Goal Translation for a Hammer for Coq (Extended Abstract)

Łukasz Czajka
lukasz.czajka@uibk.ac.at
University of Innsbruck, Austria

Cezary Kaliszyk
cezary.kaliszyk@uibk.ac.at
University of Innsbruck, Austria

Hammers are tools that provide general purpose automation for formal proof assistants. Despite the gaining popularity of the more advanced versions of type theory, there are no hammers for such systems. We present an extension of the various hammer components to type theory: (i) a translation of a significant part of the Coq logic into the format of automated proof systems; (ii) a proof reconstruction mechanism based on a Ben-Yelles-type algorithm combined with limited rewriting, congruence closure and a first-order generalization of the left rules of Dyckhoff's system LJT.

1 Introduction

Justifying small proof steps is usually a significant part of the process of formalizing proofs in an *interactive theorem proving* (ITP), or *proof assistant*, system. Many of such goals would be considered trivial by mathematicians. Still, state-of-the-art ITPs require the user to spend an important part of the formalization effort on them. The main points that constitute this effort are usually library search, minor transformations on the already proved theorems (such as reordering assumptions or reasoning modulo associativity-commutativity), as well as combining a small number of simple known lemmas. To reduce this effort various automation techniques have been conceived, including techniques from automated reasoning and domain specific decision procedures. The strongest general propose automation technique, available for various interactive theorem provers today is provided by "hammers" [10].

Hammers are proof assistant tools that employ external automated theorem provers (ATPs) in order to automatically find proofs of user given conjectures. There are three main components of a hammer:

- Lemma selection (also called relevance filtering or premise selection) that heuristically chooses a subset of the accessible lemmas that are likely useful for the given conjecture.
- Translation (encoding) of the user given conjecture together with the selected lemmas to the logics and input formats of automated theorem provers (ATPs). The focus is usually on first-order logic as the majority of the most efficient ATPs today support this foundation. The automated systems are in turn used to either find an ATP proof or just further narrow down the subset of lemmas to precisely those that are necessary in the proof.
- Proof reconstruction, which uses the obtained information from the successful ATP run, to reprove the lemma in the logic of the proof assistant.

Robust hammers exist for proof assistants based on higher-order logic (Sledgehammer [27] for Isabelle/HOL [33], HOLyHammer [20] for HOL Light [18] and HOL4 [31]) or dependently typed set theory (MizAR [21] for Mizar [7, 34, 6]). The general-purpose automation provided by the most advanced hammers is able to solve 40–50% of the top-level goals in various developments [10], as well as more than 70% of the user-visible subgoals [11], and as such has been found very useful in various proof developments [17].

Despite the gaining popularity of the more advanced versions of type theory, implemented by systems such as Agda [12], Coq [8], Lean [25], and Matita [4], there are no hammers for such systems. The

Submitted to:
HaTT 2016

construction of such a tool has so far been hindered by the lack of a usable encoding component, as well as by comparatively weak proof reconstuction.

For the proof assistants whose logics are based on the Calculus of Constructions and its extensions, the existing encodings in first-order logic so far cover only limited fragments of the source logic [1, 32, 9]. Why3 [16] provides a translation from its own logic [15] (which is a subset of the Coq logic, including features like rank-1 polymorphism, algebraic data types, recursive functions and inductive predicates) to the format of various first-order provers (in fact Why3 has been initially used as a translation back-end for HOLyHammer). Recently, an encoding of the dependently typed higher-order logic of F* into first-order logic has also been developed [2].

The built-in HOL automation is able to reconstruct the majority of the automatically found proofs using either internal proof search [19] or source-level reconstruction. The internal proof search mechanisms provided in Coq, such as the `firstorder` tactic [13], have been insufficient for this purpose so far. Matita's ordered paramodulation [5] is able to reconstruct many goals with up to two or three premises, and the congruence-closure based internal automation techniques in Lean [24] are also promising.

The SMTCoq [3] project has developed an approach to use external SAT and SMT solvers and verify their proof witnesses. Small checkers are implemented using reflection for parts of the SAT and SMT proof reconstruction, such as one for CNF computation and one for congruence closure. The procedure is able to handle Coq goals in the subset of the logic that corresponds to the logics of the input systems.

