Generic Zero-Cost Reuse for Dependent Types

LARRY DIEHL, University of Iowa, USA
DENIS FIRSOV, University of Iowa, USA
AARON STUMP, University of Iowa, USA

Dependently typed languages are well known for having a problem with code reuse. Traditional non-indexed algebraic datatypes (e.g. lists) appear alongside a plethora of indexed variations (e.g. vectors). Functions are often rewritten for both non-indexed and indexed versions of essentially the same datatype, which is a source of code duplication.

We work in a Curry-style dependent type theory, where the same untyped term may be classified as both the non-indexed and indexed versions of a datatype. Many solutions have been proposed for the problem of dependently typed reuse, but we exploit Curry-style type theory in our solution to not only reuse data and programs, but do so at zero-cost (without a runtime penalty). Our work is an exercise in dependently typed generic programming, and internalizes the process of zero-cost reuse as the identity function in a Curry-style theory.

Additional Key Words and Phrases: dependent types, generic programming, reuse

1 INTRODUCTION

Dependently typed languages (such as Agda [Norell 2007], Coq [The Coq Development Team 2008], Idris [Brady 2013], or Lean [de Moura et al. 2015]) can be used to define ordinary algebraic datatypes, as well as indexed versions of algebraic datatypes that enforce various correctness properties. For example, we can index lists by natural numbers to enforce that they have a particular length (i.e. \( \text{Vec}_A : \mathbb{N} \rightarrow \star \)). Similarly, we can index lists by two elements to enforce that they are ordered and have a lower and upper bound (i.e. \( \text{OList}_{A,R} : A \rightarrow A \rightarrow \star \)). We can even combine these two forms of indexing to enforce that lists have all of the aforementioned correctness properties (i.e. \( \text{OVec}_{A,R} : A \rightarrow A \rightarrow \mathbb{N} \rightarrow \star \)).

Which datatype a programmer uses depends upon how much correctness they wish to enforce at the time a function is written, versus proving correctness as a separate step sometime later (corresponding to intrinsic and extrinsic correctness proofs of functions). Certain types tend to be better suited to writing intrinsically correct functions than others, e.g. it is natural to define a safe lookup function that takes a \( \text{Vec} \) as an argument, and a correct sort function that returns an \( \text{OList} \).

However, once we have written a function using a suitable indexed variant of a datatype, reusing the function to define a corresponding version for the unindexed (or less indexed) datatype can be painful. We may also wish to delay extrinsic verification by choosing to pay the price later, i.e. when we reuse a function over unindexed (or less indexed) datatypes to define a corresponding function over more indexed datatypes. We refer to the former direction as \textit{forgetful reuse}, and to the latter direction as \textit{enriching reuse}.

One source of pain is \textit{manually writing} functions over some datatypes by reusing functions over differently indexed variants of the same underlying datatypes. Another source of pain is that reusing functions involves linear-time conversions between differently indexed types, resulting in a runtime \textit{performance penalty} incurred by practicing the good software engineering practice of
code reuse. 1 In this paper we address both of these problems, for both the forgetful and enriching directions of reuse, by:

(1) Defining generic combinators to incrementally attack the problem of reuse for various types, where each combinator application results in simplified subgoals (similar to tactics).

(2) Ensuring that the combinators are closed operations with respect to a type abstraction, which can be eliminated to obtain reused functions at zero-cost (i.e. no performance penalty).

Our primary contributions are:

(1) Section 4.2: Generic combinator solutions to zero-cost forgetful program reuse (combinator allArr2arr, handling the type of non-dependent functions), and proof reuse (combinator allP12pi, handling the type of dependent functions).

(2) Section 4.3: Generic combinator solutions to zero-cost enriching program reuse (combinator arr2allArrP, handling the type of non-dependent functions), and proof reuse (combinator p12allP1P, handling the type of dependent functions).

(3) Section 5.3: A generic combinator solution to zero-cost forgetful data reuse (combinator ifix2fix, handling the type of fixpoints for generically encoded datatypes).

(4) Section 5.5 & Section 6.3: Generic combinator solutions to zero-cost enriching data reuse (combinators f1x2f1x and f1x2f1xP, handling the type of fixpoints for generically encoded datatypes).

Our work has two notable limitations:

(1) We can only perform zero-cost conversions when indexed types store their index data as erased arguments (or more accurately, when the arguments of non-indexed and indexed types match after erasure). This may not always be desirable, and has consequences such as needing to recompute the length of a vector if it is needed dynamically.

