Engineering Formal Metatheory

Brian Aydemir
University of Pennsylvania

Arthur Charguéraud
INRIA

Benjamin C. Pierce
University of Pennsylvania, bcpierce@cis.upenn.edu

Randy Pollack
University of Edinburgh

Stephanie Weirich
University of Pennsylvania, sweirich@cis.upenn.edu

Follow this and additional works at: http://repository.upenn.edu/cis_papers

Recommended Citation


This paper is posted at Scholarly Commons. http://repository.upenn.edu/cis_papers/369
For more information, please contact repository@pobox.upenn.edu.
Engineering Formal Metatheory

Abstract
Machine-checked proofs of properties of programming languages have become a critical need, both for increased confidence in large and complex designs and as a foundation for technologies such as proof-carrying code. However, constructing these proofs remains a black art, involving many choices in the formulation of definitions and theorems that make a huge cumulative difference in the difficulty of carrying out large formal developments. The representation and manipulation of terms with variable binding is a key issue. We propose a novel style for formalizing metatheory, combining locally nameless representation of terms and cofinite quantification of free variable names in inductive definitions of relations on terms (typing, reduction, ...). The key technical insight is that our use of cofinite quantification obviates the need for reasoning about equivariance (the fact that free names can be renamed in derivations); in particular, the structural induction principles of relations defined using cofinite quantification are strong enough for metatheoretic reasoning, and need not be explicitly strengthened. Strong inversion principles follow (automatically, in Coq) from the induction principles. Although many of the underlying ingredients of our technique have been used before, their combination here yields a significant improvement over other methodologies using first-order representations, leading to developments that are faithful to informal practice, yet require no external tool support and little infrastructure within the proof assistant. We have carried out several large developments in this style using the Coq proof assistant and have made them publicly available. Our developments include type soundness for System F and core ML (with references, exceptions, datatypes, recursion, and patterns) and subject reduction for the Calculus of Constructions. Not only do these developments demonstrate the comprehensiveness of our approach; they have also been optimized for clarity and robustness, making them good templates for future extension.

Keywords
binding, Coq, locally nameless

Comments

This conference paper is available at ScholarlyCommons: http://repository.upenn.edu/cis_papers/369
Engineering Formal Metatheory

Brian Aydemir  
University of Pennsylvania  
baydemir@cis.upenn.edu

Arthur Charguéraud  
INRIA  
arthur.chargueraud@inria.fr

Benjamin C. Pierce  
University of Pennsylvania  
bcpierce@cis.upenn.edu

Randy Pollack  
University of Edinburgh  
rpollack@inf.ed.ac.uk

Stephanie Weirich  
University of Pennsylvania  
sweirich@cis.upenn.edu

Abstract

Machine-checked proofs of properties of programming languages have become a critical need, both for increased confidence in large and complex designs and as a foundation for technologies such as proof-carrying code. However, constructing these proofs remains a black art, involving many choices in the formulation of definitions and theorems that make a huge cumulative difference in the difficulty of carrying out large formal developments. The representation and manipulation of terms with variable binding is a key issue.

We propose a novel style for formalizing metatheory, combining locally nameless representation of terms and cofinite quantification of free variable names in inductive definitions of relations on terms (typing, reduction, . . . ). The key technical insight is that our use of cofinite quantification obviates the need for reasoning about equivariance (the fact that free names can be renamed in derivations); in particular, the structural induction principles of relations defined using cofinite quantification are strong enough for metatheoretic reasoning, and need not be explicitly strengthened. Strong inversion principles follow (automatically, in Coq) from the induction principles. Although many of the underlying ingredients of our technique have been used before, their combination here yields a significant improvement over other methodologies using first-order representations, leading to developments that are faithful to informal practice, yet require no external tool support and little infrastructure within the proof assistant.

We have carried out several large developments in this style using the Coq proof assistant and have made them publicly available. Our developments include type soundness for System F₀ and core ML (with references, exceptions, datatypes, recursion, and patterns) and subject reduction for the Calculus of Constructions. Not only do these developments demonstrate the comprehensiveness of our approach; they have also been optimized for clarity and robustness, making them good templates for future extension.

Categories and Subject Descriptors   F.4.1 [Mathematical Logic and Formal Languages]: Mathematical Logic—Mechanical theorem proving;  D.3.1 [Programming Languages]: Formal Definitions and Theory—Syntax

General Terms   Design, Documentation, Languages, Theory, Verification

Keywords   binding, Coq, locally nameless

1. Introduction

Recent years have seen burgeoning interest in the use of proof assistants for formalizing definitions of programming languages and checking proofs of their properties. However, despite several successful tours de force (Appel 2001; Crary 2003; Klein and Nipkow 2006; Leroy 2006; Lee et al. 2007, etc.), the community remains fragmented, with little synergy between groups and, for newcomers wanting to join the game, a perplexing array of choices between different logics, proof assistants, and representation techniques.

To stimulate progress on this problem, Aydemir et al. (2005) proposed the P0PL’08 challenge, a set of tasks designed to stress many of the critical issues in formalizing programming language metatheory. They laid out three criteria for evaluating proposed formalization techniques: infrastructure overhead—formalization overheads, such as additional operations and their associated proof obligations, should not be prohibitive for large developments; transparency—formal definitions and theorems should not depart radically from the usual informal conventions familiar to a technical audience; and cost of entry—the infrastructure should be usable by someone who is not an expert in theorem proving technology.

In this paper, we propose a new style for formalizing programming language metatheory that performs well on all these criteria. The style builds upon two key ingredients: locally nameless representation of syntax involving binders and a cofinite style of quantification for introducing free names in rules dealing with binders.

In locally nameless representation, bound variables are represented by de Bruijn indices while free variables are represented by names. This mixed representation combines the benefits of both approaches, avoiding difficulties associated with alpha-conversion by ensuring that each alpha-equivalence class of terms has a unique representation, while supporting formal reasoning that closely follows informal practice.

The second ingredient is the use of a cofinite quantification of free names introduced by rules dealing with binders, instead of the standard “exists-fresh” quantification—these styles are illustrated by the following variants of the typing rule for abstraction in the
Section 2. A Survey of Binding Representations

There are two main categories of approaches to representing and reasoning about syntax with binding, depending on whether variables and binders are represented as concrete nodes in first-order algebraic structures or whether these aspects are “lifted to the metalinguage” by representing the bodies of binding constructs as metalinguage functions. In each category there are many ingenious variations, and we do not attempt to be exhaustive in this survey. Outside the main categories, there is a plethora of hybrid representations, multi-level representations, …, which we do not consider.

2.1 Concrete Approaches

With concrete (or first-order) approaches, variables are typically encoded using names or natural numbers. Capture-avoiding substitution must then be defined explicitly as a function on terms, and in the case where bound variables are named, alpha-equivalence must also be defined explicitly. Concrete approaches can be subdivided roughly in three categories: those using names to represent variables, those using de Bruijn indices, and those distinguishing bound and free variables (which may use either names or indices for each kind of variable).