Contributions. We present our recently developed proof advice components for type theory and systems based on it. We first introduce an encoding of the Calculus of Inductive Constructions, including the additional logical constructions introduced by the Coq system, in untyped first-order logic with equality. We implement the translation and evaluate it experimentally on the standard library of the Coq proof assistant. We advocate that the encoding is sufficient for a hammer system for Coq: the success rates are comparable to those demonstrated by early hammer systems for Isabelle/HOL and Mizar, while the dependencies used in the ATP proofs are most often sufficient to prove the original theorems. Strictly speaking, our translation is neither sound nor complete. However, our experiments suggest that the encoding is "sound enough" to be usable. Moreover, we believe that a "core" version of the translation is sound and we are currently working on a proof of this fact.

Secondly, we present a proof reconstruction mechanism based on a Ben-Yelles-type procedure combined with a first-order generalization of the left rules of Dyckhoff's LJT, congruence closure and heuristic rewriting. With this still preliminary proof search procedure we are able to reprove almost 90% of the problems solved by the ATPs, using the dependencies extracted from the ATP output.

2 Translation

In this section we introduce an encoding of (a close approximation of) the Calculus of Inductive Constructions into untyped first-order logic with equality. The encoding should be a practical one, which implies that its general theoretical soundness is not the main focus, i.e., of course the translation needs to be "sound enough" to be usable, but it is more important that the encoding is efficient enough to provide practically useful information about the necessary proof dependencies. In particular, the encoding needs to be shallow, meaning that Coq terms of type Prop are translated directly to corresponding first-order formulas. Our translation is in fact unsound, e.g., it assumes proof irrelevance and ignores certain universe constraints. However, we believe that under the assumption of proof irrelevance a "core" version of the translation is sound, and we are currently working on a proof.

Below we present a variant of the translation for a fragment of the logic of Coq. The intention here is to provide a general idea, but not to describe the encoding in detail. In the first-order language we assume a unary predicate P, a binary predicate T and a binary function symbol $@$. Usually, we write ts instead of $@(t,s)$.

For the sake of efficiency, terms of type Prop are encoded directly as FOL formulas using a function $\mathscr{F}$. Terms that have type Type but not Prop are encoded using a function $\mathscr{G}$ as guards which essentially specify what it means for an object to have the given type. For instance, $\forall f:\tau.\varphi$ where $\tau=\Pi x:\alpha.\beta$ is translated to $\forall f.\mathscr{G}(\tau,f)\to\mathscr{F}(\varphi)$ where $\mathscr{G}(\tau,f)=\forall x.\mathscr{G}(\alpha,x)\to\mathscr{G}(\beta,fx)$. So $\mathscr{G}(\tau,f)$ says that an object f has type $\tau=\Pi x:\alpha.\beta$ if for any object x of type α, the application fx has type β. Function $\mathscr{F}$ encoding propositions as FOL formulas is defined by:

- If $\Gamma\vdash t:\text{Prop}$ then $\mathscr{F}_\Gamma(\Pi x:t.s)=\mathscr{F}_\Gamma(t)\to\mathscr{F}_{\Gamma,x:t}(s)$.
- If $\Gamma\nvdash t:\text{Prop}$ then $\mathscr{F}_\Gamma(\Pi x:t.s)=\forall x.\mathscr{G}_\Gamma(t,x)\to\mathscr{F}_{\Gamma,x:t}(s)$.
- Otherwise, if none of the above apply, $\mathscr{F}_\Gamma(t)=P(\mathscr{C}_\Gamma(t))$.