(2) Forgetful program (or proof) reuse is not possible when the domain of the indexed function is constrained, such as the vector head function, when using our allArr2arr combinator. We discuss this further in Section 7.3, including how a small extension of our work (i.e. also constraining the domain of the non-indexed function) could partially address this problem.

The remainder of our paper proceeds as follows:

• Section 2: We review background material, covering the Curry-style type theory that our results are developed within, and providing intuition for why zero-cost conversions are motivated by Curry-style type theory.

• Section 3: We explain the primary problems (linear-time reuse of programs, proofs, and data) we are solving through concrete examples, and provide manual solutions (zero-cost, or constant time, reuse), which our primary contribution combinators generalize via generic programming.

• Section 4: We generically solve the problems of (both forgetful and enriching) zero-cost program and proof reuse (as combinators for the types of non-dependent and dependent functions).

• Section 5: We generically solve the problems of (both forgetful and enriching) zero-cost data reuse (as combinators for the type of fixpoints).

• Section 6: We evaluate our work on a more complex example of reuse between unchecked and checked STLC terms, and also extend our functional data reuse enrichment to the relational setting, where there is a premise on the non-indexed data.

1 Wherever we say linear-time conversions, we are assuming reasonable implementations of the conversions. Of course, the time complexity can be worse for poorly implemented conversions.
We briefly summarize the type theory, the Calculus of Lambda Eliminations (CDLE), that the results of this paper depend on. For full details on CDLE, including semantics and soundness results, please see the previous papers [Stump 2017, 2018a,b]. The main metatheoretic property proved in the previous work is logical consistency: there are types which are not inhabited. Cedille is an implementation of CDLE, and all the code appearing in this paper is Cedille code.

CDLE is an extrinsic (i.e. Curry-style) type theory, whose terms are exactly those of the pure untyped lambda calculus (with no additional constants or constructs). The type-assignment system for CDLE is not subject-directed, and thus cannot be used directly as a typing algorithm. Indeed, since CDLE includes Curry-style System F as a subsystem, type assignment is undecidable [Wells 1999]. To obtain a usable type theory, Cedille thus has a system of annotations for terms, where the annotations contain sufficient information to type terms algorithmically. But true to the extrinsic nature of the theory, these annotations play no computational role. Indeed, they are erased both during compilation and before formal reasoning about terms within the type theory, in particular by definitional equality (see Figure 1).

CDLE extends the (Curry-style) Calculus of Constructions (CC) with implicit products, primitive heterogeneous equality, and intersection types:

- \( \forall x : T \cdot T' \), the implicit product type of Miquel [2001]. This can be thought of as the type for functions which accept an erased input of type \( x : T \), and produce a result of type

2 The Cedille formalization accompanying this paper is available at: https://github.com/larrytheliquid/generic-reuse
There are term constructs $\Lambda \ x \ . \ t$ for introducing an implicit input $x$, and $t \ - t'$ for instantiating such an input with $t'$. The implicit arguments exist just for purposes of typing so that they play no computational role and equational reasoning happens on terms from which the implicit arguments have been erased.

- $t_1 \simeq t_2$, a Curry-style heterogeneous equality type. The terms $t_1$ and $t_2$ are required to be typed, but need not have the same type. We introduce this with a constant $\beta$ which erases to $\lambda \ x \ . \ x$ (so our type-assignment system has no additional constants, as promised); $\beta$ proves $t \simeq t$ for any typeable term $t$. Combined with definitional equality, $\beta$ proves $t_1 \simeq t_2$ for any $\beta$-equal $t_1$ and $t_2$ whose free variables are all declared in the typing context. We eliminate the equality type by rewriting, with a construct $\varphi \ q \ - \ t$. Suppose $q$ proves $t_1 \simeq t_2$ and we synthesize a type $T$ for $t$, where $T$ has several occurrences of terms definitionally equal to $t_1$. Then the type synthesized for $\varphi \ q \ - \ t$ is $T$ except with those occurrences replaced by $t_2$. The construct $\varphi \ q \ - \ t_1\{t_2\}$ casts a term $t_2$ (of any type) to type $T$, provided that $t_1$ has type $T$ and $q$ proves $t_1 \simeq t_2$. The point of using the term $\varphi \ q \ - \ t_1\{t_2\}$ at type $T$, instead of the term $t_1$, is that the $\varphi$ term erases to $|t_2|$. Note that the types of the terms are not part of the equality type itself, nor does the elimination rule require that the types of the left-hand and right-hand sides are the same to do an elimination.

