Automatically Deriving Schematic Theorems for Dynamic Contexts

Olivier Savary Bélanger
McGill University
osavary@cs.mcgill.ca

Kaustuv Chaudhuri
INRIA, France
kaustuv.chaudhuri@inria.fr

Abstract
Hypothetical judgments go hand-in-hand with higher-order abstract syntax for meta-theoretic reasoning. Such judgments have two kinds of assumptions: those that are statically known from the specification, and the dynamic assumptions that result from building derivations out of the specification clauses. These dynamic assumptions often have a simple regular structure of repetitions of blocks of related assumptions, with each block generally involving one or several variables and their properties, that are added to the context in a single backchaining step. Reflecting on this regular structure can let us derive a number of structural properties about the elements of the context.

We present an extension of the Abella theorem prover, which is based on a simply typed intuitionistic reasoning logic supporting (co-)inductive definitions and generic quantification. Dynamic contexts are represented in Abella using lists of formulas for the assumptions and quantifier nesting for the variables, together with an inductively defined context relation that specifies their structure. We add a new mechanism for defining particular kinds of regular context relations, called schemas, and tacticals to derive theorems from these schemas as needed. Importantly, our extension leaves the trusted kernel of Abella unchanged. We show that these tacticals can eliminate many commonly encountered kinds of administrative lemmas that would otherwise have to be proven manually, which is a common source of complaints from Abella users.

Categories and Subject Descriptors F.4.1 [Mathematical Logic and Formal Languages]: Mathematical Logic—Computational logic, lambda calculus and related systems, proof theory

General Terms Theory

Keywords dynamic contexts; context relations; context schemas; tactics and tacticals

1. Introduction
Higher-order abstract syntax (HOAS) [13], also known as λ-tree syntax (λTS) [9], is the popular name for a representational scheme where data structures with binding constructs are represented using λ-terms in a logical framework in such a way the binding structure of the λ-terms reflects that of the represented data. In this paper we use the term HOAS in the narrow sense when the logical framework guarantees that all λ-terms are built out of variables, λ-abstractions, and applications, and that the equational theory of λ-terms identifies terms up to αβη-conversion. For example, consider the following signature (in λProlog [10]) specifying a data structure for simply typed λ-terms.

\[ \text{kind } \text{ty } \rightarrow \text{type}. \]
\[ \text{type } \text{i } \rightarrow \text{ty}. \]
\[ \text{type } \text{arr } \text{ty } \rightarrow \text{ty } \rightarrow \text{ty}. \]
\[ \text{kind } \text{tm } \rightarrow \text{type}. \]
\[ \text{type } \text{app } \text{tm } \rightarrow \text{tm } \rightarrow \text{tm}. \]
\[ \text{type } \text{abs } \text{ty } \rightarrow (\text{tm } \rightarrow \text{tm}) \rightarrow \text{tm}. \]

The term \[ \lambda f:i \rightarrow i. \lambda x:i. f(fx) \] would be represented as follows.

\[ \text{abs } (\text{arr } i i) \]
\[ (\lambda f. \text{abs } i (\lambda x. \text{app } f (\text{app } f x))) \]

Reasoning about such representations requires a logic that can support arbitrarily nested implications and universal quantification, such as the logic of higher-order hereditary Harrop formulas (HOHH) [10, sec. 5.2.2] that forms the basis of λProlog and the Abella interactive theorem prover [5, 19]. Such logics are generally presented in terms of sequents (or hypothetical judgments) of the form \[ \Gamma \vdash C \] where \( C \) is a formula and \( \Gamma \) is a context of formulas. As an illustration, suppose we wish to represent the following type-checking judgment that relates a λ-term to its type. In λProlog we write it using a relation \( \Gamma \vdash \ )1 with these program clauses.

\[ \text{type of } \text{tm } \rightarrow \text{ty } \rightarrow \text{o}. \]
\[ \text{of } (\text{app } M N) \text{ B } \equiv \]
\[ \text{of } M (\text{arr } A B), \text{of } N A. \]
\[ \text{of } (\text{abs } A R) (\text{arr } A B) \equiv \]
\[ \text{pl } \lambda x. (\text{of x A } \Rightarrow \text{ of (R x) B}). \]

The universal quantifier \( \text{pl} \) and implication \( \Rightarrow \) occur in the body of the clause for abstractions and determines a scoped assumption for a fresh variable \( x \) that is used to reason about the body of the λ-term \( R \). Note that \( (R x) \) stands for application in λProlog; if \( R \) were \[ \text{abs } i (\lambda x. \text{app } u x) \] for example, then \( (R x) \) would be equal to \( \text{abs } i (\lambda x. \text{app } x x) \), which avoids the capture of \( x \).

Let \( \Phi \) stand for this pair of defining clauses for \( \text{of} \). The typing judgment \[ \lambda f:i \rightarrow i. \lambda x:i. f(fx) : (i \rightarrow i) \rightarrow i \rightarrow i \] amounts to showing that the following sequent is derivable:

\[ \Phi \vdash \text{of } (\text{abs } (\text{arr } i i) (\lambda f. \text{abs } i (\lambda x. \text{app } f (\text{app } f x)))) ((\text{arr } (\text{arr } i i) (\text{arr } i i)). \]

1 Note: all relations have target type \( \text{o} \), the type of λProlog formulas.
To prove this sequent, we would need to backchain the second clause in \( \Phi \), which will produce the subgoal:

\[
\Phi, \text{ of } f (\text{arr } i j) \vdash \text{of} (\text{abs } i (\text{app } f (\text{app } f x))) (\text{arr } i i)
\]

where \( f \) is a fresh variable, i.e., it does not occur free in the original goal. Backchaining once more would produce the subgoal:

\[
\Phi, \text{ of } f (\text{arr } i i), \text{ of } x i \vdash \text{of} (\text{app } f (\text{app } f x)) i
\]

where \( x \) is a fresh variable, i.e., it does not occur in the first subgoal and is different from \( f \).