Function $\mathscr{G}$ encoding types as guards is defined by:

- If $t=\Pi x:t_1.t_2$ and $\Gamma\vdash t_1:\text{Prop}$ then $\mathscr{G}_\Gamma(\Pi x:t_1.t_2,s)=\mathscr{F}_\Gamma(t_1)\to\mathscr{G}_{\Gamma,x:t_1}(t_2,s)$.
- If $t=\Pi x:t_1.t_2$ and $\Gamma\nvdash t_1:\text{Prop}$ then $\mathscr{G}_\Gamma(\Pi x:t_1.t_2,s)=\forall x.\mathscr{G}_\Gamma(t_1,x)\to\mathscr{G}_{\Gamma,x:t_1}(t_2,sx)$.
- Otherwise, when t is not a product $\mathscr{G}_\Gamma(t,s)=T(u,\mathscr{C}_\Gamma(t))$.

Function $\mathscr{C}$ encoding terms as FOL terms is defined by:

- $\mathscr{C}_\Gamma(b)=b$ for b being a variable or a constant,
- $\mathscr{C}_\Gamma(ts)$ is equal to:
  - $\mathscr{C}_\Gamma(t)$ if $\Gamma\vdash s:A:\text{Prop}$ for some A,
  - $\mathscr{C}_\Gamma(t)\mathscr{C}_\Gamma(s)$ otherwise.
- $\mathscr{C}_\Gamma(\Pi x:t.s)=P\vec{y}$ for a fresh constant P where $\vec{y}=\text{FV}(\Pi x:t.s)$ and
  - if $\Gamma\vdash(\Pi x:t.s):\text{Prop}$ then $\forall\vec{y}.P\vec{y}\leftrightarrow\mathscr{F}_\Gamma(\Pi x:t.s)$ is a new axiom,
  - if $\Gamma\nvdash(\Pi x:t.s):\text{Prop}$ then $\forall\vec{y}z.P\vec{y}z\leftrightarrow\mathscr{G}_\Gamma(\Pi x:t.s,z)$ is a new axiom.
- $\mathscr{C}_\Gamma(\lambda\vec{x}:\vec{t}.s)=F\vec{y}$ where s does not start with a lambda-abstraction any more, F is a fresh constant, $\vec{y}=\text{FV}(\lambda\vec{x}:\vec{t}.s)$ and $\forall\vec{y}.\mathscr{F}_\Gamma(\forall\vec{x}:\vec{t}.F\vec{y}\vec{x}=s)$ is a new axiom.
- $\mathscr{C}_\Gamma(\texttt{case}(t,c,n,\lambda\vec{a}:\vec{\alpha}.\lambda x:c\vec{p}\vec{a}.\tau,\lambda\vec{x_1}:\vec{\tau_1}.s_1,\ldots,\lambda\vec{x_k}:\vec{\tau_k}.s_k))=F\vec{y_1}\vec{y_2}$ for a fresh constant F where
  - $I(c:\gamma:\kappa:=c_1:\gamma_1:\kappa_1,\ldots,c_k:\gamma_k:\kappa_k)\in E$,
  - $\Gamma_2=\vec{y_2}:\vec{\rho_2}=\text{FC}(\Gamma;t)$,
  - $\Gamma_1=\vec{y_1}:\vec{\rho_1}=\text{FC}(\Gamma;\lambda\vec{y_2}:\vec{\rho_2}.t(\lambda\vec{x_1}:\vec{\tau_1}.s_1)\ldots(\lambda\vec{x_k}:\vec{\tau_k}.s_k))$,
  - $\gamma_i=\Pi\vec{z_i}:\vec{\beta_i}.\Pi\vec{x_i}:\vec{\tau_i}.\sigma_i$ for $i=1,\ldots,k$,
  - the following is a new axiom:

$$\begin{array}{lll}\forall\vec{y_1}.\mathscr{F}_{\Gamma_1}(\forall\vec{y_2}:\vec{\rho_2} & . & (\exists\vec{z_1}:\vec{\beta_1}.\exists\vec{x_1}:\vec{\tau_1}.t=c_1\vec{z_1}\vec{x_1}\wedge F\vec{y_1}\vec{y_2}=s_1)\\ & \vee & \ldots\\ & \vee & (\exists\vec{z_k}:\vec{\beta_k}.\exists\vec{x_k}:\vec{\tau_k}.t=c_k\vec{z_k}\vec{x_k}\wedge F\vec{y_1}\vec{y_2}=s_k))\end{array}$$