The most standard representation on paper uses names to represent variables, generally quotienting raw expressions by some notion of alpha-equivalence. Although used since the first days of the lambda calculus, this representation has not proved convenient for formal developments. For example, naive substitution is not structurally recursive in this setting but requires well-founded recursion. Nevertheless, the Church-Rosser property in pure lambda calculus has been proved using named representations by Homeier (2001) using HOL, Ford and Mason (2001) using PVS, and Vestergaard and Brotherston (2003) using Isabelle/HOL (unusually, Vestergaard and Brotherston reason about unquotiented raw terms). Other attempts to tame this representation include work by Stoughton (1988), where simultaneous substitution is defined structurally and puts terms into a canonical alpha-normal form, and by Hendriks and van Oostrom (2003).

To overcome the clumsiness of explicit alpha-conversion or quotienting, the first large formalizations of languages with binders were based on de Bruijn indices (1972) rather than names. In this approach, each alpha-equivalence class of lambda-terms has a unique syntactic representation in which variables are encoded using natural numbers giving the depth of the variable relative to its binder. Indices greater than the number of enclosing local binders may indicate a position in some enclosing context such as a typing context. There is abundant evidence of the effectiveness of this representation, from proofs of the Church-Rosser property for lambda calculus—for example, by Shankar (1988) using the Boyer-Moore prover, by Huet (1994) in Coq, by Rasmussen (1995) in Isabelle/ZF, and by Nipkow (2001) in Isabelle/HOL—to harder results such as strong normalization for System F in Lego (Altenkirch 1993) and formalizing Coq in Coq (Barras and Werner 1997).

In de Bruijn representation, the treatment of bound variables incurs minor technical annoyances (lifting over binders, etc.), while the treatment of free variables (e.g., variables bound by a typing context) requires reasoning far from natural informal presentations. For example, consider the statement of a lemma that says thinning the typing
Nominal logic (Pitts 2003) provides another way to address the problem of alpha-conversion inherent in the named representation. Therefore, Urbán (2007) and Urban et al. (2007,a,b) have adapted these ideas to standard higher-order logic, and are developing a nominal package in Isabelle/HOL that provides facilities for formal reasoning with binders over names. The package helps by automating language definitions that equate alpha-equivalent terms, by providing lemmas about renaming and freshness of names, by deriving strengthened induction principles that provide name freshness facts for relations over these languages, and by supporting recursion over these languages. Some syntactic restrictions must be observed to use this package. See Section 4.6 for further discussion of the nominal package.

Besides these “homogeneous” concrete approaches—where there is a single notion of “variable,” either a name or a de Bruijn index—there is a third category of concrete representations that uses two distinct syntactic classes, called variables (for locally bound variables) and parameters (for free, or globally bound variables). The idea that variables and parameters should be distinguished is a natural one, but reasoning about alpha-conversion is more complex. Several combinations can be considered, using names, indices, or levels to represent bound and free variables; of these, two have been used in the literature.

The first is the locally named representation, which uses distinct species of names to represent variables and parameters respectively. McKinna and Pollack introduced this technique to formalize Pure Type Systems (PTS) (1993; 1999), following a suggestion by Coquand (1991). This representation avoids the difficulties of reasoning about capture-avoiding substitutions: since variables and parameters are syntactically distinguished, no parameter can ever be captured by a variable binder during substitution. McKinna and Pollack also introduced a technique for handling the requirement of choosing fresh global variables that often occurs in reasoning about binding (weak freshness lemmas are a prototypical example of the problem). By careful choice of definitions, they avoid the need to reason about alpha-equivalence of bound variables throughout their large formal body of PTS metatheory. With these techniques, reasoning about lambda calculus and type theory is straightforward, if heavy.

Nonetheless, the use of names for bound variables is not a perfect fit to the intuitive notion of binding: eventually one needs to reason about alpha-conversion. For example, parallel reduction has the diamond property concretely in locally named representation, but beta reduction has this property only up to alpha-conversion (Pollack 1994b, Section 3.3.6). So Pollack (1994a) suggested that the McKinna–Pollack approach to reasoning with two species of variables also works well with a representation that uses names for parameters and de Bruijn indices for bound variables. This locally nameless representation, in which alpha-equivalence classes have canonical representation, was already mentioned in the conclusion of de Bruijn’s famous paper (1972). It had been used for implementation in Huet’s Constructive Engine (1989), and later in the implementations of the Coq, LEGO (Luo and Pollack 1992), HOL 4 (Norrish and Slind 2007), Isabelle, and EPIGRAM (McBride and McKinna 2004) proof assistants.

In the context of formal proofs, Gordon (1994) appears to be the first to have used locally nameless representation. Rather than reason directly with locally nameless terms, he builds a representation of named lambda terms on top of locally nameless terms and demonstrates how the named terms may be used as a basis for representing syntax. Using this technique, Gordon formalized Abramsky’s lazy lambda calculus. Later work by Gordon and Melham (1996) also uses locally nameless terms as a model for an abstract “axiomatic” representation of named terms. Pollack (2006) has more recently emphasized the benefits of locally nameless representation in the context of the PoplMark challenge. Locally nameless representation with (variants of) McKinna–Pollack style reasoning has been used by several researchers (with several proof tools) for solutions to the PoplMark challenge, first by Leroy (2007) and later by Chipala (2006) and Ricciotti (2007).

2.2 Higher-Order Approaches

Higher-order representations use the function space of the meta-logic to encode binding in an object language, allowing issues like capture-avoidance and alpha-equivalence to be handled once and for all by the meta-logic, rather than facing them anew for each object language. There is a bewildering variety of higher-order approaches, which we survey only superficially.

In higher-order abstract syntax (HOAS) (Harper et al. 1993; Pfenning and Elliot 1988), the introduction form for lambda-abstractions has type \( \text{term} \rightarrow \text{term} \). The lambda constructor packages a function of type \( \text{term} \rightarrow \text{term} \), which should be thought of as the function substituting its argument into the body of the lambda. Not only are alpha-equivalent terms equal, but substitution is handled by the meta-language.

HOAS was introduced by Church and developed by de Bruijn, Martin-Löf, Plotkin, Pfenning, and their co-workers. The modern form, using dependent types and hypothetical and schematic judgements as well as HOAS term representation, appeared in the Edinburgh Logical Framework (LF) (Harper et al. 1993). By design, LF is a weak type theory to be used as a metalanguage, not supporting inductive definition or primitive recursion. It allows faithful representation of object languages, where LF itself adds no mathematical strength to the object language representation. The use of LF methodology for metatheory (as opposed to just representation) has been highly developed by Pfenning and his co-workers; the implementation of this approach is the Twelf system (Pfenning and Schürmann 1999), which is widely used and very successful for formalizing the metatheory of a wide variety of programming languages (Ashley-Rollman et al. 2005; Lee et al. 2007). The approach continues to be developed foundationally as well as in practice (Harper and Licata 2007).

A second mainstream of work, weak higher-order encodings, gives a type such as \( \text{name} \rightarrow \text{term} \rightarrow \text{term} \) to the lambda constructor of terms. This constructor packages a function of type \( \text{name} \rightarrow \text{term} \), which should be thought of as a function that takes a name and returns the instance of the lambda’s body instantiated with that name. Here, alpha-equivalent terms are equated, but substitution must be defined as a relation between terms. Despeyroux et al. (1995) demonstrate an early use of this approach. The approach has been most developed by Honsell, Miculan, and their co-workers as the Theory of Contexts (Honsell et al. 2002; Bucalo et al. 2006).