As should be obvious from this example, every context that occurs in derivations involving the of relation has the structure:

\[
\Phi, \text{ of } x_1 t_1, \ldots, \text{ of } x_n t_n
\]

where for \( i \in 1..n \), the variable \( x_i \) is fresh for \( \Phi \) and for \( (x_j, t_j) \) for every \( j \in 1..i-1 \). The program \( \Phi \) is a static participant in these contexts, while the rest of the context is dynamic, determined entirely by the original goal. In Abella, which can support inductive definitions and generic quantification (using the \( \nabla \) quantifier [11]), the form of this dynamic portion of the context can be specified as an inductively defined predicate \( \text{ctx} \) as follows.

**Define** \( \text{ctx} : \text{olist} \to \text{prop} \) by

\[
\text{ctx nil} ; \text{nabla} x, \text{ctx} (\text{of} x A :: G) \triangleq \text{ctx} G.
\]

The type \( \text{olist} \) denotes a list of HOHH formulas (built as usual using \( \text{nil} \) and \( ; \)) that represents the context, while the type \( \text{prop} \) denotes formulas of the meta-logic. The definition consists of a sequence of clauses separated by semi-colons; each clause contains a head and an optional body (specified using \( \triangleq \)). We follow the \( \lambda \text{Prolog} \) convention of universally closing every clause of the definition over its capitalized free variables. The \( \text{nabla} \) at the head of the second clause of this inductive definition asserts that \( x \) is fresh for—i.e., does not occur free in—\( \lambda \) and \( G \).

Using definitions such as \( \text{ctx} \) in proofs requires a number of essentially administrative inductive theorems for reasoning about the dynamic context (i.e., lists of type \( \text{olist} \)) specified by it. These lemmas amount to unfolding the \( \text{ctx} \) definition to observe the structural properties of its argument, such as that the head of the list is of the form \( f x A \), that \( x \) is a nominal constant, and that it does not occur in \( \lambda \) or in the tail of the list. In Abella 2.0, such theorems have to be proved manually. This is a common source of frustration because of the generally uninteresting nature of these theorems and their proofs. The problem is worse than it appears on the surface because in a large development there may be several relations like of that may be rewritten to each other by \( \alpha \beta \gamma \) and a systematic permutation of their nominal constants are equated:

2 This is related to a similar notion from nominal logic [18], but we retain the HOAS representation of terms.

3 http://abella-prover.org
Abella adopts the Coq style of displaying subgoals: hypotheses and conclusions are separated by a line of ‘=’s.) The inductive hypothesis, called H1, has the same form as the theorem except the assumption that must be strictly smaller is marked with * This assumption in the conclusion is modified to have annotation * . This annotation changes to • after at least one application of a backchaining step, i.e., at least one unfolding of the inductive definition of the H0H1H2H3 sequent calculus. In Abella, this is achieved using the case tactic that considers every possible way to backchain on clauses in the program or in the dynamic context G to derive the conclusion of M A. As expected, there are exactly three possibilities.

P1. Backchaining on the first program clause produces:

Variables: G M A B M1 N1 A1 B1
H1 : ctx G
H3 : (G ⊢ of (app M1 N1) B)
H4 : (G ⊢ of M1 (arr A1 B1))*
H5 : (G ⊢ of N1 A1)*

B1 = B

The eigenvariables M and A are unified with the terms app M1 N1 and B1 respectively.

P2. Backchaining on the second program clauses produces:

Variables: G M A B R A1 B1
H1 : ctx G
H3 : (G ⊢ of (app M1 N1) B)
H4 : (G ⊢ of (arr A1 R) B)*
H5 : member F G

arr A1 B1 = B

Once again, some eigenvariables get unified with other terms. Moreover, the new assumption H4 has a larger dynamic context, with a fresh assumption of n1 A1 where n1 is a nominal constant. Abella uses the convention that all identifiers beginning with n and followed by numbers are nominal constants.

P3. Finally, backchaining on an element of G itself produces:

Variables: G M A B F
H1 : ctx G
H3 : (G ⊢ of M B)
H4 : (G, of [F] ⊢ of (M A))*
H5 : member F G

A = B

The hypothesis H4 stands for the assertion that we are backchaining on F, which must be a member of G by H5.

