Tactic and (co)algebraic reasoning

Rodrigo Raya

École Polytechnique Fédéral de Lausanne

January 9, 2020

Overview

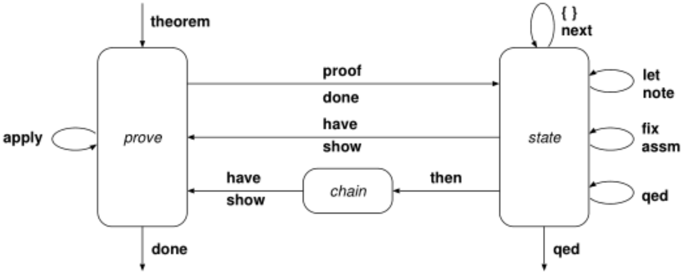
Goal

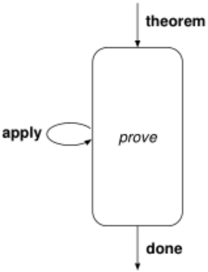
- How can I design new proof methods?
- What are some fancier proof methods?

Why tactics

- Widely adopted theorem proving methodology.
- Introduced by Milner in the Edinburgh LCF proof assistant.
- Isabelle provides them using three abstraction layers: Isabelle/ML, Isabelle/Isar and Eisbach.
- Here: Isabelle/ML. See report for technicalities.

Tactical theorem proving vs Isabelle/Isar





Demo I

- Hales.thy: *associativity* lemma, local_setups, concrete_assoc.

- Goal is to prove:

$$(p_1 \oplus_i p_2) \oplus_j p_3 = p_1 \oplus_k (p_2 \oplus_l p_3)$$

  where $i, j, k, l \in \{0, 1\}$ and $\oplus_0, \oplus_1$ are well-defined operations on elliptic curves points.

- Good example for the section on structured proofs in the Isabelle Cookbook.

- A challenging part (requires debugging): deduce what the rewrite tactic does internally.

Demo II

- Experiment3.thy: rewrite sine expressions whose arguments contain sums of multiples of π.

1. Define SIN_SIMPROC_ATOM x n = x + of_int n * pi.
2. Write a conversion sin_atom_conv rewriting of_int n * pi to SIN_SIMPROC_ATOM 0 n and everything else to SIN_SIMPROC_ATOM x 0.
3. Write a conversion that descends through +, applies sin_atom_conv to every atom, and then applies some kind of combination rule like:
   SIN_SIMPROC_ATOM x1 n1 + SIN_SIMPROC_ATOM x2 n2 = SIN_SIMPROC_ATOM (x1 + x2) (n1 + n2).
4. In the end, I have rewritten the original term to the form sin (SIN_SIMPROC_ATOM x n), and then I apply some suitable rule to that.

Some categorical notions

Let $F : Set \to Set$ be a functor.

- An F-**algebra** is a set A with a structure mapping $\alpha : F(A) \to A$.
- f is an F-**homomorphism** between (A, α) and (B, β) if this diagram commutes:

$$\begin{array}{ccc} F(A) & \xrightarrow{F(f)} & F(B) \\ \downarrow{\scriptstyle \alpha} & & \downarrow{\scriptstyle \beta} \\ A & \xrightarrow{\quad f \quad} & B \end{array}$$

- Set^F is the category formed with F-algebras and F-homomorphisms.
- A **subalgebra** of $\mathcal{A} = (A, s)$ is $S \subseteq A$ with $\beta_S : F(S) \to S$ where the inclusion $i : S \to A$ is a F-homomorphism.
- An **initial** F-**algebra** is an initial object in Set^F. It is unique up to isomorphism. The structure mapping is an isomorphism **(Lambeck)**.

- A relation $R \subseteq S \times T$ is an F-**congruence** if there exists an F-algebra structure (R, γ) such that the projections π_i are F-homomorphisms:

$$
\begin{tikzcd}
F(S) \arrow[d, "\alpha"] & F(R) \arrow[l, "F(\pi_1)"'] \arrow[r, "F(\pi_2)"] \arrow[d, dashed, "\gamma"] & F(T) \arrow[d, "\beta"] \\
S & R \arrow[l, "\pi_1"] \arrow[r, "\pi_2"'] & T
\end{tikzcd}
$$

- By reversing arrows: F-coalgebra, coalgebra homomorphism, category Set_F, terminal F-coalgebra, F-bisimulation.
- Diagonal of a set: $\Delta(A) = \{(a, a).a \in A\}$
- **Induction**: congruences on initial algebras contain Δ.
- **Coinduction**: bisimulations on final coalgebras contain Δ.
- Mathematical induction, streams (co)induction and least (greatest) fixed points characterizations are easy to derive.