2.3 Why Locally Nameless?

We have chosen to concentrate on concrete approaches to represent binding structure. Despite the attractiveness of higher-order approaches and their demonstrated success in many large developments, they are, at present, not completely general-purpose. Twelf is currently the only industrial-strength proof assistant for developing
3. Locally Nameless Representation

In this section, we describe locally nameless representation, using $\lambda$- as a running example. All the material presented is formal, in the sense that we have implemented it in Coq.\footnote{Experts will note that this definition of open works correctly only when zero is the only unbound index. The definition can be generalized to allow multiple binders to be opened simultaneously as explained in Section 5 in the paragraph about our ML development.} We write $\Rightarrow$ for logical implication.

3.1 Definitions

The fundamental idea of the locally nameless approach is to distinguish bound from free variables, using de Bruijn indices for the former and names for the latter. Figure 1 presents a locally nameless specification of $\lambda$. An occurrence of a bound variable is represented as (bvar $i$) where $i$ is a natural number index denoting the number of enclosing lambda abstractions that must be traversed before reaching the abstraction that binds variable. Abstractions do not bind names; for example, (abs (bvar 0)) is the identity function.

In its raw form, this representation is not isomorphic to ordinary named lambda terms because a bound variable may not resolve to any binder. For example, (abs (bvar 2)) does not represent any lambda-term. We therefore define a predicate, (term $t$), on pre-terms (the syntactic objects defined by the grammar in Figure 1) that selects the locally closed pre-terms—or just terms—in which all indices resolve to binders.

The key operation on pre-terms is opening an abstraction by instantiating a bound variable with a term. If (abs $t$) is an abstraction and $u$ another term, then the result of opening the abstraction with $u$, written $t^u$, is formed by replacing in the body $b$ all bound variables that point to the outermost lambda with $u$. Figure 1 gives the structurally recursive definition of the open function; the idea is to keep the term, keeping track of the current number of binders passed through.\footnote{The corresponding development is contained in files STLC_Core_*_.v.} One use of the open definition is in the conclusion of the beta reduction rule, RED-BETA, near the bottom of the figure.

The open operation is also frequently applied to a free variable; we call this specialized version variable opening, which we write with abuse of notation as ($t^\text{fvar}$) instead of ($t^{\text{bvar} i}$). Variable opening is used when passing through a binder to turn the bound variable into a free variable. In a named representation, we may have our hands on an abstraction (abs $x$ $t$) and wish to talk about its body $t$. With the locally nameless representation, the abstraction would

Syntax:

\[
\begin{align*}
S, T & \equiv A | T_1 \to T_2 \\
t, u, w & \equiv \text{bvar } i | \text{fvar } x | \text{app } t_1 t_2 | \text{abs } t \\
E, F, G & \equiv \emptyset | E, x : T
\end{align*}
\]

Open: $t^u \equiv \{0 \to u\} t$, with

\[
\begin{align*}
\{k \to u\} (\text{bvar } k) & = u \\
\{k \to u\} (\text{fvar } i) & = \{fvar \ i\} \text{ when } i \not= k \\
\{k \to u\} (\text{app } t_1 t_2) & = \text{app } (\{k \to u\} t_1) (\{k \to u\} t_2) \\
\{k \to u\} (\text{abs } t) & = \text{abs } ((\{k + 1\} \to u) t)
\end{align*}
\]

Free variables:

\[
\begin{align*}
\text{FV}(\text{fvar } i) & = \emptyset \\
\text{FV}(\text{fvar } x) & = \{x\} \\
\text{FV}(\text{app } t_1 t_2) & = \text{FV}(t_1) \cup \text{FV}(t_2) \\
\text{FV}(\text{abs } t) & = \text{FV}(t)
\end{align*}
\]

Locally closed terms:

\[
\begin{align*}
\text{TERM-ABS} & \quad \text{TERM-APP} \\
\text{TERM-VAR}
\end{align*}
\]

Well-formed environments (no duplicate names):

\[
\begin{align*}
\text{OK-NIL} & \quad \text{OK-CONS}
\end{align*}
\]

Typing:

\[
\begin{align*}
\text{OK-CONS} & \quad \text{OK-NIL} \\
\text{OK-CONS} & \quad \text{OK-NIL} \\
\end{align*}
\]

Call-by-value evaluation:

\[
\begin{align*}
& \quad \text{RED-APP-2} \\
& \quad \text{RED-APP-1} \\
& \quad \text{RED-BETA} \\
& \quad \text{TERM-ABS} \\
& \quad \text{TERM-VAR} \\
& \quad \text{TERM-APP}
\end{align*}
\]

Type soundness lemmas (preservation and progress):

\[
\begin{align*}
E \vdash t : T & \Rightarrow t \Rightarrow t' \Rightarrow E \vdash t' : T \\
\emptyset \vdash t : T & \Rightarrow (\text{value } t \vee \exists t', t \Rightarrow t')
\end{align*}
\]

Figure 1. Locally nameless presentation of $\lambda$. 
be written (abs t); to talk about its body as a term (rather than a
pre-term with an unbound index), we open t with some fresh name
x. A typical example of variable opening appears in the typing rule
for abstractions, TYPING-ABS.

In variable opening, the chosen variable should not already
appear free inside the term. Otherwise, the operation would be a
form of variable capture: the name being given to the index, a
bound variable, would be the same as a free variable. Therefore,
the rule TYPING-ABS also includes the premise x \notin FV(t),\footnote{The rule TYPING-ABS could also explicitly include a
premise to ensure that x does not appear in the domain of the
environment E. However, that restriction is already implied by
the second hypothesis.} which uses the free variable function to
collect the set of free variables in a pre-term. Like open, this
operation is defined by simple structural recursion and need not
worry about confusing free and bound variables—all variable
names in the term are free, since bound variables are represented by
indices.

To state the typing judgement of \( \lambda \_ \_ \_), we need to formalize
typing environments. In Figure 1, environments are concretely
represented as association lists, for which we assume some stan-
dard operations. The concatenation of two environments is written
\((E, F)\). The lookup of a variable in an environment is expressed by
the predicate \((x : T) \in E\), which holds when \((x, T)\) is the rightmost
pair of \(E\) whose first component is \(x\). Finally, a variable is fresh
for an environment, written \(x \notin \text{dom}(E)\), when that variable
does not appear in the first component of any pair in \(E\). In principle,
the association list representation allows multiple assumptions for
the same variable, but our typing rules ensure that environments
participating in derivable judgements are well formed (ok) in the
sense that a given variable is bound at most once in a given environ-
ment.

The final portion of Figure 1 completes the definition of \( \lambda \_ \_ \_\)_
by stating the properties we wish to prove: the standard preservation
and progress lemmas. The figure, as a whole, forms our “trusted
base” of definitions: one must believe in the correctness of every-
thing stated in the figure in order to believe that a formal proof of
type soundness is really about one’s informal notion of \( \lambda \_ \_ \_\)_.
We discuss this issue in additional detail in Section 3.4.