The proof follows the technique of unfolding H3, and then appealing to the IH on the results. For example, for P1, one of the cases would be:

Variables: G M A B M1 N1 A1 B1 A2 B2
H1 : ctx G
H4 : (G ⊢ of M1 (arr A1 B1))*
H5 : (G ⊢ of N1 A1)*
H6 : (G ⊢ of M1 (arr A2 B2))

N7 : (G ⊢ of H2 A2)

B1 = B2

Here, invoking the IH on H1, H5, H7 will unite A1 and A2, so H1, H4, H6 will then unite arr A1 B1 with arr A1 B2, which will make the conclusion true. The other case is when H3 is itself proved by backchaining on a clause in G:

Variables: G M A B M1 N1 A1 B1
H1 : ctx G
H4 : (G ⊢ of M1 (arr A1 B1))*
H5 : (G ⊢ of N1 A1)*
H6 : (G, of [F] ⊢ of (app M1 N1) B)
H7 : member F G

arr A1 B1 = B

This case is impossible, since (by H1) G contains only assumptions of the form of a C where n is a nominal constant. Nominal constants can only be united with other nominal constants up to equivariance, so (G, of n C) ⊢ of (app M1 N1) B) has no proof, since n and app M1 N1 do not unify. In Abella this can be stated as a lemma.

forall G E, ctx G → member E G → exists X A, (E = of X A) ∧ name X.

where name asserts that its argument is a nominal constant of type tm, definable in Abella as:

Define name : tm → prop by nabla x, name x.

This is an administrative lemma that can almost entirely be derived from the ctx definition.

Another example of an administrative lemma comes from case P3, where from the above lemma we know that F = of n1 C and member (of n1 C) (G n1). Note that G is raised over the new nominal constant n1. This changes H4 to (G n1, [of n1 C] ⊢ of (M n1) (A n1)), which in turn gives the solution M n1 = n1 (i.e., M = λx. x) and A n1 = C (i.e., A = λx. C). Now, if we apply the same reasoning to the hypothesis H3, we would deduce that member (of n1 D) (G n1) and of n1 D. The conclusion will require us to show that C = D. This requires the following lemma.

forall G X A B, ctx G → member (of X A) G → member (of X B) G → A = B.

This is a uniqueness lemma that guarantees that every every variable is assigned a unique type in G by the ctx definition. As before, this lemma has an uninteresting inductive proof that follows almost entirely from the definition of ctx. Indeed, its proof itself uses another administrative lemma asserting that a nominal constant that does not occur in a list cannot occur in any member of the list.

forall E A, nabla (n:tm), member (E n) G → exists F, E = λx. F.

In other words, if E n occurs in G, which cannot depend on n because of the order of forall and nabla, it must be the case that E cannot depend on n either, i.e., E begins with a vacuous λ-abstraction.
In all, the administrative lemmas and their proofs constitute about 60% of the lines of code in this reasoning file. Such lemmas occur repeatedly in the examples suite of Abella, often with slight variations in their formulation and a wide variance in their names. Larger developments contain a number of specified relations such as, e.g., each producing its own ctx definition and their associated administrative lemmas. Indeed, Abella even allows for context relations, which are inductive definitions such as ctx with multiple context arguments, which further causes an exponential proliferation of administrative lemmas. It has been clear for a long time that we require better automation to deal with such lemmas about contexts. Indeed, this is one of the criticisms of Abella in the recent survey of HOAS reasoning systems by Felty et al. [4].

3. A Framework for Plugins

This work proposes to derive a large class of these administrative lemmas automatically when the relevant ctx-like definition has a regular form. We implement this technique in terms of a plugin in an extension of the Abella system with a plugin architecture. As the architecture is rather generic, we describe it before the particular plugin for deriving administrative lemmas.

Abella is written in OCaml and has a broadly LCF-style architecture with a core family of trusted tactics that formalize the inference rules of the logic G [7]. The basic reasoning tactics case (for case-analysis) and search (for depth-bounded automated search) are implemented using these core tactics. However, Abella 2.0 lacks a mechanism for defining new tactics like case and search; users of Abella must write their proofs using the tactics that already exist. This design allows Abella to be compiled—even to act as a compiler itself—but does limit its versatility.

Our approach is to allow users to write Abella plugins that can extend both the grammar of Abella and its family of tactics. However, we do not allow arbitrary extensions of either. We require all extensions to the grammar to be explicitly delimited, and for all top level commands and proof tactics added in the plugins to function as elaborators that produce proof text for the core Abella plugins. This not only makes the plugins certifying, keeping the trusted core of Abella unaltered, but also allows developments built using plugins to be used even in versions of Abella without the plugin architecture.

Each toplevel command or tactic added by a plugin named Plug must have the form

```ocaml
module type PLUGIN = sig
val process_top :
  core:(string -> unit)
  -> string 
  -> unit
end
```

Each module of type PLUGIN has to implement two functions, process_tactic and process_top, defining its behavior on tactics and toplevel commands, respectively. Each function takes a named parameter core, a shallow wrapper around the core Abella functionality which processes the elaborated string produced by the plugin.

To add a new plugin to Abella, it is necessary to add the module implementing PLUGIN to a global plugins table. This table is stored in the file abella.ml that is the entry point for Abella, so every added plugin requires recompiling this file and relinking Abella. For instance, to add the Schemas plugin implemented as the OCaml module Schemas (described in the next section), we add the following line to abella.ml and recompile Abella.

```ocaml
Hashtbl.add plugins "Ctx" (module Schemas : PLUGIN);
```

Note that plugins is a mapping from strings to first-class OCaml modules, which were added in OCaml 3.12 and significantly improved in 4.0. We require plugin names to be valid upper-cased Abella identifiers distinct from all built-in core keywords, and their namespace is flat and global. In future work, we plan to use the dynamic loading features of OCaml 4.0— which is not yet released at the time of writing this paper—to avoid recomilation, and instead have Abella dynamically initialize the table of plugins from a configuration file.

