Logipedia: a system-independent encyclopedia of formal proofs

Gilles Dowek

Formats

In the early ages: write a piece of software (for example: text processing system) chose a representation for the data
Involuntarily defined a format

In modern times: define a format first
ASCII, TCP / IP, HTTP, HTML, UNICODE...
Software has to comply to the format

But not yet in the realm of formal proofs: "A Coq proof of the four color theorem"

Problems: interoperability, sustainability

Why is it more difficult with formal proofs?

Because we cannot go too far

Euclidean geometry $\not\leftrightarrow$ Hyperbolic geometry

$\mathrm{ZF} \not\leftrightarrow \mathrm{ZFC}$

But...

A proof in ZF can be "translated" to ZFC

A proof in ZFC that does not use the axiom of choice can be "translated" to ZF

Proof transformation

There exists a basis of $\mathbb{R}^2$

- by the incomplete basis theorem (axiom of choice)
- $\langle 1, 0\rangle$, $\langle 0, 1\rangle$

automatically (for example: constructivization) or by hand

Reverse mathematics as the basis of interoperability

Reformulating the project of reverse mathematics

- Formal proofs, not pencil-paper-LaTeXones
- Expressive theories (Set theory, Type theory...) and not fragments of arithmetic
- Analyze proofs before (possibly) transforming them

Logical Frameworks

The interoperability ZF / ZFC possible because ZF and ZFC expressed in the same logical framework: predicate logic

In predicate logic, a theory: several axioms

Permits to raise the question: which axioms are used in a proof π

The revolution of predicate logic

Since Euclid: geometry, arithmetic, set theory... each system its syntax, its notion of proof...

Hilbert and Ackermann (1928): a common predicate logic

A common framework for geometry, arithmetic, set theory...
Sharing connectives, deduction rules...

A theory: symbols and axioms

But a short revolution

Predicate logic: simplification of Type theory (*Principia Mathematica*)
But no reformulation of Type theory in predicate logic

Soon (1940) Church: a new formulation of Type theory (based on λ-calculus) impossible to express in predicate logic (λ binds)

1970, 1985... Martin-Löf's type theory, the Calculus of constructions... not in predicate logic

Three attitudes

- Consider logical framework as a dead concept
- Express Russell's type theory, Church's, Martin-Löf's, the Calculus of constructions... in predicate logic by will of by force (Henkin, Davis, D...)
- Extend predicate logic to a better logical framework

The limits of predicate logic

- No bound variables ($\lambda x\ x$)
- No syntax for proofs
- No notion of computation
- No good notion of cut
- Classical and not constructive

New logical frameworks

- No bound variables ($\lambda x\ x$): λ-Prolog, Isabelle, $\lambda\Pi$-calculus
- No syntax for proofs: $\lambda\Pi$-calculus
- No notion of computation: Deduction modulo theory
- No good notion of cut: Deduction modulo theory
- Classical and not constructive: Ecumenical logic

The $\lambda\Pi$-calculus modulo theory that generalizes them all
DEDUKTI: an implementation of it

Defining a theory in Dedukti

No universal method
But several paradigmatic examples

- Any (finite) theory expressed in Predicate logic
- Axiom schemes
- Simple type theory (without and with polymorphism)
- Pure type systems (CoC...)
- Inductive types
- Universes

Ongoing: universe polymorphism, proof irrelevance, predicate subtyping

Simple type theory in Dedukti

$$
\begin{array}{rcl}
\mathit{type} & : & \mathit{Type} \\
\eta & : & \mathit{type} \to \mathit{Type} \\
o & : & \mathit{type} \\
\mathit{nat} & : & \mathit{type} \\
\mathit{arrow} & : & \mathit{type} \to \mathit{type} \to \mathit{type} \\
\varepsilon & : & (\eta\ o) \to \mathit{Type} \\
\Rightarrow & : & (\eta\ o) \to (\eta\ o) \to (\eta\ o) \\
\forall & : & \Pi x : \mathit{type}\ (((\eta\ x) \to (\eta\ o)) \to (\eta\ o))
\end{array}
$$

$$
\begin{array}{rcl}
(\eta\ (\mathit{arrow}\ x\ y)) & \longrightarrow & (\eta\ x) \to (\eta\ y) \\
(\varepsilon\ (\Rightarrow\ x\ y)) & \longrightarrow & (\varepsilon\ x) \to (\varepsilon\ y) \\
(\varepsilon\ (\forall\ x\ y)) & \longrightarrow & \Pi z : (\eta\ x)\ (\varepsilon\ (y\ z))
\end{array}
$$