3.2 Substitution

The definition of \( \lambda \_ \_ \_\)_ depends only on the open and FV
operations. In proofs, however, the operation of free variable substi-
tution, replacing free names with terms, is also needed. Written \([x \rightarrow u] t\),
it is defined by structural recursion in Figure 2. As in the definition
of open, there is no arbitrary choice of name for the bound variable
in the abstraction case, so we need not worry that this choice will
affect the result. Therefore, the behavior of substitution is natural
and easy to reason about.

The properties of the substitution function, shown in the bottom
portion of Figure 2, are fundamental when working with a locally
nameless representation. Lemma subst_fresh states that substitu-
tion for an unused variable has no effect, and lemma subst_open
states that substitution distributes over the opening operation. The
remaining two lemmas are immediate corollaries of the former
two. Lemma subst_open_var permutes a substitution with variable
opening, and lemma subst_intro decomposes an open operation
into a variable opening operation and a substitution. Examples of
when these lemmas are used can be found in the proof of the substi-
tution lemma for \( \lambda \_ \_ \_\_) (Section 4.1) and in the proof of a renaming
lemma (Section 4.2).

We occasionally need a related operation—intuitively, a recip-
rcal of variable opening—that abstracts the name \(x\) from the term
t, replacing it with a bound variable. We call this operation variable
closing and define it in Figure 3, again by structural recursion.

\begin{figure}[h]
\centering
\begin{align*}
\text{Substitution of a term for a free name:} & \\
\{z \rightarrow u\} (\text{bvar i}) &= \text{bvar i} \\
\{z \rightarrow u\} (\text{fvar z}) &= u \\
\{z \rightarrow u\} (\text{fvar x}) &= \text{fvar x} \quad \text{when } x \neq z \\
\{z \rightarrow u\} (\text{app t1 t2}) &= \text{app} ([z \rightarrow u] t1) ([z \rightarrow u] t2) \\
\{z \rightarrow u\} (\text{abs t}) &= \text{abs} ([z \rightarrow u] t)
\end{align*}
\end{figure}

\begin{figure}[h]
\centering
\begin{align*}
\text{Properties of substitution:} & \\
\text{subst_fresh} & x \notin \text{FV(t)} \Rightarrow \{x \rightarrow u\} t = t \\
\text{subst_open} & \{x \rightarrow u\} (t^{\omega}) = ([x \rightarrow u] t)^{(x \rightarrow u) \omega} \\
& \text{when term } u \\
\text{subst_open_var} & \{x \rightarrow u\} (t^y) = ([x \rightarrow u] t)^{y} \\
& \text{when } x \neq y \text{ and term } u \\
\text{subst_intro} & \{x \rightarrow u\} (t^x) = t^u \\
& \text{when } x \notin \text{FV(t)}
\end{align*}
\end{figure}

\begin{figure}[h]
\centering
\begin{align*}
\text{3.3 Working With Local Closure} & \\
\text{We use an inductive definition of local closure, given in Figure 1. A}
\text{bound variable on its own is never locally closed: there is no case
for (bvar i). An abstraction is locally closed if the body of that
abstraction, once opened up with a fresh name, is locally closed.}
\text{There are alternative ways to define local closure, but we favor
this style which is similar to the VCLosed relation of McKinna
and Pollack (1993). As they point out, this definition provides a
structural induction principle for locally closed terms:}
\forall x. P(x) \\
\forall t_1 t_2. P(t_1) \Rightarrow P(t_2) \Rightarrow P(\text{app } t_1 t_2) \\
\forall x t. x \notin \text{FV(t)} \Rightarrow P(t) \Rightarrow P(\text{abs } t)
\end{align*}
\end{figure}

\begin{figure}[h]
\centering
\begin{align*}
\text{This definition can be used in many different logics: intensional
or extensional, dependent types or simply typed HOL.}^{\footnote{The freshness assumption is not actually required for this
judgement; we include it to maintain the invariant that a term is never
opened with a name that occurs free in that term.}^{\footnote{In extensional higher order logic the (non-empty) predicate term can
be made into a type of locally closed terms. In intensional dependent
type systems (including Coq), this construction isn’t available. However,
in such systems we can define the inductive family of pre-terms indexed
by the set of bvars occurring free (e.g. McBride and McKinna (2004);
for discussion of inductive families see Dybjer (1994)). Then the section
of this family indexed by the empty set of bvars is exactly the locally
closed terms. Improvements to our style may be possible along these
lines, but for the moment we stick with the representation given above,
for its simplicity of use and uniformity of presentation across different
logics.}}}
\end{align*}
\end{figure}
for theorems versus adequacy for derivations.) In our setting, you can carry this out (on paper, not in Coq) with a fairly simple bijection.

4. Cofinite Quantification

Although the definitions in the previous section adequately represent $\lambda_\text{..}$, they yield insufficiently strong induction principles: in many situations we must rename the variable used to open an abstraction to complete a proof. The need to reason about renaming motivates our use of “cofinite quantification” for inductively defined relations.

In this section, we present a complete style of proof development based on cofinite quantification. We first explain cofinite quantification and give a new definition of $\lambda_\text{..}$ using it. Definitions in this style are sufficient, by themselves, to develop significant amounts of metatheory—e.g., to prove weakening and substitution results—since they naturally balance strong induction and inversion principles with useful introduction forms. In a few situations, however, renaming lemmas are required to build derivations using the cofinitely quantified rules. We show how these lemmas can be derived directly from weakening and substitution. Finally, to be confident that relations defined with cofinite quantification are adequate encodings of $\lambda_\text{..}$, we use renaming to show that they are equivalent to their counterparts shown in Figure 1.

Having described our complete style of development, we then observe that some significant streamlining is possible, notably in cases where one believes in the adequacy of the cofinite presentation of a relation without requiring a formal proof of equivalence with the “exists-fresh” variant. We conclude the section with comparisons to previous work, arguing that the added value of our approach is significant in each case.

4.1 Stronger Induction Principles

The definitions of the local closure and typing relations in Figure 1 use an “exists-fresh” style of quantification. When an abstraction is opened in the premise of a rule (TERM-ABS and TYPING-ABS), the name used in the premise is required to be fresh for the body of the abstraction. Thus, to build a derivation using one of these rules, it suffices to show that there exists a single sufficiently fresh name for which the premise holds. The premises of rules written in this way are easy to satisfy, making them ideal introduction forms.

However, such rules give rise to weak elimination forms (inversion and induction principles). The fact that the premises of a rule defined using exists-fresh quantification need only hold for one particular name means that the corresponding induction hypothesis also holds for one particular name. That name will only be as fresh as the premise of the rule requires, which may not be enough.