Here t is the term matched on, the type of t has the form $c\vec{p}\vec{u}$, the integer n denotes the number of parameters (which is the length of $\vec{p}$), the type $\tau[\vec{u}/\vec{a},t/x]$ is the return type, i.e., the type of the whole case expression, $\vec{a}\cap \mathrm{FV}(\vec{p}) = \emptyset$, and $s_i[\vec{v}/\vec{x_i}]$ is the value of the case expression if the value of t is $c_i\vec{p}\vec{v}$. The free variable context $\mathrm{FC}(\Gamma;t)$ of t in Γ is defined inductively: $\mathrm{FC}(\emptyset;t) = \emptyset$; $\mathrm{FC}(\Gamma,x:\tau;t) = \mathrm{FC}(\Gamma;\lambda x:\tau.t),x:\tau$ if $x \in \mathrm{FV}(t)$; and $\mathrm{FC}(\Gamma,x:\tau;t) = \mathrm{FC}(\Gamma;t)$ if $x \notin \mathrm{FV}(t)$.

In the data exported from Coq there are three types of declarations: definitions, typing declarations and inductive declarations. We briefly describe how all of them are translated.

A definition $c = t:\tau:\kappa$ is translated as follows.

- If $\kappa = \mathrm{Prop}$ then add $\mathscr{F}(\tau)$ as a new axiom with label c.
- If $\kappa \neq \mathrm{Prop}$ then
  - add $\mathscr{G}(\tau,c)$ as a new axiom,
  - if $\tau = \mathrm{Prop}$ then add $c \leftrightarrow \mathscr{F}(t)$ as a new axiom with label c,
  - if $\tau = \mathrm{Set}$ or $\tau = \mathrm{Type}$ then add $\forall f.cf \leftrightarrow \mathscr{G}(t,f)$ as a new axiom with label c,
  - if $\tau \notin \{\mathrm{Prop},\mathrm{Set},\mathrm{Type}\}$ then add the equation $c = \mathscr{C}(t)$ as a new axiom with label c.

A typing declaration $c:\tau:\kappa$ is translated as follows.

- If $\kappa = \mathrm{Prop}$ then add $\mathscr{F}(\tau)$ as a new axiom with label c.
- If $\kappa \neq \mathrm{Prop}$ then add $\mathscr{G}(\tau,c)$ as a new axiom with label c.

An inductive declaration $I(c:\tau:\kappa := c_1:\tau_1:\kappa_1,\ldots,c_n:\tau_n:\kappa_n)$ is translated as follows.

- Translate the typing declaration $c:\tau:\kappa$.
- Translate each typing declaration $c_i:\tau_i:\kappa$ for $i = 1,\ldots,n$.
- Add axioms stating injectivity of constructors, axioms stating non-equality of different constructors, and the "inversion" axioms for elements of the inductive type.

For inductive types also induction principles and recursor definitions are translated.

The above only gives a general outline of the translation. In practice, we make a number of optimisations, e.g., the arity optimisation by Meng and Paulson [23], or translating fully applied functions with target type Prop directly to first-order predicates.

3 Reconstruction

We report on our work on proof reconstruction. We evaluate the Coq internal reconstruction mechanisms including `tauto` and `firstorder` [13] on the original proof dependencies and on the ATP found proofs, which are in certain cases more precise. In particular `firstorder` seems insufficient for finding proofs for problems created using the advice obtained from the ATP runs. This is partly caused by the fact that it does not fully axiomatize equality, but even on problems which require only purely logical first-order reasoning its running time is sometimes unacceptable.

The formulas that we attempt to reprove usually belong to fragments of intuitionistic logic low in the Mints hierarchy [29]. Most of proved theorems follow by combining a few known lemmas. This raises a possibility of devising an automated proof procedure optimized for these fragments of intuitionistic logic, and for the usage of the advice obtained from the ATP runs. We implemented a preliminary version of a Ben-Yelles-type procedure (essentially `eauto`-type proof search with a looping check) augmented with

Prover	Solved%	Solved	Sum%	Sum	Unique
Vampire	32.9	6839	32.9	6839	855
Z3	27.6	5734	34.9	7265	390
E Prover	25.8	5376	35.3	7337	72
any	35.3	7337	35.3	7337	

Table 1: Results of the experimental evaluation on the 20803 FOL problems generated from the propositions in the Coq standard library.

a first-order generalization of the left rules of Dyckhoff's system LJT [14], the use of the `congruence` tactic, and heuristic rewriting using equational hypotheses.