4. Regular Context Relations

The ctx definition of Fig. 1 is a unary context relation. Abella allows definitions of context relations of arbitrary arity, and even relations between contexts and other inductively defined structures such as natural numbers. From this zoo of possibilities, we select a class of regular context relations for which we can automatically derive the administrative lemmas. A regular context relation of arity \( n \geq 1 \):

- is an inductively defined predicate on \( n \) arguments of type \( \text{list} \); and
- relates \( n \text{ nil}'s \) as the base case; and
- each non-base case clause of the predicate completely specifies the heads of all the argument lists and whose bodies are just recursive invocations on the tails of the lists.

The Schemas plugin of Abella adds a new toplevel declaration, Schema, for declaring such regular context relations. This declaration has the following general form

```ocaml
Schema <name> \( \triangleq \) <clause> \( ; \) \( \ldots \) <clause>;
```

where each <clause> has the form:

```ocaml
exists A_1 \ldots A_m, nbla \ x_1 \ldots \ x_n, (F_1, \ldots, F_k)
```

where the \( F_i \) are either arbitrary HOHH formulas built using the variables \( A_1, \ldots, A_m, x_1, \ldots, x_n \) or left blank, indicating that

\[ \exists x_1 \ldots x_n. \]

\[ \forall y_1 \ldots y_m. \]

\[ P(y_1) \]

\[ \forall z_1 \ldots z_k. \]

\[ Q(z_1) \]

\[ \ldots \]

\[ \forall w_1 \ldots w_l. \]

\[ R(w_1) \]

\[ \ldots \]

\[ \forall t_1 \ldots t_p. \]

\[ S(t_1) \]

\[ \ldots \]

\[ \forall u_1 \ldots u_q. \]

\[ T(u_1) \]

\[ \ldots \]

\[ \forall v_1 \ldots v_r. \]

\[ U(v_1) \]

\[ \ldots \]

\[ \forall w_1 \ldots w_s. \]

\[ V(w_1) \]

\[ \ldots \]

\[ \forall x_1 \ldots x_n. \]

\[ \forall y_1 \ldots y_m. \]

\[ P(y_1) \]

\[ \ldots \]

\[ \forall z_1 \ldots z_k. \]

\[ Q(z_1) \]

\[ \ldots \]

\[ \forall w_1 \ldots w_l. \]

\[ R(w_1) \]

\[ \ldots \]

\[ \forall t_1 \ldots t_p. \]

\[ S(t_1) \]

\[ \ldots \]

\[ \forall u_1 \ldots u_q. \]

\[ T(u_1) \]

\[ \ldots \]

\[ \forall v_1 \ldots v_r. \]

\[ U(v_1) \]

\[ \ldots \]

\[ \forall w_1 \ldots w_s. \]

\[ V(w_1) \]

\[ \ldots \]

\[ \forall x_1 \ldots x_n. \]

\[ \forall y_1 \ldots y_m. \]

\[ P(y_1) \]

\[ \ldots \]

\[ \forall z_1 \ldots z_k. \]

\[ Q(z_1) \]

\[ \ldots \]

\[ \forall w_1 \ldots w_l. \]

\[ R(w_1) \]

\[ \ldots \]

\[ \forall t_1 \ldots t_p. \]

\[ S(t_1) \]

\[ \ldots \]

\[ \forall u_1 \ldots u_q. \]

\[ T(u_1) \]

\[ \ldots \]

\[ \forall v_1 \ldots v_r. \]

\[ U(v_1) \]

\[ \ldots \]

\[ \forall w_1 \ldots w_s. \]

\[ V(w_1) \]

\[ \ldots \]

\[ \forall x_1 \ldots x_n. \]

\[ \forall y_1 \ldots y_m. \]

\[ P(y_1) \]

\[ \ldots \]

\[ \forall z_1 \ldots z_k. \]

\[ Q(z_1) \]

\[ \ldots \]

\[ \forall w_1 \ldots w_l. \]

\[ R(w_1) \]

\[ \ldots \]

\[ \forall t_1 \ldots t_p. \]

\[ S(t_1) \]

\[ \ldots \]

\[ \forall u_1 \ldots u_q. \]

\[ T(u_1) \]

\[ \ldots \]

\[ \forall v_1 \ldots v_r. \]

\[ U(v_1) \]

\[ \ldots \]

\[ \forall w_1 \ldots w_s. \]

\[ V(w_1) \]

\[ \ldots \]

\[ \forall x_1 \ldots x_n. \]

\[ \forall y_1 \ldots y_m. \]

\[ P(y_1) \]

\[ \ldots \]

\[ \forall z_1 \ldots z_k. \]

\[ Q(z_1) \]

\[ \ldots \]

\[ \forall w_1 \ldots w_l. \]

\[ R(w_1) \]

\[ \ldots \]

\[ \forall t_1 \ldots t_p. \]

\[ S(t_1) \]

\[ \ldots \]

\[ \forall u_1 \ldots u_q. \]

\[ T(u_1) \]

\[ \ldots \]

\[ \forall v_1 \ldots v_r. \]

\[ U(v_1) \]

\[ \ldots \]

\[ \forall w_1 \ldots w_s. \]

\[ V(w_1) \]

\[ \ldots \]

\[ \forall x_1 \ldots x_n. \]
the clause does not modify this projection of the context relation. The number of \( F_i \) determines the arity of the definition; each clause must specify exactly \( k \) projections for a relation of arity \( k \). Note that the nesting order of \( \text{exists} \) and \( \text{nabla} \) is fixed and guarantees that every \( x \) is fresh for each \( A \).