For example, consider the proof of the weakening lemma for the typing relation, which allows one to insert additional typing assumptions anywhere into a typing environment:

$$\text{typing\_weaken} : \quad E, G \vdash t : T \Rightarrow \text{ok}(E, F, G) \Rightarrow E, F, G \vdash t : T$$

When proving this by induction on the given typing derivation, in the case for TYPING-ABS we are given a derivation ending with

$$x \notin \text{FV}(t_1) \quad E, G, x: T_1 \vdash t_{1t} : T_2$$

and an induction hypothesis

$$\text{ok}(E, F, G, x:T_1) \Rightarrow (E, F, G, x: T_1 \vdash t_{1t} : T_2)$$

and we must derive

$$E, F, G \vdash \text{abs} \; t_1 : T_1 \rightarrow T_2$$

---

8 These include both low-level technical properties such as subst\_open and more interesting properties such as reflexivity of typing in System $F_\kappa$. 

---

footnote text
for an arbitrary \( F \). By \textsc{typing-ABS}, it suffices to derive
\[
E,F,G,y;T_1 \vdash t_1^y : T_2,
\]
for some name \( y \not\in \text{FV}(t_1) \). Intuitively, we want to use \( y = x \), since the induction hypothesis applies to \( x \). However, the given typing derivation shows that \( t_1^x \) is well typed in the environment \((E,G,x;T_1)\), so we can only derive that \( x \) is fresh for \((E,G)\). We need to type \( t_1^x \) in the extended environment \((E,F,G,x;T_1)\), but \( x \) is not guaranteed to be sufficiently fresh for that environment as that it could appear in \( F \).

The particular name \( x \) appearing in the induction hypothesis is (notionally) the name that occurs in that position of the particular derivation of \( E, G \vdash t : T \) being eliminated in the proof. In an informal proof, we avoid the problem by assuming this derivation uses sufficiently fresh names. (Depending on our informal view, this is usually justified by appeal to the Barendregt Variable Convention (1984), which allows us to assume that the name of the bound variable in the abstraction case is sufficiently fresh.)

Our solution is to strengthen the induction principle for the typing relation by changing its definition to the one at the bottom of Figure 4. This definition differs only in the rule for abstractions, \textsc{typing-ABS}. This new rule intuitively states: "to show that \((\text{abs } t)\) is well typed it suffices to choose some finite set of names, \( L \), and show that \( t^x \) is well-typed for every \( x \) not in \( L \)." We say that this definition of the typing relation uses "cofinite quantification" since we must show that the premise of \textsc{typing-ABS} (""cofinite") holds for all but finitely many names. The set of names \( L \) may be different for each use of the rule. To similarly strengthen the structural induction principle for terms, we redefine term in the cofinite style at the top of Figure 4, and replace all references to term in the reduction relation with this one. Although this new definition of \( \lambda \) differs from the one given in Figure 1, it is equivalent, as we show in Section 4.3.

Using cofinite quantification in the rules provides a stronger induction principle because, in the cases that open abstractions, the induction hypotheses will hold for all names except those in some finite set \( L \), rather than just for a single name. On the other hand, a rule defined using cofinite quantification is nearly as easy to use for introduction as the exists-fresh version because \( L \) holds the finitely many names that could potentially lead to clashes.

Consider again the proof of the weakening lemma—this time, for the cofinite typing relation \( \vdash_c \). In the case for \textsc{typing-ABS}, we are given the assumption
\[
\forall x \not\in L'. (E,G,x;T_1 \vdash_c t_1^x : T_2)
\]
for some finite set of names \( L' \), and an induction hypothesis
\[
\forall x \not\in L'. \text{ ok } (E,F,G,x;T_1) \Rightarrow E,F,G,x;T_1 \vdash_c t_1^x : T_2,
\]
and we must derive
\[
E,F,G \vdash_c \text{ abs } t_1 : T_1 \rightarrow T_2.
\]

By \textsc{typing-ABS}, it suffices to find an \( L \) for which we can show
\[
\forall x \not\in L. (E,F,G,x;T_1 \vdash_c t_1^x : T_2).
\]
Choosing \( L = L' \cup \text{ dom}(F) \) does the trick: any \( x \not\in L \) is also not in \( L' \), so the induction hypothesis directly gives us the result.

As an additional example of the usefulness of the stronger induction hypothesis, we prove the standard "substitution preserves typing" lemma for the cofinite typing relation:

\[
\begin{align*}
\text{typing\_subst} : \\
& (E,z;S,F \vdash_c t : T) \Rightarrow (E \vdash_c u : S) \Rightarrow (E,F \vdash_c [z \mapsto u] t : T).
\end{align*}
\]

The proof of this lemma also demonstrates the ease of working with substitution in a locally nameless setting. It proceeds by induction on the given typing derivation for \( t \). In the case for \textsc{typing-ABS}, we have a derivation that ends in
\[
\forall x \not\in L'. (E,z;S,F,x;T_1 \vdash_c t_1^x : T_2)
\]
for some finite set of names \( L' \), and an induction hypothesis
\[
\forall x \not\in L'. (E,F,x;T_1 \vdash_c [z \mapsto u] t_1^x : T_2),
\]
and we must derive
\[
E,F \vdash_c [z \mapsto u] \text{ abs } t_1 : T_1 \rightarrow T_2.
\]

By the definition of substitution, we simplify the substitution to \( \text{abs } ([z \mapsto u] t_1) \). Then to apply \textsc{typing-ABS}, we must find an \( L \) for which we can show
\[
\forall x \not\in L. (E,F,x;T_1 \vdash_c ([z \mapsto u] t_1^x) : T_2).
\]

We use \texttt{subst\_open\_var} to write \([z \mapsto u] t_1^x\) as \([z \mapsto u] (t_1^x)\) (to match the induction hypothesis). To do this rewriting, we need \( x \) and \( z \) to be distinct. Since the induction hypothesis holds only for \( x \not\in L' \), we choose \( L = L' \cup \{ z \} \).

In this proof and in the proof of weakening, we carefully chose \( L \) when applying \textsc{typing-ABS} to create a new typing derivation. In practice, we use a tactic to instantiate \( L \) with the set of all names appearing in the current proof context: intuitively, showing that a judgement holds for fewer names is easier than showing that it holds for more.

4.2 Renaming

There are a few situations in which we need to apply a derivation rule with a cofinitely quantified premise but only know that this premise holds for one particular fresh name. Alternatively, knowing that an abbreviation type checks, we may wish to use inversion to show that the body of an abbreviation type checks with a particular name, but we do not know that that particular name is excluded from the set \( L \) used with that invocation of \textsc{typing-ABS}. In such situations, we need to explicitly invoke the fact that our judgments are equivariant through a renaming lemma. We present now an example of such a situation and explain how to prove the renaming lemma by reusing the corresponding substitution lemma.
Suppose that we are trying to show that type checking is decidable.\footnote{Coq is a constructive logic, so a proof of this lemma is an algorithm that produces either a typing derivation, or a proof that one isn’t possible.} To show (1), we use C-TYPING-VAR. To show (2), we use the weakening lemma, by which it suffices to show that \(E, x : T_1 \vdash c \ t^\prime : T_2\). This result follows from the assumptions of typing_rename.

In our experience, renaming lemmas are rarely needed. Aside from typing_decidable, the only situations in our developments where they are required are in the proofs of confluence for beta reduction in the lambda calculus and the Calculus of Constructions.