It is important to note that while the external ATPs we employ are classical and the translation assumes proof irrelevance, the proof reconstruction phase does not assume any additional axioms. We reprove the theorems in the intuitionistic logic of Coq, effectively using the output of the ATPs merely as hints for our hand-crafted proof search procedure. Therefore, if the ATP proof is inherently classical then proof reconstruction will fail. Currently, the only information from ATP runs we use is a list of lemmas needed by the ATP to prove the theorem (these are added to the context) and a list of constant definitions used in the ATP proof (we try unfolding these constants and no others).

Another thing to note is that we do not use the information contained in the Coq standard library during reconstruction. This would not make sense for our evaluation of the reconstruction mechanism, since we try to reprove the theorems from the Coq standard library. In particular, we do not use any preexisting hint databases available in Coq, not even the core database (we use the `auto` and `eauto` tactics with the `nocore` option). Also, we do not use any domain-specific decision procedures available as Coq tactics, e.g., `field`, `ring` or `omega`.

4 Evaluation

We evaluated our translation on the problems generated from all declarations of terms of type Prop in the Coq standard library of Coq version 8.5. We used the following classical ATPs: E Prover version 1.9 [30], Vampire version 4.0 [22] and Z3 version 4.0 [26]. The methodology was to measure the number of theorems that the ATP could reprove from their extended dependencies within a time limit of 30 s for each problem. The extended dependencies of a theorem are obtained by taking all constants occuring in the proof term of the theorem in Coq standard library, and recursively taking all constants occuring in the types and non-proof definitions of any dependencies extracted so far. Because of the use of extended dependencies, the average number of generated FOL axioms for a problem is 193. We limited the recursive extraction of extended dependencies to depth 2.

The evaluation was performed on a 48-core server with 2.2 GHz AMD Opteron CPUs and 320 GB RAM. Each problem was always assigned one CPU core. Table 1 shows the results of our evaluation. The column "Solved%" denotes the percentage (rounded to the first decimal place) of the problems solved by a given prover, and "Solved" the number of problems solved out of the total number of 20803 problems. The column "Sum%" denotes the percentage, and "Sum" the total number, of problems solved by the prover or any of the provers listed above it. The column "Unique" denotes the number of problems the given prover solved but no other prover could solve.

We also evaluated various proof reconstruction mechanisms on the problems originating from ATP

Tactic	Time	Solved%	Solved
`yreconstr0`	10s	26.8	1965
`yreconstr`	1s	83.1	6097
`yreconstr`	2s	85.8	6296
`yreconstr`	5s	87.5	6421
`yreconstr`	10s	88.1	6466
`yreconstr`	15s	88.2	6473
`simple`	1s	50.1	3674
`firstorder'`	10s	69.6	5103
`jprover`	10s	56.1	4114
any		90.1	6609

Table 2: Results of the evaluation of proof reconstruction on the 7337 problems solved by the ATPs.

proofs of lemmas in the Coq standard library. In our setting, the Ben-Yelles-type algorithm mentioned in the previous section tends to perform significantly better than the available Coq's tactics. The results of the evaluation are presented in Table 2. Our tactic (`yreconstr`) manages to reconstruct about 88% of the reproved theorems. However, it needs to be remarked that if we use the advice obtained from ATP runs then about 50% of the the reproved theorems follow by a combination of hypothesis simplification, the tactics `intuition`, `auto`, `easy`, `congruence` and a few heuristics (tactic `simple`). Moreover, the `yreconstr` tactic without any hints (`yreconstr0`), i.e., without using any of the information obtained from ATP runs, achieves a success rate of about 26%. The reconstruction success rate of the `firstorder` tactic combined with various heuristics is about 70% if generic axioms for equality are added to the context (tactic `firstorder'`). The `jp` tactic (which integrates the intuitionistic first-order automated theorem prover JProver [28] into Coq) combined with various heuristics and equality axioms (tactic `jprover`) achieves a reconstruction success rate of about 56%. This low success rate is explained by the fact that in contrast to the `firstorder` tactic the `jp` tactic cannot be parameterised by a tactic used at the leaves of the search tree when no logical rule applies.

