$\forall x^\sigma \forall y \leq_\rho s(x) \forall z^\tau (\forall u \leq \Phi(x) B_\forall \to \exists v \leq \Phi(x) C_\exists)$$

holds in every model for the corresponding system.

Fixed points

We now describe the specific nonlinear analysis problem used as an application.
Suppose we have a Hilbert space H, a closed, convex set $C \subseteq H$ and a finite family of self-mappings $(T_i : C \to C)_{1 \leq i \leq N}$ with a common fixed point.

The problem is: to find one common fixed point.

We will work with the following two conditions on a map $T : C \to C$.

Definition

The map T is called **nonexpansive** if for all $x, y \in C$, we have that $\|Tx - Ty\| \leq \|x - y\|$.

Definition (Browder and Petryshyn, 1967)

Let $k \in [0, 1)$. The map T is called k-**strictly pseudocontractive** if for all $x, y \in C$, we have that:

$$\|Tx - Ty\|^2 \leq \|x - y\|^2 + k\|(x - Tx) - (y - Ty)\|^2.$$

We can see that nonexpansive $\Leftrightarrow$ 0-strictly pseudocontractive.

The parallel algorithm

A way to solve the problem of finding the fixed point for mappings of these two classes is the *parallel algorithm*, a generalization of the ubiquitous Mann algorithm.

We define a **family of weights** to be a family $(\lambda_i^{(n)})_{\substack{n\in\mathbb{N}\\1\leq i\leq N}}$ that satisfies, for each i,

$$\inf_{n\in\mathbb{N}} \lambda_i^{(n)} > 0,$$

and for any $n \in \mathbb{N}$,

$$\sum_{i=1}^{N} \lambda_i^{(n)} = 1.$$

The parallel algorithm

Let $(\lambda_i^{(n)})_{\substack{n\in\mathbb{N}\\1\le i\le N}}$ be a family of weights, $(t_n)_{n\in\mathbb{N}} \subseteq (0,1)$ be a sequence and $x \in C$ be a point, called the *starting point*.
Write, for all $n \geq 0$:

$$A_n := \sum_{i=1}^{N} \lambda_i^{(n)} T_i.$$

We define the output of the parallel algorithm for this data to be the sequence $(x_n)_{n\in\mathbb{N}} \subseteq C$, defined recursively by:

$$x_0 := x$$

$$x_{n+1} := t_n x_n + (1 - t_n) A_n x_n$$

(Some authors use the "opposite convention", i.e. t_n is interswitched with $1 - t_n$.)
The Mann algorithm is the special case where $N = 1$ (and therefore all $\lambda_1^{(n)}$'s are 1). (This case was treated by Ivan and Leuștean (2015).)

Our framework

From now on, we will fix:

- a Hilbert space H;
- a closed, convex set $C \subseteq H$;
- a finite family of self-mappings $(T_i : C \to C)_{1 \leq i \leq N}$;
- a starting point $x \in C$;
- a sequence $(t_n)_{n \in \mathbb{N}} \subseteq (0, 1)$;
- a family of weights $\left(\lambda_i^{(n)}\right)_{\substack{n \in \mathbb{N} \\ 1 \leq i \leq N}}$.

In this context, we shall denote by $(x_n)_{n \in \mathbb{N}}$ the output of the parallel algorithm associated to the data above.

The concrete result that we shall analyse is the following.

Theorem (López-Acedo and Xu, 2007)

Let $k \in (0,1)$ and suppose that each T_i is k-strictly pseudocontractive, with $\bigcap_{i=1}^N Fix(T_i) \neq \emptyset$. Impose the conditions:

$$\sum_{n=0}^{\infty}(t_n - k)(1 - t_n) = \infty$$

and

$$\sum_{j=0}^{\infty}\sqrt{\sum_{i=1}^{N}|\lambda_i^{(j+1)} - \lambda_i^{(j)}|} < \infty.$$

Then the sequence $(x_n)_n$ weakly converges to a common fixed point of the family $(T_i : C \to C)_{1 \leq i \leq N}$.

Goal: asymptotic regularity

An intermediate result in the proof (which is often a crucial lemma for proving such convergence theorems) is the feature of $(x_n)_n$ to be (relatively to our context) *asymptotically regular*, i.e.

$$\lim_{n \to \infty} \|x_n - A_n x_n\| = 0.$$

It is for this result that we shall derive the quantitative version.

Also: we will at first analyse the proof for the easier case where the T_i's are nonexpansive (that is, $k = 0$).

What does "quantitative" mean?

Let us see what kind of information we might hope to extract. The limit before means that:

$$\forall \varepsilon > 0 \exists N_\varepsilon \forall N \geq N_\varepsilon (\|x_n - A_n x_n\| < \varepsilon).$$

What we want to get is a "formula" for N_ε in terms of (obviously) ε and of some other arguments parametrizing our situation (for those, see the next slide). Such a function is called a **rate of convergence** for the sequence.
We now see how the logical tools might help us.