- $\iota \ x : T, T'$, the dependent intersection type of Kopylov [2003]. This is the type for terms $t$ which can be assigned both the type $T$ and the type $[t/x]T'$, the substitution instance of $T'$ by $t$. In the annotated language, we introduce a value of $\iota \ x : T, T'$ by construct $[ t, t' ]$, where $t$ has type $T$ (algorithmically), $t'$ has type $[t/x]T'$, and the erasure $|t|$ is definitionally equal to the erasure $|t'|$. There are also annotated constructs $t.1$ and $t.2$ to select either the $T$ or $[t.1/x]T'$ view of a term $t$ of type $\iota \ x : T, T'$.

It is important to understand that the described constructs are erased before the formal reasoning (e.g. when checking if 2 terms are definitionally equal), according to the erasure rules in Figure 1.

### 2.2 Curry-Style Typing

There is an intuitive explanation for why zero-cost (i.e. no performance penalty) conversion should be possible between differently indexed data (i.e. List and Vec) and differently indexed programs (i.e. appl and appV). In a Curry-style theory, the same underlying untyped term can be typed multiple different ways. Therefore, if it is possible to type a term as both a list and a vector, then there is actually no need to do any conversion at all because the same term can inhabit both types! In a type-annotated (rather than type-assignment) setting, this translates to having 2 distinct terms at two distinct types, whose erasures are equal.

**Curry-Style Data.** As an example of Curry-style data, consider the standard definitions of Church-encoded lists and vectors below. Note that a left-pointing triangle ($\triangleleft$) is type ascription syntax for Cedille definitions, rather than a conventional colon (:). The direction of the triangle is meant to convey that the definition will be checked using the checking (rather than inferring) mode of a bidirectional type checker.

```
List $\triangleleft \triangleleft \ * \rightarrow * = \Lambda \ A. \ \forall \ X : * . \ X \rightarrow (A \rightarrow X \rightarrow X) \rightarrow X.

\mathsf{nilL} \triangleleft \forall \ A : * . \ \mathsf{List} \ A = \Lambda \ A, X. \ \lambda \ cN, cC . \ cN.

\mathsf{consL} \triangleleft \forall \ A : * . \ A \rightarrow \mathsf{List} \ A \rightarrow \mathsf{List} \ A = \Lambda \ A. \ \lambda \ x, xS . \ \Lambda \ X. \ \lambda \ cN, cC . \ cC \times (xs \cdot X cN cC).

Vec $\triangleleft \triangleleft \ * \rightarrow \mathsf{Nat} \rightarrow * = \Lambda \ A, n . \ \forall \ X : \mathsf{Nat} \rightarrow * .

\ X \ \mathsf{zero} \rightarrow (\forall \ n : \mathsf{Nat} . \ A \rightarrow X \ n \rightarrow X \ (\mathsf{suc} \ n)) \rightarrow X \ n .

\mathsf{nilV} \triangleleft \forall \ A : * . \ \mathsf{Vec} \ A \ \mathsf{zero} = \Lambda \ A, X. \ \lambda \ cN, cC . \ cN.
```
Generic Zero-Cost Reuse for Dependent Types

\[ \text{consV} \triangleq \forall A : \star. \forall n : \text{Nat}. A \rightarrow \text{Vec} A n \rightarrow \text{Vec} A (\text{suc} n) = \Lambda A,n. \lambda x,xs. \Lambda X. \lambda cN,cC. cC -n x (xs -X cN cC). \]

Notice that the only difference between the list constructor terms (\text{nilL} and \text{consL}) and vector constructor terms (\text{nilV} and \text{consV}) is the number of implicit abstractions (e.g. \( \Lambda n \)) and implicit applications (e.g. \(-n\)). According to the erasure rules of Figure 1, this means that after erasure, \text{nilL} and \text{nilV} share the same underlying untyped term (and the same holds for \text{consL} and \text{consV}):

\[
|\text{nilL}| = |\text{nilV}| = \lambda cN,cC. cN
|\text{consL}| = |\text{consV}| = \lambda x,xs,cN,cC. cC x (xs cN cC)
\]