In these proofs, the conclusion of the theorem is an existential statement: “there exists a common reduct…” The cases which require renaming involve rules with binding, where we have an induction hypothesis of the form “for all \(x\) not in \(L\), there exists \(t\) such that…” which we need to use to prove a statement of the form “there exists \(t\) such that for all \(y\) not in \(L\),…” The induction hypothesis is strictly weaker than the statement we are trying to prove. For each name \(x\), it gives us an object \(t\) for which some relation holds. We must use the renaming lemma for that relation to show that the relation holds for all sufficiently fresh names \(y\).

4.3 Equivalence of Exists-Fresh and Cofinite Definitions

At this point we have defined two sets of rules for \(\lambda\)…: one consisting of the original exists-fresh definitions, which corresponds directly to the “paper presentation,” and one that uses cofinite quantification, which yields stronger induction principles. To be confident that the cofinite definitions are adequate, we verify (formally) that they define the same relations as the exists-fresh variants.\footnote{The formal proofs can be found in file STLC_Core_Adequacy.v.}

Note that these proofs of equivalence are straightforward consequences of renaming lemmas.

Let us take the typing relation as an example. The theorem to prove is the following:

\[E \vdash_c t : T \iff E \vdash t : T.\]

Both directions are straightforward inductions. The TYPING-ABS case of the \(\iff\) direction is the only interesting case, where we must show \((E \vdash_c \text{abs } t_1 : T_1 \arrow T_2)\) from the induction hypothesis \((E, x : T_1 \vdash_c t_1^\prime : T_2)\) and an assumption \((x \notin \text{FV}(t_1))\). This case follows by applying C-TYPING-ABS, instantiating \(L\) to be \(\text{FV}(t_1) \cup \text{dom}(E) \cup \{x\}\). We are left with the subgoal

\[\forall y \notin L. (E, y : T_1 \vdash_c t_1^y : T_2)\]

and conclude using typing_rename.

4.4 Streamlining Developments

In general, then, developing metatheory in the locally nameless style involves the following steps.

1. State definitions using exists-fresh quantification.
2. Restate the definitions using cofinite quantification.
3. Prove basic lemmas, such as subst_intro.
4. Prove substitution and weakening lemmas.
5. Prove renaming lemmas.
6. Prove the equivalence of the cofinite and exists-fresh presentations.

However, we can sometimes get by with even less. To begin with, observe that the exists-fresh definitions are needed only in the first and last steps above. The main reason for stating the exists-fresh definitions is to check (informally) that the cofinite definitions adequately encode the relations of interest. Thus, it is only necessary to state the exists-fresh definitions for relations whose adequacy is crucial—those relations that appear hereditarily in the statements of theorems of interest. The style can be streamlined by omitting
steps 1 and 6 for definitions of technical relations that are introduced only to help with proofs, as long as the equivalence of the existential and universal forms of definition is not actually needed in proofs (see Section 4.2).

In fact, the style can be streamlined further, according to need and personal taste. One may (after gaining some familiarity) feel convinced that the cofinite definitions are adequate encodings of one’s informal ideas, without the need of a formal proof of equivalence with their “more obviously adequate” exists-fresh versions. In this case, steps 1 and 6 can be omitted. By doing this, one may be risking adequacy of representation, but one is not risking consistency: the soundness of cofinite-quantified relations checked by Coq is no more in question than the soundness of Coq itself.

Finally, if we omit step 6, we can then prove renaming lemmas—step 5—only as needed, rather than going to the trouble of stating and proving them for all relations. Indeed, as we described in Section 4.2, our experience has been that they are rarely needed.

In short, the core of our style is to define relations using cofinite quantification and to reason about those particular definitions, i.e., steps 2 through 4. The remaining steps need not be carried out for every relation.

4.5 Comparison

As we mentioned in Section 2, we are not the first to propose a locally nameless representation for languages with binders. Nevertheless, our style of formal reasoning using this representation differs from previous work in important ways, requiring much less infrastructure as a result. In this subsection, we give a comparison to the styles used by Gordon (1994), McKinna and Pollack (1993; 1999), and Leroy (2007). To obtain a meaningful comparison, we sometimes apply their styles to our definitions.

Gordon was the first to use a locally nameless representation for formal reasoning about binders: he used them as a concrete datatype upon which to erect a type of named terms up to alpha equivalence. In contrast to our judgements (both the exists-fresh and cofinite versions), his are stated using the close operation instead of open. For example, Gordon’s local closure rule for abstraction (in our notation) is:

\[
\frac{\text{term } t}{\text{term } (\text{abs } (\forall x t))}
\]

Gordon observed that the rule induction principle arising from this form is “too weak” (Gordon 1994, Section 4) in that renaming a variable with a fresh name is sometimes required. He derives a strengthened induction principle by well-founded induction on the length of a term (as is standard). Then, to avoid further reasoning about length, he uses this strengthened induction principle to derive another strengthened induction principle that picks a name fresh for some finite set. However, Gordon’s strengthened induction principle is not the same as that arising from our form of cofinite quantification, since the induction hypothesis only applies to a single variable in the abstraction case. In this respect, Gordon’s strengthened principle is more similar to that of Urban et al. (2007b) than to ours. We get a stronger induction principle and stronger inversion principles automatically from the cofinite quantified forms.

McKinna and Pollack (1993, 1999) proposed a style of development for working with locally named terms. The most significant difference from our style is that it obtains strengthened induction principles by stating universal definitions of inductive relations, which quantify bound variables over as many names as possible. For example, they would state the typing rule for abstractions as follows:

\[
\frac{\forall x \not\in \text{dom}(E). \ (E, x : S \vdash t^x : T)}{E \vdash \text{abs } t : S \rightarrow T} \quad \text{A-TYPING-ABS}
\]

This form of quantification is natural, as there is no arbitrary choice of fresh name in rule A-TYPING-ABS, hence there is at most one derivation of a judgement\( E \vdash t : T \), and the associated induction principle is as strong as possible. The problem with relation \( \vdash \) is that rule A-TYPING-ABS is too weak as an introduction rule, since it requires its premise to hold for every sufficiently fresh name, where cofinite quantification (rule C-TYPING-ABS) allows a finite number of problematic names to be excluded. In McKinna-Pollack style one must prove the equivalence between \( \vdash \) and \( \vdash_a \) (at significant cost in infrastructure) before metatheoretic results—step 4 in our style—can be developed, since \( \vdash \) is too weak in elimination and \( \vdash_a \) is too weak in introduction for either to prove both weakening and substitution lemmas for themselves. Thus our strategy for proving equivalence between definitions, which piggybacks on standard metatheoretic results, fails in the McKinna-Pollack style.

Leroy’s locally nameless solution to the POPLMARK Challenge uses McKinna and Pollack’s ideas. He takes universal definitions as the official ones, obtaining strong induction principles “for free”, and does not explicitly define the exist-fresh definitions at all. However, to make his proofs go through he needs to show that exists-fresh forms are admissible for the universal forms. E.g., Lehrer proves the following lemma for his universally quantified typing relation:

\[
t_{\text{abs}} : \quad x \notin \text{FV}(t) \Rightarrow E, x : T_1 \vdash t^x : T_2 \Rightarrow E \vdash_a \text{abs } t : T_1 \rightarrow T_2.
\]

Such lemmas, which are essentially equivalent to showing that \( \vdash \) and \( \vdash_a \) define the same relation, must be stated and proved for every rule with binding in the development. The proof of such lemmas requires a significant amount of infrastructure related to renaming, so Leroy’s style is a minor streamlining of the McKinna-Pollack style.