Acknowledgments. We thank the organizers of the First Coq Coding Sprint, especially Yves Bertot, for the help with implementing Coq export plugins. We wish to thank Thibault Gauthier for the first version of the Coq exported data, as as well as Claudio Sacerdoti-Coen for improvements to the exported data and fruitful discussions on Coq proof reconstruction. This work has been supported by the Austrian Science Fund (FWF) grant P26201.

References

[1] Andreas Abel, Thierry Coquand & Ulf Norell (2005): *Connecting a Logical Framework to a First-Order Logic Prover*. In Bernhard Gramlich, editor: *Frontiers of Combining Systems (FroCoS 2005)*, *LNCS* 3717, Springer, pp. 285–301.

[2] Alejandro Aguirre, Cătălin Hriţcu, Chantal Keller & Nikhil Swamy (2016): *From* F* *to SMT (Extended Abstract)*. In: *Hammers for Type Theories, HaTT 2016*.

[3] Michaël Armand, Germain Faure, Benjamin Grégoire, Chantal Keller, Laurent Théry & Benjamin Werner (2011): *A Modular Integration of SAT/SMT Solvers to Coq through Proof Witnesses*. In Jean-Pierre Jouannaud & Zhong Shao, editors: *Certified Programs and Proofs (CPP 2011)*, *LNCS* 7086, Springer, pp. 135–150.

[4] Andrea Asperti, Wilmer Ricciotti & Claudio Sacerdoti Coen (2014): *Matita Tutorial*. *J. Formalized Reasoning* 7(2), pp. 91–199.

[5] Andrea Asperti & Enrico Tassi (2007): *Higher order Proof Reconstruction from Paramodulation-Based Refutations: The Unit Equality Case*. In Manuel Kauers, Manfred Kerber, Robert Miner & Wolfgang Windsteiger, editors: *Mathematical Knowledge Management (MKM 2007)*, *LNCS* 4573, Springer, pp. 146–160.

[6] Grzegorz Bancerek (2003): *On the structure of Mizar types*. *Electr. Notes Theor. Comput. Sci.* 85(7), pp. 69–85, doi:10.1016/S1571-0661(04)80758-8. Available at `http://dx.doi.org/10.1016/S1571-0661(04)80758-8`.

[7] Grzegorz Bancerek, Czeslaw Bylinski, Adam Grabowski, Artur Kornilowicz, Roman Matuszewski, Adam Naumowicz, Karol Pak & Josef Urban (2015): *Mizar: State-of-the-art and Beyond*. In: *Intelligent Computer Mathematics - International Conference, CICM 2015, Washington, DC, USA, July 13-17, 2015, Proceedings*, pp. 261–279, doi:10.1007/978-3-319-20615-8_17. Available at `http://dx.doi.org/10.1007/978-3-319-20615-8_17`.

[8] Yves Bertot (2008): *A Short Presentation of Coq*. In Otmane Aït Mohamed, César A. Muñoz & Sofiène Tahar, editors: *Theorem Proving in Higher Order Logics (TPHOLs 2008)*, *LNCS* 5170, Springer, pp. 12–16.

[9] Marc Bezem, Dimitri Hendriks & Hans de Nivelle (2002): *Automated Proof Construction in Type Theory Using Resolution*. *J. Autom. Reasoning* 29(3-4), pp. 253–275, doi:10.1023/A:1021939521172. Available at `http://dx.doi.org/10.1023/A:1021939521172`.

[10] Jasmin C. Blanchette, Cezary Kaliszyk, Lawrence C. Paulson & Josef Urban (2016): *Hammering towards QED*. *J. Formalized Reasoning* 9(1), pp. 101–148. Available at `http://jfr.unibo.it/article/view/4593`.