Curry-Style Programs. As an example of Curry-style programs, consider the standard definitions of the append function for Church-encoded lists and vectors below:

\[
\text{appL} \triangleq \forall A : \star. \text{List} A \rightarrow \text{List} A \rightarrow \text{List} A
= \Lambda A. \lambda xys. xys -(\text{List} A \rightarrow \text{List} A)
(\lambda ys. ys)
(\Lambda xs. \lambda x,ih,ys. \text{consL} -A x (ih ys)).
\]

\[
\text{appV} \triangleq \forall A : \star. \forall n : \text{Nat}. \text{Vec} A n \rightarrow \forall m : \text{Nat}. \text{Vec} A m \rightarrow \text{Vec} A (\text{add} n m)
= \Lambda A. \lambda xxs. xxs -(\lambda n. \forall m : \text{Nat}. \text{Vec} A m \rightarrow \text{Vec} A (\text{add} n m))
(\Lambda m. \lambda yys. yys)
(\Lambda n. \Lambda xs,x,ih. \Lambda m. \lambda yys. \text{consV} A -(\text{add} n m) x (ih -m yys)).
\]

Like before, \text{appL} and \text{appV} only differ by implicit abstractions and applications. An additional difference is that \text{appL} uses \text{consL} in its second branch, while \text{appV} uses \text{consV} in its second branch. Because (as seen above) the erasure \(|\text{consL}|\) is equal to the erasure \(|\text{consV}|\), it follows that \text{appL} and \text{appV} also share the same underlying untyped term:

\[
|\text{appL}| = |\text{appV}| = \lambda xs. xs (\lambda yys. yys) (\lambda x,ih,ys,cN,cC. cC x (ih ys cN cC))
\]

2.3 Inductive Datatypes

The enriching direction of reuse requires dependent function types, which must be proven by induction on their inputs using eliminators. The Church-encoded \text{List} and \text{Vec} datatypes of Section 2.2 do not support induction, due to a result by Geuvers [2001]. However, Stump [2018a] shows that the dependent intersection (using the \( \iota \)-type from Figure 1) of an impredicative Church-encoded type with a predicate, representing what it means for the type to be inductive, does support induction (or an eliminator):

\[ \text{List} \triangleq \star \rightarrow \star = \lambda A. \lambda xs : \text{ListChurch} A. \text{ListInductive} A xs. \]
\[ \text{elimList} \triangleq \forall A : \star. \forall P : \text{List} A \rightarrow \star.
P (\text{nilL} -A) \rightarrow
(\forall xs : \text{List} A. \Pi x : A. P xs \rightarrow P (\text{consL} -A x xs)) \rightarrow
\Pi xs : \text{List} A. P xs \]

Above, \text{ListChurch} is the renamed definition of \text{List} from Section 2.2. Intersection-type versions of the constructors \text{nilL} and \text{consL} can also be defined. We refer readers interested in the definitions of \text{nilL}, \text{consL}, and \text{elimList} to Stump [2018a], as this section only depends on their type-level interface (rather than their term-level implementation). We also assume a corresponding \( \iota \)-type definition of \text{Vec} (in terms of \text{VecChurch}), its constructors (\text{nilV} and \text{consV}), and its eliminator (\text{elimVec}).

An important thing to point out is that \text{List} is defined as the intersection of the \text{ListChurch} and \text{ListInductive} types, and that intersection pairs (i.e. \([\ell_1, \ell_2]\)) erase to their first components
(i.e. \( t_1 \) of type List\( \text{Church} \)) by Figure 1. Hence, the erased \( i \)-style \( \text{nilL} \) is the same as the erased Church-style \( \text{nilL} \) (and the same holds for both styles of cons\( \text{L} \)).

3 THE PROBLEM & MANUAL SOLUTION

Section 2.2 shows how differently indexed data (e.g. List and Vec) and programs (e.g. app\( \text{L} \) and app\( \text{V} \)) can share the same erased untyped terms in a Curry-style dependent type theory. Now we consider the problem of manually reusing data and programs, in both the forgetful (e.g. Vec to List, and app\( \text{V} \) to app\( \text{L} \)) and enriching (e.g. List to Vec, and app\( \text{L} \) to app\( \text{V} \)) directions.

3.1 The Problem: Manual Linear Time Reuse

First, we review how to manually reuse data and programs using linear-time conversions, which is already possible in popular dependently typed languages. Then (in Section 3.2), we show how Cedille lets us manually derive zero-cost (or constant time) conversions from the linear-time conversions. After erasure, Cedille programs consist of untyped lambda calculus terms. Hence, the cost model of Cedille is similar to that of other untyped lambda calculus implementations, such as Racket [Flatt and PLT 2010], which is the output language of compiled Cedille programs. Finally, to aid readability, from now on we omit implicit abstractions (e.g. \( \Lambda \text{A} \)) and implicit applications (e.g. \( -\text{A} \)).