To stress the capabilities of this model, note:

Concept	Category theory	Functional programming
Datatype	Initial algebra	$\begin{aligned} \text{datatype } T &= c_1 \; of \; A_1 \times T^{n_1} \\ &\vdots \\ &\mid = c_k \; of \; A_k \times T^{n_k} \end{aligned}$
Iteration	$\begin{tikzcd} 1+\mathbb{N} \arrow[r, "F(f)"] \arrow[d, "{[zero,succ]}"'] & 1+X \arrow[d, "\alpha"] \\ \mathbb{N} \arrow[r, dashed, "\exists! f"'] & X \end{tikzcd}$	$f(0) = \alpha(*)$ $f(x+1) = \alpha(f(x))$
Recursion	$\begin{tikzcd} 1+\mathbb{N} \arrow[r, dashed, "F\langle h,id\rangle"] \arrow[d, "{[zero,succ]}"'] & 1+B\times\mathbb{N} \arrow[d, "g"] \\ \mathbb{N} \arrow[r, dashed, "\exists! h"'] & B \end{tikzcd}$	$h(0) = g_1(*)$ $h(succ(n)) = g_2(h(n), n)$
Case analysis	$\begin{tikzcd} 1+\mathbb{N} \arrow[d, "{[zero,succ]}"'] \arrow[dr, "g"] & \\ \mathbb{N} \arrow[r, dashed, "\exists! h"'] & B \end{tikzcd}$	$h(0) = g_1(*)$ $h(succ(n)) = g_2(n)$

Existence of initial algebras

Theorem:
Let $\mathcal{C}$ be a category with initial object 0 and colimits for any ω-chain. If $F : \mathcal{C} \to \mathcal{C}$ preserves the colimit of the initial ω-chain, then the initial F-algebra is $\mu(F) = \mathrm{colim}_{n<\omega} F^n 0$.
Corollary:
Any polynomial functor on *Set* admits an initial algebra.
In Isabelle:

- Arbitrary limits require reasoning about infinite type families, which goes beyond HOL capabilities.
- This does not include functors of interest: finite powerset ('a fset), countable powerset ('a cset), finite multisets ('a multiset) or discrete probability distributions ('a pmf). Example: datatype 'a tree = Node 'a ('a tree fset)
- There exist results for bounded endofunctors. Both approaches are related by transfinite induction.

Isabelle approach

- HOL as a category: universe of types U as objects + functions between types as morphisms.
- A functor is a type constructor $(\alpha_1, \ldots, \alpha_n)F$ together with a mapping:

$$\mathrm{Fmap} : \overline{\alpha} \to \overline{\beta} \to \overline{\alpha}F \to \overline{\beta}F$$

  such that $\mathrm{Fmap}\ \mathrm{id} = \mathrm{id}$ and $\mathrm{Fmap}(\overline{g} \circ \overline{f}) = \mathrm{Fmap}\,\overline{g} \circ \mathrm{Fmap}\,\overline{f}$.
- An n-ary bounded natural functor is a tuple (F, Fmap, Fset, Fbd) where:
  - F is an n-ary type constructor.
  - $\mathrm{Fmap} : \overline{\alpha} \to \overline{\beta} \to \overline{\alpha}F \to \overline{\beta}F$
  - $\forall i \in \{1, \ldots, n\}.\ \mathrm{Fset}_i : \overline{\alpha}F \to \alpha_i$ set
  - Fbd is an infinite cardinal number.

  satisfying the following:
  - (F,Fmap) is a binary functor.
  - (F,Fmap) preserves weak pullbacks.
  - The following cardinal bound conditions hold: $\forall x : \overline{\alpha}F, i \in 1, \ldots, n. |\mathrm{Fset}_i\, x| \leq \mathrm{Fbd}$

- $\forall a \in \mathsf{Fset}_i\, x, i \in \{1, \ldots, n\}. f_i\, a = g_i\, a \implies \mathsf{Fmap}\, \overline{f}\, x = \mathsf{Fmap}\, \overline{g}\, x$
- $\mathsf{Fset}_i : \overline{\alpha} F \to \alpha_i$ set is a natural transformation from: $((\alpha_1, \ldots, \alpha_{i-1}, \_, \alpha_{i+1}, \ldots, \alpha_n)F, \mathsf{Fmap})$ to $(\mathsf{set}, \mathsf{image})$.

Shape and content intuition

- The definition of natural transformation for the inclusion mapping $f = i$ gives $\mathsf{Fmap}\, i = i$.
- So inclusion lifts to the inclusion.
- If (A, t) is a F-subalgebra of (B, s), then the inclusion $i : A \to B$ is an F-algebra homomorphism and $\mathsf{Fmap}\, i = i$.
- So the subalgebra equation simplifies to $s \circ i = i \circ t$ which implies that $t = s|_{F(A)}$.
- Thus, for a BNF, a subalgebra can be given by a subset together with a particular restriction of the structure mapping.

Constructing the initial algebra: minimal algebra

- Let $\mathcal{A} = (A, s)$ be an F-algebra.
  Set $M_s = \bigcap_{B.(B,s|_B)\text{is a subalgebra of (A,s)}} B$ then:

  $$\mathcal{M}(\mathcal{A}) = \left(M_s, s\Big|_{M_s}\right)$$

  is the F-subalgebra generated by $\emptyset$.

- We call it the minimal algebra generated by $\mathcal{A}$.

- $\mathcal{M}(\mathcal{A})$ is said to be the subalgebra generated by $\emptyset$ in the sense that it is the intersection of all subalgebras containing $\emptyset$.