The Calculus of constructions in DEDUKTI

$$
\begin{array}{rcl}
\mathit{type} & : & \mathit{Type} \\
\eta & : & \mathit{type} \to \mathit{Type} \\
o & : & \mathit{type} \\
\mathit{nat} & : & \mathit{type} \\
\mathit{arrow} & : & \Pi x : \mathit{type}\ (((\eta\ x) \to \mathit{type}) \to \mathit{type}) \\
\varepsilon & : & (\eta\ o) \to \mathit{Type} \\
\Rightarrow & : & \Pi x : (\eta\ o)\ (((\varepsilon\ x) \to (\eta\ o)) \to (\eta\ o)) \\
\forall & : & \Pi x : \mathit{type}\ (((\eta\ x) \to (\eta\ o)) \to (\eta\ o)) \\
\pi & : & \Pi x : (\eta\ o)\ (((\varepsilon\ x) \to \mathit{type}) \to \mathit{type})
\end{array}
$$

$$
\begin{array}{rcl}
(\eta\ (\mathit{arrow}\ x\ y)) & \longrightarrow & \Pi z : (\eta\ x)\ (\eta\ (y\ z)) \\
(\varepsilon\ (\Rightarrow\ x\ y)) & \longrightarrow & \Pi z : (\varepsilon\ x)\ (\varepsilon\ (y\ z)) \\
(\varepsilon\ (\forall\ x\ y)) & \longrightarrow & \Pi z : (\eta\ x)\ (\varepsilon\ (y\ z)) \\
(\eta\ (\pi\ x\ y)) & \longrightarrow & \Pi z : (\varepsilon\ x)\ (\eta\ (y\ z))
\end{array}
$$

A comparison

- *arrow* dependent in the Calculus of constructions but not in Simple type theory
- Same for $\Rightarrow$
- An extra symbol π in the Calculus of constructions: express functions mapping proofs to terms

Reverse mathematics in Dedukti

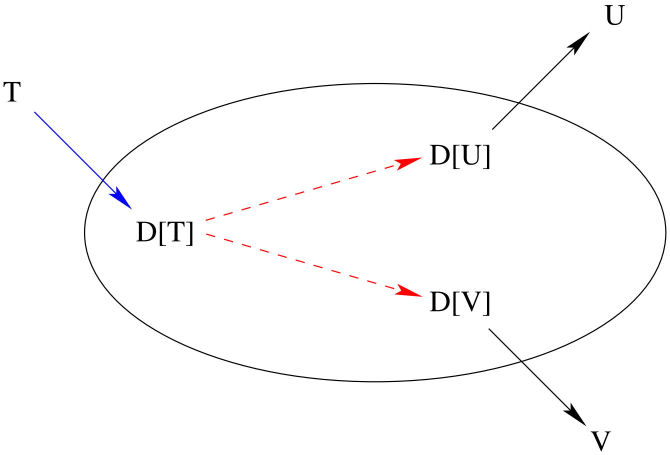
- All proofs in Simple type theory can be translated to the Calculus of constructions
- The proofs in the Calculus of constructions that do not use these three features can be translated to Simple type theory

(not the others: genuine Calculus of constructions proofs)

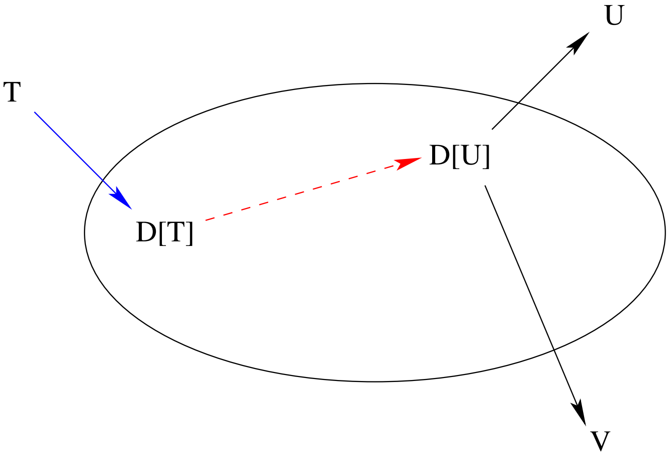
For example: all the proofs of the arithmetic library of Matita

"First" proof of Fermat's little theorem in constructive Simple type theory (further: predicative, PA, fragments of PA...)