**Linear Time Forgetful Data Reuse.** We can convert a vector to a list by iteration:

\[
\text{v2l} : \forall A : \star. \forall n : \text{Nat}. \text{Vec A } n \to \text{List A} = \text{elimVec} \text{nilL} (\lambda x,ih. \text{consL } x \text{ ih}).
\]

The conversion above only requires iteration, rather than induction, because the codomain \( \text{List A} \) does not depend on the domain \( \text{Vec A } n \). If we explicitly supplied the motive (or, predicate) \( P \) to \( \text{elimVec} \), it would ignore its argument (i.e. \( P = \lambda xs. \text{List A} \)).

**Linear Time Enriching Data Reuse.** We can convert a list to a vector by induction:

\[
\text{l2v} : \forall A : \star. \Pi xs : \text{List A}. \text{Vec A } (\text{len } xs) = \text{elimList} \text{nilV} (\lambda x,ih. \text{consV } x \text{ ih}).
\]

The conversion above requires induction, rather than iteration, because the codomain \( \text{Vec A } (\text{len } xs) \) depends on the domain \( \text{List A} \). If we explicitly supplied the motive \( P \) to \( \text{elimList} \), it would depend on its argument (i.e. \( P = \lambda xs. \text{Vec A } (\text{len } xs) \)).

**Linear Time Forgetful Program Reuse.** After defining the type synonyms \( \text{AppL} \) and \( \text{AppV} \) for the types of list and vector append, respectively, forgetful reuse of vector append to define list append corresponds to writing a function from \( \text{AppV} \) to \( \text{AppL} \):

\[
\text{appV2appL} : \forall \text{AppV} \to \text{AppL}
\]

\[
\text{appV2appL} = \lambda \text{appV},xs,ys. \text{v2l} (\text{appV } (\text{l2v } xs) (\text{l2v } ys)).
\]

The function \( \text{appV2appL} \) first reuses vector append (app\( V \)) by applying app\( V \) to the result of translating both list arguments (\( xs \) and \( ys \)) to vectors (via \( \text{l2v} \)). Then, it translates the result of app\( V \) from a vector to a list (via \( \text{v2l} \)).

---

3 While we omit most implicit abstractions and applications in this paper, the current implementation of Cedille only supports a limited form of type inference. Our accompanying Cedille formalization does not omit any implicits.
Generic Zero-Cost Reuse for Dependent Types

**Linear Time Enriching Program Reuse.** Enriching reuse of list append to define vector append is the difficult direction, which requires proving a lemma stating that once a vector has been converted to a list (via v2l, or forgetful data reuse), the length of the output list is equal to (or, preserves) the length index of the input vector:

\[
\text{v2lPresLen} \triangleq \forall A : \star. \forall n : \text{Nat}. \Pi \mathit{xs} : \text{Vec} A n. n \equiv \mathit{len} (\mathit{v2l} \mathit{xs})
\]

Recall (from Section 2.1) that \(\beta\) is the reflexivity constructor of an equality of type \(t \equiv t\) for any term \(t\), and that \(\rho\) is a rewrite primitive that exchanges occurrences of \(t\) with occurrences of \(t/\text{Var}\) in the goal, when given evidence that \(t \equiv t/\text{Var}\). The proof of \(\text{v2lPresLen}\) is an easy induction, which rewrites by the inductive hypothesis (ih) in the cons case of the input vector.

It is not possible to reuse a function of type \(\text{AppL}\) to define a function of type \(\text{AppV}\) in general, because the result of the second function has specific index requirements (namely, that the output vector length is the sum of the input vector lengths). Enriching function reuse must be modulo an additional premise, which establishes a relationship between the input and output datatype indices.\(^4\) The premise necessary to define \(\text{AppV}\) in terms of \(\text{AppL}\) requires list length \((\mathit{len})\) to distribute through list append \((\mathit{appL})\):

\[
\text{LenDistAppL} \triangleq \text{AppL} \to \star = \lambda \mathit{appL}. \Pi A : \star. \Pi \mathit{xs}, \mathit{ys} : \text{List} A. \mathit{add} (\mathit{len} \mathit{xs}) (\mathit{len} \mathit{ys}) \equiv \mathit{len} (\mathit{appL} \mathit{xs} \mathit{ys}).
\]

\[
\text{appL2appV} \triangleq \Pi \mathit{appL} : \text{AppL}. \text{LenDistAppL} \mathit{appL} \to \text{AppV}
\]

After binding the arguments to \(\text{appL2appV}\), the initial goal type, and the resulting goal type after each rewrite (using \(\rho\)), appears as a comment (i.e. to the right of the syntactic comment delimiter // on each line).