[11] Jasmin Christian Blanchette, David Greenaway, Cezary Kaliszyk, Daniel Kühlwein & Josef Urban (2016): *A Learning-Based Relevance Filter for Isabelle/HOL*. J. Autom. Reasoning, to appear. `http://cl-informatik.uibk.ac.at/cek/mash2.pdf`.

[12] Ana Bove, Peter Dybjer & Ulf Norell (2009): *A Brief Overview of Agda - A Functional Language with Dependent Types*. In Stefan Berghofer, Tobias Nipkow, Christian Urban & Makarius Wenzel, editors: *Theorem Proving in Higher Order Logics (TPHOLs 2009)*, *LNCS* 5674, Springer, pp. 73–78.

[13] Pierre Corbineau (2003): *First-Order Reasoning in the Calculus of Inductive Constructions*. In Stefano Berardi, Mario Coppo & Ferruccio Damiani, editors: *Types for Proofs and Programs (TYPES 2003)*, *LNCS* 3085, Springer, pp. 162–177.

[14] Roy Dyckhoff (1992): *Contraction-Free Sequent Calculi for Intuitionistic Logic*. *J. Symb. Log.* 57(3), pp. 795–807, doi:10.2307/2275431. Available at `http://dx.doi.org/10.2307/2275431`.

[15] Jean-Christophe Filliâtre (2013): *One Logic to Use Them All*. In Maria Paola Bonacina, editor: *International Conference on Automated Deduction (CADE 2013)*, *LNCS* 7898, Springer, pp. 1–20.

[16] Jean-Christophe Filliâtre & Andrei Paskevich (2013): *Why3 - Where Programs Meet Provers*. In Matthias Felleisen & Philippa Gardner, editors: *European Symposium on Programming (ESOP 2013)*, *LNCS* 7792, Springer, pp. 125–128, doi:10.1007/978-3-642-37036-6_8. Available at `http://dx.doi.org/10.1007/978-3-642-37036-6_8`.

[17] Thomas Hales (2013–2014): *Developments in Formal Proofs*. *Séminaire Bourbaki* 1086. abs/1408.6474.

[18] John Harrison (2009): *HOL Light: An Overview*. In Stefan Berghofer, Tobias Nipkow, Christian Urban & Makarius Wenzel, editors: *Theorem Proving in Higher Order Logics (TPHOLs 2009)*, *LNCS* 5674, Springer, pp. 60–66.

[19] Joe Hurd (2003): *First-Order Proof Tactics in Higher-Order Logic Theorem Provers*. In Myla Archer, Ben Di Vito & César Muñoz, editors: *Design and Application of Strategies/Tactics in Higher Order Logics (STRATA 2003)*, *NASA Technical Reports* NASA/CP-2003-212448, pp. 56–68. Available at `http://techreports.larc.nasa.gov/ltrs/PDF/2003/cp/NASA-2003-cp212448.pdf`.

[20] Cezary Kaliszyk & Josef Urban (2014): *Learning-Assisted Automated Reasoning with Flyspeck*. *J. Autom. Reasoning* 53(2), pp. 173–213, doi:10.1007/s10817-014-9303-3. Available at `http://dx.doi.org/10.1007/s10817-014-9303-3`.

[21] Cezary Kaliszyk & Josef Urban (2015): *MizAR 40 for Mizar 40*. *J. Autom. Reasoning* 55(3), pp. 245–256, doi:10.1007/s10817-015-9330-8. Available at `http://dx.doi.org/10.1007/s10817-015-9330-8`.

[22] Laura Kovács & Andrei Voronkov (2013): *First-Order Theorem Proving and Vampire*. In Natasha Sharygina & Helmut Veith, editors: *Computer-Aided Verification (CAV 2013)*, *LNCS* 8044, Springer, pp. 1–35. Available at `http://dx.doi.org/10.1007/978-3-642-39799-8_1`.

[23] Jia Meng & Lawrence C. Paulson (2008): *Translating Higher-Order Clauses to First-Order Clauses*. *Journal of Automated Reasoning* 40(1), pp. 35–60, doi:10.1007/s10817-007-9085-y. Available at `http://dx.doi.org/10.1007/s10817-007-9085-y`.