As a simple example, here is how the \( \text{ctx} \) definition of Fig. 1 can be written as a schema.

\[
\text{Schema} \; \text{ctx} \triangleq \\
\text{exists} \; A, \; \text{nabla} \; x, \; \text{(of } x \; A) .
\]

Using the \text{Ctx} plugin, we would in fact write it as follows:

\[
\text{Ctx}! \; \text{Schema} \; \text{ctx} \triangleq \\
\text{exists} \; A, \; \text{nabla} \; x, \; \text{(of } x \; A) !.
\]

When the \text{Ctx} plugin processes this declaration, it instructs \textit{Abella}'s kernel (using \text{process\_top}, cf. Section 3) to process exactly the inductive definition of \text{ctx} in Fig. 1. The \text{ctx nil nil} case is implicitly added, and is therefore not part of the schema declaration. In the rest of this section, we will elide the \text{Ctx}! delimiters.

A more complex example comes from the \text{normal.thm} file from \textit{Abella}'s example suite that shows how to partition \( \lambda \)-terms into normal and neutral (aka. atomic) forms:

\[
\text{Define} \; \text{ctxs} : \; \text{ol} \rightarrow \text{ol} \rightarrow \text{prop} \; \text{by} \\
\text{ctxs} \; \text{nil nil} ; \\
\text{nabla} \; x, \; \text{ctxs} \; \text{(term } x \; :: \; L) \\
(\text{neutral } x \; :: \; K) \triangleq \text{ctxs} \; L \; K .
\]

Here is how it is depicted as a context relation schema.

\[
\text{Schema} \; \text{ctxs} \triangleq \\
\text{nabla} \; x, \; \text{ctxs} \; \text{(term } x \; :: \; L, \; \text{neutral } x \; :: \; K) .
\]

Note that if any of the \text{exists} or \text{nabla} bound variables list is empty, the corresponding \text{exists} or \text{nabla} prefix may be dropped. The important feature of this schema is that the nominal variable \( x \) is shared between the two contexts in the relation.

For a yet more complex example to illustrate that the formulas at the heads of the lists representing the related contexts need not be atomic, take the \text{ctx2} definition from \text{breduce.thm}\footnote{In: \text{examples/lambda-calculus/term-structure/normal.thm}} [19].

\[
\text{Define} \; \text{ctx2} : \; \text{ol} \rightarrow \text{ol} \rightarrow \text{prop} \; \text{by} \\
\text{ctx2} \; \text{nil nil} ; \\
\text{nabla} \; x, \; \text{ctx2} \; \text{(bred } x \; x \; :: \; G) \\
(\text{path } x \; p \; :: \; D) \triangleq \text{ctx2} \; G \; D .
\]

Here is its depiction as a context relation schema.

\[
\text{Schema} \; \text{ctx2} = \\
\text{nabla} \; x, \; \text{ctx2} \; \text{(bred } x \; x, \; \text{path } x \; p) \\
(\text{exists } \; x, \; \text{nabla} \; x , \\
(\text{pi } \lambda \; u. \; \text{bred } N \; u \Rightarrow \text{bred } x \; u) : \; :: \; G) \\
(\text{pi } \lambda \; q. \; \text{path } N \; q \Rightarrow \text{path } x \; q) : \; :: \; D) \triangleq \text{ctx2} \; G \; D .
\]

The second clause above has three kinds of variables: existential (\( \exists \)), nominal (\( \lambda \)), and bound (\( \forall \) and \( \circ \)). Only the existential and nominal variables can be shared between the related contexts.

For a final example, take the \text{ctx3} definition from \text{cr.thm}\footnote{In: \text{examples/lambda-calculus/cr.thm}}.

\[
\text{Define} \; \text{ctxs} : \; \text{ol} \rightarrow \text{ol} \rightarrow \text{ol} \rightarrow \text{prop} \; \text{by} \\
\text{ctxs} \; \text{nil nil nil} ; \\
\text{nabla} \; x, \; \text{ctxs} \; \text{(tm } x \; :: \; L) \\
(\text{pr1 } x \; x \; :: \; K) \\
(\text{cd1 } x \; x \; :: \; \text{notabs } x \; :: \; J) \triangleq \\
\text{ctxs} \; L \; K \; J .
\]

The third argument of the second clause adds two elements to the head. We use the conjunction operator (\( \& \)) of Prolog in the corresponding schema.

\[
\text{Schema} \; \text{ctxs} = \\
\text{nabla} \; x, \; (\text{tm } x, \; \text{pr1 } x \; x, \; \text{cd1 } x \; x \& \text{notabs } x) .
\]

It should be noted that the support for reasoning about \( \& \) is currently rather primitive in \textit{Abella}. While the above declaration is accepted by the plugin, the automatically derived theorems currently are not accepted by \textit{Abella}. (The generated definition itself is accepted.)

We end this section by noting a number of ways in which regular context relations given as schemas do not capture the full generality of definable relations in \textit{Abella}:

- Multiple schemas may not be mutually recursive.
- Schemas can only relate dynamic contexts (\( \circ \)), not other inductively defined objects such as natural numbers.

None of these restrictions is significant as there is exactly one instance of each kind in the current \textit{Abella} examples. Moreover, removing these restrictions appears to add considerable complications to the automatic derivation of theorems. We therefore leave them to future work.