Initially, the length of the goal vector is the sum of the lengths of both input vectors \(\mathit{xs}\) and \(\mathit{ys}\). First, we use the previously proven lemma \(\text{v2lPresLen}\) to state our goal in terms of the lists resulting from converting input vectors \(\mathit{xs}\) and \(\mathit{ys}\) (via \(\mathit{v2l}\)). After reusing \(\mathit{appL}\) applied to both converted list, we would like to convert the result to a vector (via \(\mathit{l2v}\)) and return it. However, the dependent data reuse function \(\mathit{l2v}\) returns a vector indexed by the \(\mathit{len}\) of its input list, but the current goal is stated in terms of a sum (i.e. \(\mathit{add}\) rather than \(\mathit{len}\)). Therefore, we must first rewrite the goal using our premise that length distributes through append, so that we may finally return the result of applying \(\mathit{l2v}\).

### 3.2 Manual Solution: Zero-Cost Reuse

Now we derive zero-cost (constant-time) data and program conversions from the linear-time equivalents of Section 3.1. Linear time reuse (e.g. in Section 3.1) is already possible in conventional Church-style type theories, but zero-cost reuse is additionally possible in Curry-style type theories. This is semantically motivated because a Curry-style term can inhabit multiple types, so conversion is semantically unnecessary (as explained in Section 2.2). The zero-cost (data and program) conversions in this section are all defined in two parts:

\[^4\] Enriching data reuse of a list (List) as a vector (Vec) does not require a premise, but in general enriching data reuse is also modulo a premise. For example, a list can only be enriched to an ordered list (OList) modulo a premise that the list is sorted. In Section 6, we give an example of enriching a raw lambda-term to a typed lambda-term, provided a premise that the raw term is well-typed with respect to a typing relation.
An extensional identity proof about the corresponding linear-time conversion.

The actual zero-cost conversion, defined using $\phi$ from Figure 1, the linear-time conversion, and the extensional identity proof.

Zero-Cost Forgetful Data Reuse. First, we prove that the v2l conversion is extensionally the identity function:

$$v2lId \triangleq \forall A : \star. \forall n : \text{Nat.} \Pi xs : \text{Vec} A n. v2l \; xs \equiv xs$$

= elimVec $\beta$ $(\lambda x,ih. \rho \; ih - \beta)$.

Next, we use the $\phi$ primitive (of Figure 1) to return the vector input $xs$ at type $\text{List} A$, by appealing to the proof ($v2lId$) that $v2l \; xs$ is equal to $xs$.

$$v2l! \triangleq \forall A : \star. \forall n : \text{Nat.} \text{Vec} A n \rightarrow \text{List} A$$

= $\lambda xs. \phi \; (v2lId \; xs) - (v2l \; xs) \{xs\}$.

The $\phi$ expression erases to the term within the braces ($\{xs\}$) by the erasure rules of Figure 1, hence the erasure $|v2l!|$ is the identity function. Thus, $v2l!$ converts a vector to a list in constant time, as applying $v2l!$ is definitionally equal to applying the identity function in CDLE:

$$|v2l!| = \lambda xs. xs$$

By convention, we suffix a conversion function with a bang (!) to denote its zero-cost equivalent.

Zero-Cost Enriching Data Reuse. The enriching direction of zero-cost data reuse follows the same pattern as the forgetful direction, by first proving an extensional identity ($l2vId$), and then using it to define a zero-cost version ($l2v!$) via $\phi$:

$$l2vId \triangleq \forall A : \star. \Pi xs : \text{List} A. l2v \; xs \equiv xs$$

= elimVec $\beta$ $(\lambda x,ih. \rho \; ih - \beta)$.

$$l2v! \triangleq \forall A : \star. \Pi xs : \text{List} A. \text{Vec} A (\text{len} \; xs)$$

= $\lambda xs. \phi \; (l2vId \; xs) - (l2v \; xs) \{xs\}$.

