Computers in the service of logic: the Coq proof-assistant

Jacek Chrząszcz
Institute of Informatics
University of Warsaw

Happy Birthday Marek!!!!
Proofs…

- writing proofs is hard (boring details)
- reading proofs is hard (boring details)
- some proofs and theorems are wrong
List of published false theorems (from Wikipedia):

• In 1803, Gian Francesco Malfatti proved that a certain arrangement of three circles would cover the maximum possible area inside a right triangle. It was shown in 1930 that circles in a different configuration could cover a greater area, and in 1967 that Malfatti’s configuration was never optimal.

• In his *Cours d’analyse* of 1821, Cauchy “proved” that if a sum of continuous functions converges pointwise, then its limit is also continuous. However, Abel observed *three years later* that this is not the case. For the conclusion to hold, “continuous” must be replaced with “uniformly continuous”.
• Kurt Gödel proved in 1932 that the truth of a certain class of sentences of first-order arithmetic, known in the literature as $[\exists^*\forall^2\exists^*, \text{all}, (0)]$, was decidable. In the final sentence of that paper, he asserted that the same proof would work for the decidability of the larger class $[\exists^*\forall^2\exists^*, \text{all}, (0)]_\equiv$. However, in the mid-1960s, Stål Aanderaa showed that Gödel’s proof would not go through for the larger class, and in 1982 Warren Goldfarb showed that validity of formulas from the larger class was in fact undecidable.

• In 1961, Jan-Erik Roos published an incorrect theorem about the vanishing of the first derived functor of the inverse limit functor under certain general conditions. However, over forty years later, Amnon Neeman and Pierre Deligne constructed a counterexample.
Let computers verify proofs!

- Computers are patient
- Computers are fast

BUT

- Computers do not understand English
- Computers are stupid
Proof assistants

They use formal logical systems.

Once a theorem is formalized as a formula (main goal)

- the user may ask the system to try to prove the goal
- the user writes a proof-step
- the system verifies the step is correct
- the system returns new (sub)goals to prove
- if there is nothing left to prove — success!

In some proof assistants the complete proof is re-checked.
Coq proof assistant

Based on higher-order logic (calculus of constructions) + inductive definitions

\[ \lambda \rightarrow \lambda \omega \]

\[ \lambda P \]

\[ \lambda P \omega \]

\[ F \]

\[ F_\omega \]

\[ \lambda P^2 \]

\[ \lambda \omega \]

\[ \lambda P \]

\[ CC \]

↑ polymorphism

↗ type constructors

→ dependent types
Proving in Coq

Curry-Howard Isomorphism

\[
\begin{array}{ccc}
\text{proof} & \text{of given} & \text{formula} \\
\downarrow & \downarrow \\
\text{term} & \text{of given} & \text{type}
\end{array}
\]

Lemma A3 : \((A \to B \to C) \to (A \to B) \to (A \to C)\)

Proof.

auto.

Qed.

proof-term: \(\lambda x^{A\to B\to C} \lambda y^{A\to B} \lambda z^A \ x z(y z)\)
DEMO
Coq — architecture

Written in Objective Caml (200 KLOC)

⇒ “logical” part (12 KLOC)
  • definitions, inductives, etc.
  • typing algorithm

⇒ extra-logical part (190 KLOC)
  • proof mode, tactics
  • helpers: implicit arguments, coercions
  • parsing, pretty-printing, etc.
  • disk operations
Coq – history

1984  CoC - calculus of constructions - G. Huet, T. Coquand

1989  first public release (version 4.10)

1991  Coq - calculus of inductive constructions - C. Paulin
      (version 5.6)

2000  version 7.0 with new (safe) architecture

2003  version 7.4 with modules!

2004  version 8.0 with new syntax

2009  version 8.2pl1 with “type classes” (‘a la Haskell)
Important formalisations in Coq

- Fundamental Theorem of Algebra, Nijmegen 2000
- Formalizacja platformy JavaCard, Trusted Logic 2003
  
  September 2007: a big step in program certification in the real world: The Technology and Innovation group at Gemalto has successfully completed a Common Criteria (CC) evaluation on a Java Card based commercial product. This evaluation is the world’s first CC certificate of a Java product involving EAL7 components (the official press release).

- Four-colour theorem, Cambridge 2004
- The CompCert verified C compiler, INRIA 2005-2010
- Gödel’s first incompleteness theorem, Nijmegen 2005
Coq — research subjects

- formalization of theorems / proofs of programs
- writing libraries and / or tactics [lists, group theory]
- writing tools for Coq [Web interface, doc generator]
- developing new features of Coq [modules, type classes]
- extending the Coq theory [equality]
Other proof assistants

- Automath (1967) - N.G.de Bruijn, Eindhoven
- Mizar (1973) - Trybulec, Białystok - the largest library
- LCF (1972) - Milner, currently HOL4 (1988), HOL Light
  ⇒ Isabelle (≤ 1989), Paulson, Cambridge
  ⇒ PVS (≤ 1992), Rushby, Shankar, Owre, Stanford
  ⇒ Coq (≤ 1980), Huet, Coquand, (Paulin, Herbelin), France
Beyond proof-assistants

Automatic provers

- model-checkers (SMV, SPIN, ...)
- logical frameworks (Twelf, ...)
- automated theorem provers (Simplify, Vampyre, Otter, ...)

Biggest challenges:

- “Kepler Conjecture”, problem: 1611, proof: 1998 (Hales), published: 2005 (after 4 years of “verification” work!) incl. 3GB of computer programs, data and results.
  - FlySpeck project (estimated effort: 20 years)

- Fermat Last Theorem, problem 1637, proof 1994, published 1995 (few months of verification)
  - Web page (estimated effort: 50 years)