5. Derived Administrative Lemmas

In this section we inventory the administrative lemmas that are automatically derived from the schema declarations by the \textit{Schemas} plugin. These lemmas are of two basic kinds: those that arise from the types of the existentially and nominally quantified variables in a schema declaration, and those that arise from its logical structure. Lemmas of the first kind are mainly used in the automatically derived proofs of the lemmas of the second kind, but are sometimes also useful in the general toolset.

5.1 Lemmas from Types

Consider again the simple schema below corresponding to the \text{ctx} predicate of Fig. 1.

\[
\text{Schema} \; \text{ctx} \triangleq \\
\text{exists} \; A, \; \text{nabla} \; x, \; (\text{of } x \; A) .
\]

As we already mentioned, from this declaration (and its induced inductive definition), we intend to derive, automatically, that \( x \) is a nominal constant and that \( x \) is fresh for \( A \). In \textit{Abella}, these properties are easily defined, but because \textit{Abella} is not polymorphic, these definitions have to be manually monomorphized to the types in question. For instance, in the above schema, type-inference would derive the fact that \( A \) has type \( \text{ty} \) and \( x \) has type \( \text{tm} \). Thus, we would need the following instances of the \text{name} and \text{fresh} predicate:

\[
\text{Define} \; \text{name\_tm} : \; \text{tm} \rightarrow \text{prop} \; \text{by} \\
\text{nabla} \; x, \; \text{name} \; x .
\]

\[
\text{Define} \; \text{fresh\_tm\_in\_ty} : \; \text{tm} \rightarrow \text{ty} \rightarrow \text{prop} \; \text{by} \\
\text{nabla} \; x, \; \text{fresh} \; x \; \text{A} .
\]

As a side-effect of processing the \text{Schema} declaration above, the \textit{Schemas} plugin automatically adds these definitions. Precisely, a \text{name} predicate is generated for each type of nominally quantified variable in any clause of a schema, and a \text{fresh} predicate for every pair of types of nominal variables and existentially variables in each clause of the schema. Note that these instances apply to basic types declared in the signature, not to arbitrary types; if the schema uses such types, then no such definitions are generated. The \textit{Schemas} plugin keeps track of all such administrative definitions to prevent duplicates, but it may add definitions that are not ever used. This is because \textit{Abella} only allows inductive definitions at the top-level, not during a proof, and no plugin is allowed to "rewind" the state of \textit{Abella} to retroactively add definitions.

For each type of a nominally quantified variable, the \textit{Schemas} plugin also generates a \text{prune} lemma, as explained in Section 2. Here is the version for \( \text{tm} \).
Theorem member_prune_tm :  
\forall E \, G. \, \nabla x. (x : \text{tm}),

member (E x) G \rightarrow \exists F, E = \lambda x. F.

Note that no member_prune_tm is generated as there is never a nominally quantified variable of type \(\text{tm}\) in the schema. The \textit{Schemas} plugin uses process_top and process_tactic to both state and prove the member_prune_tm theorem.

We note here that much of this administrative boilerplate can be removed if \textit{Abella} were to support type-polymorphism. The definitions of \texttt{name} and \texttt{fresh}, and the statement and proof of member_prune, are identical for every type, and ideally should be part of the standard prelude. However, this does not mean that type-based administrative definitions and lemmas are completely worthless. For instance, \textit{Abella} does not currently allow induction on typing itself, which means that induction on the structure of a term must be mediated by a somewhat redundant inductive definition of the structure of well-typed terms. Such definitions and their corresponding lemmas can be automatically derived by the \textit{Schemas} plugin (or by a different specialized plugin) in the future.

5.2 Lemmas from Logical Structure

The remaining administrative lemmas in the \textit{Schemas} tactic come from the logical structure of the \textit{Schema} declaration. These are implemented in the \textit{Schemas} plugin as new \texttt{tacticals} that can be invoked in the process of a proof. Each tactical reflects on the structure of the subgoal being proved and the schema declarations known so far to introduce new assumptions into the context, which are then used using the standard \textit{Abella} tactics to continue the proof.

\textit{Inversion}. The inversion \texttt{tactical} reflects on two assumptions, one of which is a context atom \(\text{ctx} \, G_1 \ldots G_n\), where \(\text{ctx}\) is produced by a \textit{Schema} declaration, and the other is of the form \(\text{member} \, E \, G\), for some \(i \in \{1..n\}\). The result of the tactical is that \(E\) must be one of the formulas that occur in the \(i\)th position in the clauses of the \textit{Schema} declaration. For example, given the schema:

\texttt{Schema ctx_ofev} \(\triangleq\)

\(\exists A, \, \nabla x. (of \, x \, A, \, eval \, x \, \text{x})\)

and a subgoal with:

\(H_1 : \text{ctx_ofev} \, G \, D\)
\(H_2 : \text{member} \, E \, D\)

the tactical \texttt{inversion} \(H_1 \, H_2\) produces the new assumption:

\(H_3 : (\exists A \, X. \, (E = eval \, X \, X) \wedge \text{member} \, (of \, A \, X) \, G \wedge \text{fresh_tm_in_ty} \, X \, A) \)  
\(\vee (\exists A \, V \, X. \, (E = eval \, X \, V) \wedge \text{member} \, (of \, A \, X) \, G \wedge \text{fresh_tm_in_ty} \, X \, A) \wedge \text{fresh_tm_in_tm} \, X \, V)\)

Each disjunct produced by this tactical therefore contains:

- the corresponding \texttt{member}(s) of the other context(s) in the context relation; and
- the necessary assumptions about freshness of the \texttt{fresh}\texttt{val}-quantified variables in the \textit{Schema} declaration corresponding to the clause.