Rates for the compact arguments

A feature of the logical metatheorems of proof mining is that universal premises can be added to the system without influencing the overall extraction. This addition is compulsory here, as the result above requires the convergence and divergence of some series – we must therefore use the divergence rate $\theta : \mathbb{N} \to \mathbb{N}$ and the Cauchy modulus $\gamma : (0, \infty) \to \mathbb{N}$, which satisfy for all $N \in \mathbb{N}$,

$$\sum_{n=0}^{\theta(N)} (t_n - k)(1 - t_n) \geq N.$$

and for all $n \in \mathbb{N}$ and $\varepsilon > 0$,

$$\sum_{j=\gamma(\varepsilon)+1}^{\gamma(\varepsilon)+n} \sqrt{\sum_{i=1}^{N} |\lambda_i^{(j+1)} - \lambda_i^{(j)}|} \leq \varepsilon.$$

to put these premises into an universal form.

Proof interpretations

We have mainly two proof interpretations at our disposal:

- monotone modified realizability, which:
    - can extract bounds for all kinds of formulas;
    - does not permit the use of excluded middle;
- monotone functional interpretation (combined with negative translation), which:
    - can extract bounds only for Π_2 (that is, $\forall\exists$) formulas;
    - permits the use of excluded middle.

The problem is that the proof of asymptotic regularity both uses excluded middle (reductio ad absurdum) *and* has a non-Π_2 result (we saw two slides ago that it is of a $\Pi_3/\forall\exists\forall$ form).
A similar problem was encountered and solved by Leuștean (2014) for the Ishikawa iteration.

The solution

The good news is that excluded middle is only used in the beginning of the proof to derive that:

$$\liminf_{n\to\infty} \|x_n - A_n x_n\| = 0,$$

which *is* a Π_2 statement, as it can be written as:

$$\forall \varepsilon > 0 \forall m \exists N \geq m (\|x_N - A_N x_N\| < \varepsilon).$$

A bound on the N above, depending on ε and m, will be called a *modulus of liminf* for the sequence.

The quantitative modulus of liminf

Monotone functional interpretation can therefore be used to obtain the following result.

Theorem (A.S.)

Let $b > 0$ and $p \in \bigcap_{i=1}^{N} Fix(T_i)$ such that $\|x\| \leq b$ and $\|x - p\| \leq b$. Set, for all $\varepsilon > 0$ and $m \in \mathbb{N}$:

$$\Delta_{b,\theta}(\varepsilon, m) := \theta(m + \lceil \frac{b^2}{\varepsilon^2} \rceil).$$

Then, for all $\varepsilon > 0$ and $m \in \mathbb{N}$, there is an $N \in [m, \Delta_{b,\theta}(\varepsilon, m)]$ such that $\|x_N - A_N x_N\| \leq \varepsilon$.

Switching the interpretation

The liminf statement, once the N has been bounded, becomes a (monotonely) purely universal statement, which can be added as an axiom to the remainder of the proof, analyzed using modified realizability. We obtain the following result.

Theorem (A.S.)

Set, for all $\varepsilon > 0$,

$$\Phi_{b,\theta,\gamma}(\varepsilon) := \Delta_{b,\theta}\left(\frac{\varepsilon}{2}, \gamma\left(\frac{\varepsilon}{6b}\right) + 1\right) = \theta(\gamma\left(\frac{\varepsilon}{6b}\right) + \lceil\frac{4b^2}{\varepsilon^2}\rceil + 1).$$

Then for all $\varepsilon > 0$ *we have that for all* $n \geq \Phi_{b,\theta,\gamma}(\varepsilon)$, $\|x_n - A_n x_n\| \leq \varepsilon$.

Remember, all this is for $k = 0$ (nonexpansivity)!

Towards strictly pseudocontractive maps

By doing the proof for the nonexpansive case, we have greatly streamlined it and removed the square root in the convergence assumption. We will now use a trick to prove it also for the general k-strictly pseudocontractive case.

Definition

If $T : C \to C$ is a mapping and $t \in [0, 1]$, define $T_t : C \to C$ as $T_t x := tx + (1 - t)Tx$.

Note that we used the same convention as before. We can easily check that $T_{1-t_1} T_{1-t_2} = T_{1-t_1 t_2}$, for all $t_1, t_2 \in [0, 1]$. A nontrivial property is the following:

Lemma (Browder and Petryshyn, 1967)

If T is k-strictly pseudocontractive, then T_k is nonexpansive.

Using these properties (and the fact that T_k has the same fixed points as T, for any T and k), we can derive the general quantitative version.

Theorem (A.S.)

Set, for all $\varepsilon > 0$,

$$\Phi'_{b,k,\theta,\gamma}(\varepsilon) := \theta(\gamma\left(\frac{\varepsilon(1-k)}{6b}\right) + \lceil \frac{4b^2}{(1-k)^2\varepsilon^2} \rceil + 1).$$

Then for all $\varepsilon > 0$ *we have that for all* $n \geq \Phi'_{b,k,\theta,\gamma}(\varepsilon)$, $\|x_n - A_n x_n\| \leq \varepsilon$.

Similar examples of reduction to nonexpansiveness can be seen in the paper arXiv:1605.02237.

Thank you for your attention.