And similarly, $l2v!$ converts any list ($xs$) to a vector at zero-cost:

$$|l2v!| = (\lambda xs. xs)$$

Zero-Cost Forgetful Program Reuse. For zero-cost forgetful program reuse of vector append, we prove the following extensional identity: Applying the conversion appV2appL to any implementation of vector append ($f$), and both list arguments, is equal to applying vector append ($f$) to both list argument that have been zero-cost converted to vectors (via $l2v!$).

$$\text{appV2appLId} \triangleq \Pi f : \text{AppV}. \forall A : \star. \Pi xs,ys : \text{List} A.$$

= $\lambda f,xs,ys. \quad / \quad v2l \; (f \; (l2v \; xs) \; (l2v \; ys)) \equiv f \; xs \; ys$

$\rho \; (l2vId \; xs) - \quad / \quad v2l \; (f \; xs \; (l2v \; ys))$

$\rho \; (l2vId \; ys) - \quad / \quad v2l \; (f \; xs \; ys)$

$\rho \; (v2lId \; (f \; xs \; ys)) - \quad / \quad f \; xs \; ys \equiv f \; xs \; ys$

$\beta$

The right-side of the equality in the goal begins with $f \; xs \; ys$, because the zero-cost conversions $l2v! \; xs$ and $l2v! \; ys$ definitionally reduce to $xs$ and $ys$, respectively. We rewrite twice (for $xs$ and $ys$) by the extensional identity lemma for $l2v$ (using $l2vId$). Then, we rewrite once (for $f \; xs \; ys$) by the extensional identity lemma for $v2l$ (using $v2lId$), after which our goal is solvable by reflexivity ($\beta$).

We define the zero-cost conversion appV2appL! using $\phi$ and the identity lemma appV2appLId applied to the vector append argument ($f$) and both list arguments ($xs$ and $ys$):
Generic Zero-Cost Reuse for Dependent Types

$$\text{appV2appL!} \equiv \text{AppV} \rightarrow \text{AppL}$$

$$= \lambda f, x, y. \varphi (\text{appV2appLId} f x y) - (\text{appV2appL} f x y) \{f (\text{l2v!} x) (\text{l2v!} y)\}.$$ 

The erased zero-cost conversion \(\text{appV2appL!}\) also definitionally reduces to the identity function:

$$|\text{appV2appL!}| = \lambda f. \lambda x, y. |f (\text{l2v!} x) (\text{l2v!} y)|$$

The \(\text{l2v!}\) zero-cost conversions reduce to applications of the identity function. Then, the body of the \(\lambda f\) abstraction \(\eta\)-contracts to \(f\), such that the entire expression reduces to the identity function.

**Zero-Cost Enriching Program Reuse.** The zero-cost enriching program reuse of list append requires the following extensional identity: Applying the conversion \(\text{appL2appV}\) to any implementation of list append \((f)\), a proof of the length distributivity premise \((p)\), and both vector arguments, is equal to applying list append \((f)\) to both vector argument that have been zero-cost converted to lists (via \(\text{v2l!}\)).

$$\text{appL2appVId} \equiv \Pi f : \text{AppV}. \Pi q : \text{LenDistAppL} f.$$ 

$$= \forall A : \ast. \forall n, m : \text{Nat}. \Pi x, y : \text{Vec} A n. \Pi x, y : \text{Vec} A m.$$ 

$$\text{appL2appV} f q x y \equiv (v2l! x) (v2l! y).$$

$$\varphi (v2lId x) - // 12v (f (v2l x) (v2l y)) \equiv f x y$$

$$\varphi (v2lId y) - // 12v (f (x y) (v2l y)) \equiv f x y$$

$$\varphi (12vId (f x y)) - // f x y \equiv f x y$$ 

Once again, the zero-cost conversion \((\text{appL2appV!})\) is defined in terms of the linear time conversion \((\text{appL2appV})\), \(\varphi\), and the extensional identity \((\text{appL2appVId})\):

$$\text{appL2appV!} \equiv \Pi f : \text{AppL. LenDistAppL} f \Rightarrow \text{AppV}$$

$$= \lambda f. \Lambda q. \lambda x, y. (\text{appL2appVId} f q x y) - (\text{appL2appV} f q x y) \{f (\text{v2l!} x) (\text{v2l!} y)\}. $$

The implication \((\Rightarrow)\) to the right of the premise \((\text{LenDistAppL} \ \text{appL})\) of \(\text{appL2appV}\) is syntax for a non-dependent implicit (or, erased) product (i.e. a \(\forall\) with no dependency on the quantified variable). The fact that the zero-cost conversion \(\text{appL2appV}\) uses an erased premise (compared to the non-erased premise in the linear-time conversion \(\text{appL2appV}\)) is crucial, allowing \(\text{appL2appV}\) to also erase to the identity function:

$$|\text{appL2appV!}| = (\lambda f. \lambda x, y. |f (\text{v2l!} x) (\text{v2l!} y)|)$$

The implicit abstraction \(\Lambda q\) is discarded by erasure, allowing the erasure \(|\text{appL2appV!}|\) to \(\eta\)-contract to the identity function (similar to how \(|\text{appV2appL!}|\) reduces).