4.6 Nominal Isabelle and Local Representation

The nominal Isabelle package (Urban 2007; Urban et al. 2007a,b) supports representation in Isabelle/HOL of languages with binding (as an extended notion of datatype) using a single species of name for both bound and free variables. Using the HOL type definition facility, the package arranges that HOL equality of terms in these represented languages is up-to alpha equivalence, while users need not define alpha equivalence or see any quotient structure. The package supports multiple classes of names (e.g., type variables and term variables) in mutually inductive datatypes (e.g., terms and dependent types) and multiple relations on them (e.g., typing and reduction). Using the Isabelle typeclass facility, the package automatically provides definitions of name freshness and name permutation (renaming) polymorphic over these multiple aspects of a formalization. Importantly, given some syntactic restrictions on the represented languages, the package automatically derives equi-variability lemmas and strengthened induction principles for formalized languages. The nominal package involves no axiomatic extension to Isabelle/HOL, hence is as safe as Isabelle/HOL itself. This nominal representation is a serious alternative to our style for users who want to do formal metatheory. However it is based on a great deal of bespoke infrastructure only available in Isabelle/HOL, there are some syntactic restrictions, and certain aspects of formalization (such as rule inversion and recursive definition over terms) are not

---

\(^{14}\)It is worth emphasizing that the exists-fresh versions are more obviously adequate, in a technical sense: they can be formalized in a very weak theory—Primitive Recursive Arithmetic—while cofinite presentations involve rules with “infinitely many hypotheses” that cannot be formalized in any conservative extension of PRA.
well supported (at time of writing) because of the need to respect alpha equivalence.

A locally nameless formalization (as we propose) or a locally named formalization (à la McKinna-Pollack) can be formalized in many logics and proof tools, depending only on basic notions like datatypes and inductively defined relations.\(^{15}\) In particular, the local representation can be seen as a nominal representation that doesn’t happen to use any nominal binding; in this case the nominal relation of freshness of names for language structures (terms, types, contexts, . . . ) is the same as the local representation’s notion of non-occurrence. This leads to an interesting possibility: local representations can be defined and reasoned about using the nominal Isabelle package, getting the definitions of freshness and renaming and the proofs of equivariance for free from the package.

Interestingly, the nominal package can also automatically infer a strengthened induction principle for inductive relations over local representations (with some syntactic restrictions); see (Urban and Pollack 2007) for more details.

A detailed comparison between nominal representation and locally nameless representation is beyond the scope of this paper, but, as we are focusing on strengthened induction principles, we can remark that the locally nameless strengthened induction principle (e.g., from rule C-TYPING-ABS or rule A-TYPING-ABS) is stronger than the strengthened induction principle automatically derived by the nominal Isabelle package. For a full explanation of the nominal strengthened induction, see Urban et al. (2007b); here it is enough to observe that, when using such an induction principle, one specifies not only the derivation being eliminated (“induction on . . . ”), but also an arbitrary finitely supported structure to be used as a freshness context. Whenever a chosen name appears in an induction hypothesis (e.g., the name that is bound in an abstraction typing rule), that name is known to be fresh for the specified freshness context. This solves many problems, as in the example above of weakening for the typing judgement. However, you must specify the freshness context when invoking induction. In contrast, the locally nameless strengthened induction principles, with their universally quantified fresh names, allow fresh instantiation of these names later, at the point where you have seen the particular cases that must be proved. Since particular cases may involve existential quantifiers, you may not know what the necessary freshness context is until after seeing a case and eliminating an existential quantifier. In such rare examples, the automatically derived nominal strengthened induction principle is not adequate (a detailed hand proof is necessary) while the locally nameless strengthened induction principle works perfectly.

5. Practical Formal Metatheory

We close with a concrete discussion of our Coq realization of working with cofinite quantification. We present the major developments we have carried out—System \(F_{\epsilon}\), the Calculus of Constructions, and extensions of ML—to demonstrate that our techniques scale to larger languages and are expressive enough for wide use. We discuss issues specific to each development (e.g., dealing with multi-binders) as well as properties common to all of them (e.g., the conciseness and robustness on change of our scripts). Finally, we give an overview of the structure of our developments for readers who may wish to use them as a basis for their own formalizations.

5.1 System \(F_{\epsilon}\)

The heart of the POPLMARK Challenge is a proof of type soundness of System \(F_{\epsilon}\). This task was designed as a stress test for formalized metatheory. Our solution closely follows the structure proposed in the appendix of the challenge’s description.

To get a sense of how our solution compares to other solutions, we measured the number of lemmas and steps of reasoning involved in several developments;\(^{16}\) the results are summarized in Figure 5. Our study is restricted to Coq developments since there are no meaningful metrics across different proof assistants. In this comparison, a step of reasoning is defined as the invocation of one tactic, excluding trivial ones: intros (introduces hypotheses into the context), auto (performs automated proof search), and simple variations of these two. We also exclude from the counts material from libraries that are reusable across metatheory developments. We isolate part 1A of the challenge (technical properties of the subtyping relation) from parts 1A+2A (preservation and progress) because some submissions cover only part 1A.

Even when attention is restricted to Coq developments, the comparison is necessarily rough, since the developments were carried out by different people with different preferences in the amount of automation they used, etc. Nevertheless, a few interesting points emerge. The first line of the table shows that our style compares well against Vouillon’s development, which uses pure de Bruijn indices—a representation commonly thought to be quite effective. The solutions that do not use a standard de Bruijn or locally nameless representation (Stump and Hirschowitz & Maggesi) require significantly more steps. The last line shows that our style requires significantly less infrastructure than Leroy’s solution, a locally nameless solution faithful to McKinna and Pollack’s way of dealing with the quantification of names. Chlipala’s is a variation on an earlier version of Leroy’s development and is the only one that is shorter than ours, mainly because Chlipala’s proof script relies heavily on development-specific tactics.

Although clever automation and development-specific tactics can make scripts extremely small (Nipkow [2001] gives an impressively automated proof of confluence, for example), they can also make them slow (hard to work with) and brittle (hard to reuse). Brittness emerges when automation fails after some changes to earlier definitions or lemmas: it can be hard to know exactly what has been broken and how to fix it. Of course, the other extreme of no automation at all is also unmanageable, especially for metatheory proofs, which involve many related arguments. Moreover, too little automation can also make scripts brittle: even a tiny change in definitions or statements of lemmas will require update. Aiming for maximal robustness, our convention is to rely as much as possible on automation for low-level details (e.g., proving that a term is locally closed) while giving all the “important” arguments (the ones that would appear in a rigorous paper proof) manually.