Internally, the \texttt{process_tactic} function of the plugin is used to first assert\(^5\) and prove the general form of an \textit{inversion lemma}; this asserted lemma is then used for the particular hypotheses indicated in the arguments of the tactical. The proof of the assertion is by a

\(^5\)The \texttt{assert} tactic of \textit{Abella} is used to assert and prove a lemma in a subproof and then to continue the proof with the lemma as a hypothesis, \textit{i.e.}, it is an instance of the cut rule of the \textit{G} logic.
Schema $\text{ctx} \triangleq \exists x \ A$, nabla $x$, (of $x \ A$).

In this case, if we are in a subgoal with:

- $H_1 : \text{ctx} \ G$
- $H_2 : \text{member} \ (\text{of} \ X \ A) \ G$
- $H_3 : \text{member} \ (\text{of} \ X \ B) \ G$

then it must be that $A$ and $B$ are equal, since there is only one clause of $\text{ctx}$ that could have introduced any member of the form of $X \ C$ into $G$. This is achieved by the tactical application unique $H_1 H_2 H_3$, which has the side effect of uniting the terms $A$ and $B$.

While easily explained, this tactic has several subtleties. First, we require that the contexts—the $G$ above—be identical in all three arguments to unique, and that each member—the of $X \ A$ and of $X \ B$ above—be unifiable with one of the formulas in the contexts related in the Schema declaration. If the latter assumption is not true, then we can just use inversion to rule out this entire subgoal as impossible. Second, we do not require the term corresponding to the nabla-quantified variables—the $x$ above—to be a nominal constant; if it is not a nominal constant, then the inversion lemma rules out the subgoal as impossible. Finally, the generated lemma and its proof requires the use of the member prune lemma explained in the previous section: in the inductive argument, the case where one of the members is the first element of the context while the other member is not is impossible, and member prune very succinctly rules it out.

Projection. It is a common design pattern in Abella to prove inductive theorems for the smallest context relations that suffice. Thus, theorems about typing using a specified relation $\sigma$ would use a unary context schema about $\sigma$, while those about evaluation using eval would use a unary schema for eval. However, if a theorem has to relate typing to evaluation, such as in proofs of type-preservation, then it is necessary to state the theorem using a binary schema relating the two contexts. Unfortunately, in Abella there is no automatic way to “import” a theorem proved using a unary context relation into one with a binary relation, nor “export” theorems the other way. Such facts must be proved by hand.

A common denominator of such facts is that there exist mappings between two context relations that existentially close over the contexts in the target of the mapping that are not present in the source. We call such mappings projections. The projas tactical applies to an assumption:

$H_1 : \text{rel1} \ G_1 \ldots \ G_n$

where $\text{rel1}$ is a schematic context relation. The tactical application

projas (rel2 $D_1 \ldots D_m$) $H_1$

where each $D_i$ is either one of the $G_i$ or is a new eigenvariable, has the effect of adding the assumption

$H_2 : \text{rel2} \ D_1 \ldots D_m$

to the goal, when justified.

This tactical application is interpreted into a general projection lemma that has the following form. Let $D_{\sigma(1)}, \ldots, D_{\sigma(k)}$ be the eigenvariables that are distinct from all the $G_i$. Then, the following lemma is proved by induction:

forall $G_1 \ldots G_n$, rel1 $G_1 \ldots G_n$ exists $D_{\sigma(1)} \ldots D_{\sigma(k)}$, rel2 $D_1 \ldots D_m$.

This proof proceeds by induction on the definition of $\text{rel1}$, but is rather straightforward. Of course, if all the $D_i$ are distinct from the $G_i$, then this tactic is useless. Like the other tactics, projas detects invocations which are invalid or outside its fragment and only generate proofs which will be accepted by the Abella kernel.

To illustrate one of the limitations to its fragment, consider the following pair of schemas:

Schema $\text{rel1} \triangleq (\ell, \ell)$.
Schema $\text{rel2} \triangleq (\ell, \ell) : (\ell, \ell)$.

where $\ell$ is an atomic $\text{HOH}$ formula of type $\ell$. Clearly, forall $G$, rel1 $G$ $\rightarrow$ rel2 $G$.

is provable. However, as no single clause of rel2 matches the non-trivial clause of rel1, projas would not apply to this theorem.

6. Experimental Evaluation

We based our implementation of the plugin architecture on Abella version 2.0.1. Our initial experiments are promising. For instance, using the Schemas plugin to rewrite the breduce example from [19] removes over 40% of the lines of code from the file breduce.thm. Table 1 contains a summary of improvements in a few other examples from the Abella examples suite. In addition to this quantitative reduction in size, we can also compare the plugin qualitatively: the Schemas tacticals free us from the tedium of writing and proving the administrative lemmas that make Abella developments both tedious to write and hard to read. Our experience using the plugin has been entirely positive, so we plan to integrate the plugin architecture into the next release (2.1.x) of Abella.