[24] Leonardo de Moura & Daniel Selsam (2016): *Congruence Closure in Intensional Type Theory*. In: *International Joint Conference on Automated Reasoning, IJCAR 2016*.

[25] Leonardo Mendonça de Moura, Soonho Kong, Jeremy Avigad, Floris van Doorn & Jakob von Raumer (2015): *The Lean Theorem Prover*. In Amy P. Felty & Aart Middeldorp, editors: *International Conference on Automated Deduction (CADE 2015)*, *LNCS* 9195, Springer, pp. 378–388.

[26] Leonardo Mendonça de Moura & Nikolaj Bjørner (2008): *Z3: An Efficient SMT Solver*. In C. R. Ramakrishnan & Jakob Rehof, editors: *TACAS 2008*, *LNCS* 4963, Springer, pp. 337–340.

[27] Lawrence C. Paulson & Jasmin Blanchette (2010): *Three Years of Experience with Sledgehammer, a Practical Link between Automated and Interactive Theorem Provers*. In: *8th IWIL*. Available at `http://www4.in.tum.de/~schulz/PAPERS/STS-IWIL-2010.pdf`.

[28] Stephan Schmitt, Lori Lorigo, Christoph Kreitz & Aleksey Nogin (2001): *JProver : Integrating Connection-Based Theorem Proving into Interactive Proof Assistants*. In Rajeev Goré, Alexander Leitsch & Tobias Nipkow, editors: *Automated Reasoning, First International Joint Conference, IJCAR 2001, Siena, Italy, June 18-23, 2001, Proceedings*, *Lecture Notes in Computer Science* 2083, Springer, pp. 421–426, doi:10.1007/3-540-45744-5_34. Available at `http://dx.doi.org/10.1007/3-540-45744-5_34`.

[29] Aleksy Schubert, Paweł Urzyczyn & Konrad Zdanowski (2015): *On the Mints Hierarchy in First-Order Intuitionistic Logic*. In Andrew M. Pitts, editor: *Foundations of Software Science and Computation Structures (FoSSaCS 2015)*, *Lecture Notes in Computer Science* 9034, Springer, pp. 451–465.

[30] Stephan Schulz (2013): *System Description: E 1.8*. In Kenneth L. McMillan, Aart Middeldorp & Andrei Voronkov, editors: *Logic for Programming, Artificial Intelligence (LPAR 2013)*, *LNCS* 8312, Springer, pp. 735–743, doi:10.1007/978-3-642-45221-5_49. Available at `http://dx.doi.org/10.1007/978-3-642-45221-5_49`.

[31] Konrad Slind & Michael Norrish (2008): *A Brief Overview of HOL4*. In Otmane Ait Mohamed, César Muñoz & Sofiène Tahar, editors: *TPHOLs 2008*, *LNCS* 5170, Springer, pp. 28–32.

[32] Tanel Tammet & Jan M. Smith (1998): *Optimized Encodings of Fragments of Type Theory in First-Order Logic*. *J. Log. Comput.* 8(6), pp. 713–744, doi:10.1093/logcom/8.6.713. Available at `http://dx.doi.org/10.1093/logcom/8.6.713`.

[33] Makarius Wenzel, Lawrence C. Paulson & Tobias Nipkow (2008): *The Isabelle Framework*. In Otmane Aït Mohamed, César A. Muñoz & Sofiène Tahar, editors: *Theorem Proving in Higher Order Logics (TPHOLs 2008)*, *LNCS* 5170, Springer, pp. 33–38. Available at `http://dx.doi.org/10.1007/978-3-540-71067-7_7`.

[34] Freek Wiedijk (2007): *Mizar's Soft Type System*. In: *Theorem Proving in Higher Order Logics, 20th International Conference, TPHOLs 2007, Kaiserslautern, Germany, September 10-13, 2007, Proceedings*, pp. 383–399, doi:10.1007/978-3-540-74591-4_28. Available at `http://dx.doi.org/10.1007/978-3-540-74591-4_28`.