5.2 Calculus of Constructions (CoC)

This formalization contains three main components. The first is a proof of confluence for parallel beta-reduction. Its use of the variable-closing operation requires slightly more advanced reasoning about binders. We have also used this development to evaluate our scripts for robustness: we first completed a proof of confluence for the pure lambda calculus and then migrated it to CoC, preserving the same flow of arguments and just adding the extra cases and updating the names of some hypotheses.

The second part consists of a set of inversion results for the typing and the conversion judgments, which are quite technical in themselves but do not involve much difficulty with respect to

\(^{15}\) Some specialized tactics, as in the accompanying Coq code http://arthur.chargueraud.org/research/2007/binders/, are useful, but not essential. In any case these tactics are very much simpler than the nominal Isabelle package.

\(^{16}\) All are publicly available from http://alliance.seas.upenn.edu/~plclub/cgi-bin/poplmark/.
binders. This part is representative of the effort required for the CoC formalization: we felt that the largest share of the work was to understand and be able to state clearly the arguments involved. The fact that the proofs were machine-checked appeared to be only a way to force oneself to understand all the details of the proofs formalized, not as an extra burden to be dealt with.

The third and last part is the core of the preservation proof, which shares a significant amount of structure with the corresponding proof for System $F_{\Sigma}$ (only the steps of reasoning are more complex to follow). Compared to the System $F_{\Sigma}$ development, the fact that terms and types are represented in a uniform way in CoC and that proofs are more complex reduces the binding infrastructure from about 60% of the total development down to 35% percent.

Compared to Barras’s formalization of CoC in Coq (1997) using a pure de Bruijn representation, our style saves us from numerous issues associated with shifting indices—particularly bothersome when working with CoC’s dependent types—and so, unlike Barras, we did not need clever engineering of the statements of lemmas to make them fit the representation used.

### 5.3 Extensions of ML

We also investigated how to extend the basic $\lambda_{\omega}$ development with features such as datatypes, fixpoints, references, exceptions, pattern matching, and ML-style polymorphic types. Then, building on the individual successes of each of these extensions, we set up the formalization of a language containing all of these features together. This experiment demonstrates that our style can be applied to full-featured programming languages, not just tiny lambda-calculi.

Pleasingly, most of the arguments that would be omitted as “trivial” on paper and which do not involve binders are solved by the automatic tactic. In particular, the infrastructure lemmas do not need any change when we add features like pairs or references. The key proofs—preservation and progress—contain respectively in 60 and 80 lines, which seems reasonable for a system with 24 typing rules.

A key idea involved in this development is a treatment of multi-binders, used to encode ML type schemes, patterns, and recursion. The basic operation on a multi-binder is to open it with a $\text{bind}$ binder above the current position is represented as $(\text{tm}_{\,i\,,j})$. Free variables are still represented as one single name and substitution lemmas are still stated in terms of an atomic substitution from one name to one term. Then, to prove subject reduction, for example, the substitution lemma is applied iteratively in the cases involving multi-binders.

### 5.4 Organization of Our Developments

Each of our developments is divided into three main parts: trusted definitions, infrastructure, and core lemmas.

The trusted definitions part contains the description of the language formalized—the syntax of the language, its semantics, and its type system, if any—and the statements of the main theorems which are to be proven about the language. This part is the only material that needs to be trusted, in the following sense: if one is convinced about the adequacy of these definitions and trusts that Coq correctly type-checks all proofs in the development, then one can have confidence that the system has been formally certified.

The infrastructure part sets up the machinery required for the core lemmas and consists of several components:

1. Language-specific specializations of tactics for working with cofinite quantification, e.g., to automatically choose a set $L$ when applying a rule that uses cofinite quantification.

2. Proofs about properties of substitution (Figure 2).

3. Proofs that local closure is preserved by various operations, e.g., substitution (Section 3.3).

4. Regularity lemmas which state that relations contain only locally closed terms (Section 3.3).

5. Hints to enable Coq’s automation to use regularity lemmas.

Note that the lemmas about substitution are similar from one language to the next. When setting up a new development, those who wish to get to “interesting” proofs quickly may state as axioms properties of substitution and regularity lemmas, proving them only after they believe that their main proofs will go through. In this manner, one can “test-drive” a language definition without having to revise infrastructure proofs as the definition is modified.

Finally, the core lemmas part contains the lemmas that would normally be stated in an informal presentation. The statements closely match their informal counterparts, and the proofs contain few uninteresting details—facts about local closure and freshness side conditions are almost always handled automatically.
We conclude this section with a breakdown of our developments by the number of definitions and lemmas in each part, as well as the number of proof steps required to prove the lemmas; the results are summarized in Figure 6. The number of trusted definitions gives an idea of the size of the language formalized. The number of core lemmas gives an idea of the theoretical complexity of the development; recall that these lemmas correspond closely with results that would be proven informally. The amount of infrastructure, in terms of both lemmas and steps, is proportional to the number of binding constructs and relations in the language. A comparison between figures from the last two columns suggests that the amount of infrastructure work is reasonable compared to the core proof work—especially since infrastructure proofs follow a standard pattern and are easily set up. More precisely, the ratio “infrastructure over core proofs” increases with the number of binding constructs involved but decreases with the inherent complexity of the development.

6. Conclusion

We have argued for a novel style for formalizing programming-language metatheory, based on a combination of locally nameless representation with a cofinite idiom for quantifying free names in inductive definitions.

This design satisfies the evaluation criteria of POPLMARK. First, our presentation is transparent: the proofs closely follow their informal equivalents. Second the overheads of the approach are low: we do not need manual proofs of freshness side-conditions nor reasoning about alpha-equivalence, and only on rare occasions do we need to justify well-formedness of objects (e.g. local closure) or explicitly rename variables. When we do, the required proofs of renaming lemmas follow with virtually no infrastructure. At the same time, there is no need for external tools, and the style works in any general purpose theorem prover (although we found Coq to be well-suited to the task). In our scripts, definitions and results about variables, freshness, and environments are factored into a reusable library. Finally, experience with a number of large developments suggests that the approach is “complete” in the informal sense that any language definition and accompanying reasoning techniques that would be accepted as informally correct can be carried out in this style. Our formalizations provide a good starting point for new work—reports from early adopters confirm that modifying them is much easier than starting a new development from scratch.

In the future, we would like to integrate our approach with the Otter tool (Sewell et al. 2007), which automatically generates Coq (and other) definitions from compact, high-level language descriptions. A more ambitious next step is to study the possibility of automatically generating the required infrastructure proofs, which follow a fairly simple pattern. Finally, we would like to explore certified programming of tools such as type checkers and compilers that must deal with binders. In the long term, we hope that our techniques will help with the verification of both specifications and implementations of programming languages.

Acknowledgments Many thanks to the members of the Penn PLClub for extensive feedback and discussion, and to Michael Norrish and Christian Urban for important technical insights and to the POPL reviewers, Bob Harper, Peter Sewell, and Karl Crary for helping improve the final version. This work was supported by NSF grant 545886, CRI: Machine Assistance for Programming Language Research. Pollack thanks the International Centre for Mathematical Sciences (Edinburgh) Workshop on Abstraction, Substitution and Naming and the EU Framework 6 Coordination Action 510996, TYPES.

References