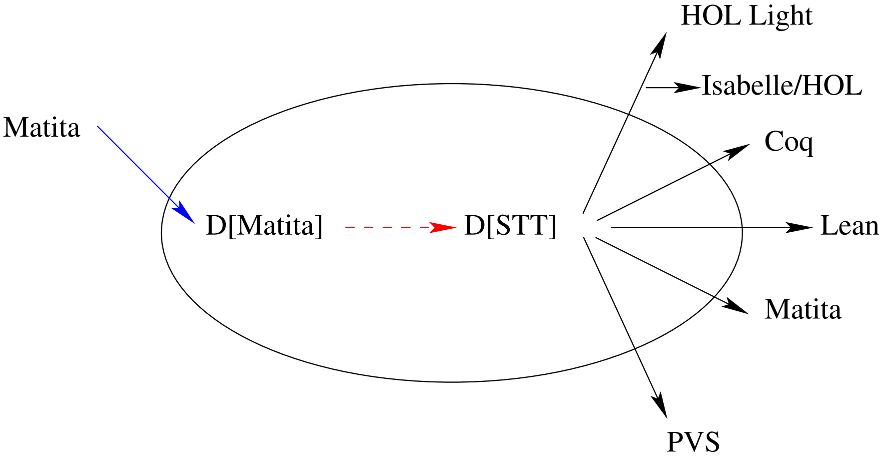
Proof translation



But also



An example


HOL Light
Isabelle/HOL
Matita
Coq
D[Matita]
D[STT]
Lean
Matita
PVS

Why does it work so well?

Because proof systems implement very expressive theories and use only a tiny part of it

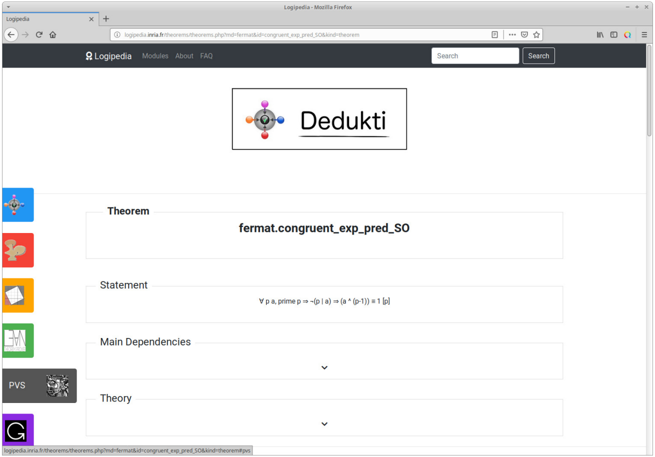
Three early empirical evidences

- Proof systems: very different theories, but very similar libraries
- Mathematicians are not very interested in the axioms used in their proofs: any theory seems to fit
- Mathematician are not even interested in definitions (real numbers must be constructed, but who cares how)

Collecting all the proofs in a single data base

LOGIPEDIA: an encyclopedia of proofs expressed

- in various theories
- in DEDUKTI



http://logipedia.science

Towards concept alignment in Logipedia

Connectives and quantifiers: inductive types / Q_0
Should be ignored by the library

Making formal the saying: Cauchy sequences or Dedekind cuts immaterial (isomorphic and only structural statements)

But classical and constructive disjunctions (ecumenical logic)

Already concrete results

While QED (1993) did not go very far

- Better understanding of the theories implemented in the various proof systems
- A new logical framework to express the these theories
- Analyzing the proofs (reverse mathematics) before we share them (partial translations)

Interoperability is not just a question of committees, negotiations, and standards: it is a research problem