<table>
<thead>
<tr>
<th>File</th>
<th># schemas</th>
<th># lemmas derived</th>
<th>LOC removed</th>
<th>% removed</th>
</tr>
</thead>
<tbody>
<tr>
<td>breduce</td>
<td>3</td>
<td>11</td>
<td>124</td>
<td>42</td>
</tr>
<tr>
<td>copy</td>
<td>1</td>
<td>4</td>
<td>28</td>
<td>43</td>
</tr>
<tr>
<td>cr</td>
<td>1</td>
<td>3</td>
<td>32</td>
<td>19</td>
</tr>
<tr>
<td>type_uniq</td>
<td>1</td>
<td>3</td>
<td>27</td>
<td>63</td>
</tr>
</tbody>
</table>

Table 1. Quantitative evaluation of the Schemas plugin on some examples from the Abella examples suite

7. Related Work

The concept of regular context relations, at least in the unary case, is similar to that of regular worlds from Twelf, introduced in version 1.4 [14, Section 9.1]. A regular world is an arbitrary repetition of a sequence of blocks, which are individually named in Twelf and correspond to the clauses of our Schema declarations. Despite the superficial similarity of Twelf’s block and world declarations and our Schema declarations, there are some significant differences: first, we use nominal abstraction (“nabla at the head”) [7] to interpret our nabla-quantified variables, rather than universal quantification as in Twelf, which allows us to directly use the logical principles of G to derive pruning, inversion, and uniqueness theorems; second, regular worlds in Twelf are tied to a particular inductive type family and cannot be reused as such for different families, nor can a family have different regular world declarations; third, the regularity is at the level of local extensions to the dynamic context rather than to the entire dynamic context as a whole; and finally, because Twelf contexts contain both variable declarations and ordinary assumptions, the rigid list structure of regular worlds forces the use of somewhat unnatural placement of quantifiers in the specification, explained in [14, p. 49]. Twelf also has a concept of world-checking, where the constructors of an inductive type family in a signature are automatically checked (using a trusted checker) to conform to the declared world for that family. This feature is sometimes useful as a sanity check on specifications, but is ultimately orthogonal to formal (logical) reasoning about the specifications.

Regular contexts are given a more principled foundational treatment in the Beluga system [16], which is a dependently typed programming language for reasoning about contextual modal LF terms [12]. Indeed, we appropriated the term “schema” from Beluga. Schemas in Beluga, like regular worlds in Twelf, are treated as

\footnote{See: http://abella-prover.org/schemas}
classifiers of individual contexts, which make them similar to unary context relation schemas in our plugin. Schemas are tightly integrated into the Beluga type system, and it does not make much sense to ask for its treatment of schemas to be certifying with respect to a system without schemas. However, this does raise the level of trust required in the Beluga implementation. For instance, administrative lemmas such as uniqueness and inversion are unnecessary in Beluga as they are built into the type-checker, which is therefore necessarily more complex than the rather straightforward implementation of the core tactics of Abella, which are themselves direct implementations of the \( \mathcal{G} \) inference rules [6, 7].

Our plugin architecture is a restricted form of tacticals—functions on tactics—initially designed in the LCF family of theorem provers (such as HOL) but now pervasive in Coq, Isabelle, NuPRL, etc. There is a particular similarity to tactics languages such as Lttac [1] of Coq that allow building tactics libraries that can reflect on the state of the prover and construct proofs from meta-procedures, which are then checked by the Coq kernel. However, Coq provides no support for reasoning about higher-order abstract syntax, which is our main interest. The Hybrid system [2] is one approach for reasoning about HOAS in Coq by using an intermediate De Bruijn representation; indeed, several of the tactics of the Schemas plugin are reminiscent of similar operations in Hybrid, but a formal correspondence seems difficult given the difference in nature between Hybrid and Abella.

Meta-theorems for reasoning about HOAS-specified object logics has been the topic of a recent survey article sequence [3, 4] that both identifies a family of essential theorems and compares the performance of Twelf, Beluga, Abella, and Hybrid. We believe most of the lemmas and theorems in the associated ORBI library can be automatically derived from a simple composition of the tacticals in the Schemas plugin. For instance, admissibility of reflexivity [4, Theorems 7, 14, 21], context inversion [4, Lemma 9], and completeness [4, Lemma 22] follow immediately from our sync tactic while relational strengthening [4, Theorems 15, 20] correspond to proj}s. Other theorems such as context membership [4, Lemma 6] require both inversion and unique, while transitivity [4, Theorem 10] requires a sequence of applications of inversion. While our plugin does not change the logic of Abella, it requires much of the tediousness of an explicit representation of contexts that was criticized in this survey.

8. Conclusion and Perspectives

We have described an extension to Abella with a backwards-compatible and certifying plugin architecture, which we have used to implement regular context relations in a Schemas plugin, and have given a preliminary experimental evaluation using existing examples from the Abella examples suite.

The main missing feature in this plugin is the ability to reason about context strengthening using subordination, which is a built-in (trusted) feature of both Twelf and Beluga. Since Abella relies strongly on its logical foundations in the \( \mathcal{G} \) logic, the first step would be to give a logical characterization of strengthening and subordination, which is currently an open problem. To an extent strengthening and subordination are not strictly necessary in Abella since we can tailor the context relations to fit the theorems, instead of using a common global context for all theorems. Nevertheless, there are instances in the Abella examples suite where, for instance, one needs to show that the addition of natural numbers remains commutative even when there are assumptions about \( \lambda \)-terms in the dynamic context. Such lemmas are an easy consequence of subordination—the type of \( \lambda \)-terms is not subordinate to that of numbers—but still currently require manual proofs.

Acknowledgements. This work was partially supported by the INRIA Associated Team grant RAPT and by the ERC Advanced Grant ProofCert.

References