Constructing the initial algebra: minimal algebra lemma

Lemma:
There exists at most one morphism from $\mathcal{M}(\mathcal{A})$ to any other F-algebra (Y, t).
Proof:
Since, if f, g are two such morphisms, we can show that:

$$B = \mathcal{M}(\mathcal{A}) \cap \{x \in \mathcal{A}.f(x)_= g(x)\}$$

is a F-subalgebra of $\mathcal{A} = (A, s)$. Indeed, by our remarks, it suffices to note that $M_s \cap \{x \in \mathcal{A}.f(x)_= g(x)\} \subseteq M_s$ and consider the structure map $s|_B$. This leads to a subalgebra of $\mathcal{M}(\mathcal{A})$ which can be naturally seen as a subalgebra of $\mathcal{A}$. By definition of $\mathcal{M}(\mathcal{A})$, $M_s \subseteq B$ and thus $\forall x \in M_s.f(x) = g(x)$. Thus, $f = g$.

Construction of the initial algebra: naive approach

1. Set $\mathcal{R} = \prod\{\mathcal{A}.\mathcal{A} \text{ is an algebra}\}$.
2. Given an algebra $\mathcal{A}$, note h the projection morphism from $\mathcal{R}$ to $\mathcal{A}$.
3. Then $h|_{\mathcal{M}(\mathcal{R})}$ is the unique morphism between $\mathcal{M}(\mathcal{R})$ and $\mathcal{A}$.
4. Since the construction does not depend on the chosen algebra $\mathcal{A}$, $\mathcal{M}(\mathcal{R})$ is the desired initial algebra.

Problems in HOL:

1. One cannot quantify over infinite type collections.
2. The product of the carrier sets of all algebras, fails itself to be a set.

Construction of the initial algebra: for a fixed algebra

- Given an F-algebra $\mathcal{A}$ we know that there exists at most one morphism $\mathcal{M}(\mathcal{A}) \to \mathcal{A}$. But from the shape and content intuition, for bounded natural functors, the inclusion is one such morphism. So there is exactly one morphism $g : \mathcal{M}(\mathcal{A}) \to \mathcal{A}$.
- Goal: give a set of algebras $\mathcal{R}$ such that from $\mathcal{R}$ there is a unique morphism to any $\mathcal{M}(\mathcal{A})$.
- Strategy: find a sufficiently large type T_0 such that its cardinality is an upperbound for any $\mathcal{A}$.

Construction of the initial algebra: for an arbitrary algebra

$$|A| \leq_o (r ::' b\,\text{set}) \implies \exists f\ B ::' b\,\text{set}.\text{bij\_betw} f\ B\ A \quad (\textit{ex\_bij\_betw})$$

If we can bound the cardinality of a set by some ordinal then the set has a bijective representation on the carrier of the wellorder inducing the ordinal.
For any algebra $\mathcal{A}$, with M the carrier of $\mathcal{M}(\mathcal{A})$, $|M| \leq_o 2 \wedge_c k$.
The package witnesses a type T_0 with this cardinality and sets:

$$\mathcal{R} = \prod\{\mathcal{A}.\mathcal{A} = (A, s) \text{ is an algebra with structure map } s : T_0 F \rightarrow T_0\}$$

By means of *ex_bij_betw*, the minimal algebras $\mathcal{M}(\mathcal{A})$ have isomorphic representants on a component of $\mathcal{R}$. Thus, the corresponding projection from the product to $\mathcal{M}(\mathcal{A})$ restricted to $\mathcal{M}(\mathcal{R})$ is the unique morphism f between the two.
Then, $f \circ g : \mathcal{M}(\mathcal{R}) \rightarrow \mathcal{A}$ is a suitable morphism. One shows it is the unique morphism between the two with a similar argument as the previous lemma.

Conclusion

- We have explored the Isabelle/ML: including tactics, parsing of specifications, new proof commands and definitional packages.
- We have seen a practical use of category theory in a real-world tool formalizing (co)datatypes as bounded natural functors.
- Next natural step: explore the logic foundations of several proof assistants.

References I

- Michael Barr and Charles Wells. *Category theory for computing science*. Vol. 49. Prentice Hall New York, 1990.
- Jasmin Christian Blanchette, Andrei Popescu, and Dmitriy Traytel. "Cardinals in Isabelle/HOL". In: *International Conference on Interactive Theorem Proving*. Springer. 2014, pp. 111–127.
- Klaus Denecke and Shelly L Wismath. *Universal algebra and coalgebra*. World Scientific, 2009.
- Herman Geuvers and Erik Poll. "Iteration and primitive recursion in categorical terms". In: (2007).
- Bart Jacobs. "Introduction to coalgebra". In: *Towards mathematics of states and observations* (2005).
- Jan Rutten. "The Method of Coalgebra: exercises in coinduction". In: (2019).

References II

Dmytro Traytel. "A category theory based (co)datatpye package for Isabelle/HOL". Master Thesis. TUM, 2012.

Markus M Wenzel. "Isabelle/Isar—a versatile environment for human-readable formal proof documents". PhD thesis. Technische Universität München, 2002.