4 GENERIC PROGRAM & PROOF REUSE

Section 3.2 gives a zero-cost solution to the problem of linear-time data and program reuse problem presented in Section 3.1. However, the reused definitions in Section 3.2 are manually derived. Beginning with this section, and for the remainder of this paper, we solve the problem of zero-cost reuse generically.
In Section 4.1 we review the type of dependent identity functions (IdDep), which captures a pattern appearing in the manual zero-cost solution to reuse (Section 3.2). The IdDep type is the dependent generalization of the non-dependent Id type introduced by Firsov et al. [2018]. Section 4.2 generically solves the problem of forgetful program and proof reuse, which corresponds to defining IdDep-closed combinators for the type of non-dependent functions (for program reuse), and the type of dependent functions (for proof reuse). Section 4.3 defines 2 additional combinators to generically solve the problem of enriching program and proof reuse.

4.1 Type of Dependent Identity Functions

As explained in Section 2.2, an (erased) term may have several possible types in a Curry-style theory. Of particular importance to our work is that the identity function, represented by the untyped lambda term (λ x. x), can have many possible types. We have seen several examples of this in Section 3.1, where the zero-cost conversions v2l!, l2v!, appV2appL!, and appL2appV! all erase to the identity function. Thus, it makes sense to define a type of dependent identity functions for any domain A : ✷ and codomain B : A → ✷. We informally denote the type of dependent identity functions by (a : A) ≤ B a. Inhabitance of the type (a : A) ≤ B a represents the existence of a term F, such that |F| = (λ x. x), and the existence of a typing derivation for the judgement Γ ⊢ F : Π a : A. B a.

Section 3.1 manually defines zero-cost conversions using a proof that the linear-time conversion is (after erasure) an identity operation. Hence, the zero-cost conversion depends on two parts:

1. The linear-time conversion.
2. A proof that the linear-time conversion is extensionally an identity function.

Now we formally derive the type of dependent identity functions (a : A) ≤ B a in Cedille as IdDep A B, which abstractly represents both zero-cost conversion parts as a dependent function (Π) returning a dependent pair (Σ):

\[ \text{IdDep} :: \forall A : ✷. \forall B : A \rightarrow ✷. ✷ \]

The type IdDep A B is defined when A is a type and B is a family of types indexed by A. Inhabitants of IdDep A B take elements of (a : A) to elements of (b : B a), and a proof that b is propositionally equal to a (using the heterogeneous equality type ≃ from Figure 1). We can represent the two parts more explicitly by deriving an introduction rule that takes the (conversion) function f and its extensional identity proof as arguments:

\[ \text{intrIdDep} :: \forall A : ✷. \forall B : A \rightarrow ✷. \Pi f : \Pi a : A. B a. f a ≃ a \rightarrow \text{IdDep} A B \]

\[ = \lambda f,q,a. \Pi (f a) (q a). \]

In practice, it is more convenient to introduce elements of IdDep directly in terms of the underlying ΠΣ representation, rather than using intrIdDep.

Now we define the crucial elimination rule elimIdDep, which exposes the witness F at type Π a : A. B a, whose erasure is the identity function:

\[ \text{elimIdDep} :: \forall A : ✷. \forall B : A \rightarrow ✷. \Pi a : A. B a \rightarrow \Pi c,a. \varphi (\text{proj2} (c a)) - (\text{proj1} (c a)) \{a\}. \]

The elimination rule elimIdDep uses \varphi to return the input a, originally at type A, at type (B a) using the extensional identity proof (proj2 (c a)), where c : IdDep A B. From the erasure rules

5 The dependent pair type Sigma can be derived in Cedille just like the inductive List and Vec types, as explained in Section 2.3.
of CDLE (in Figure 1), it follows that for any dependent identity function \( c \) of type \( \text{IdDep} \ A \ B \), \( |\text{elimIdDep} \ c| = |F| = (\lambda \ a. \ a) \).

Finally, notice how the definition of \( \text{elimIdDep} \) abstracts out a part (i.e. the use of \( \varphi \) and the extensional identity proof) of the zero-cost conversion definitions \((v2l!, 12v!, \text{appV2appL}!, \text{and appL2appV}! \)) from Section 3.2. In subsequent sections we define \( \text{IdDep} \)-closed combinators, taking \( \text{IdDep} \) inputs and producing an \( \text{IdDep} \) output. Because the combinators always return an \( \text{IdDep} \), well typed combinator definitions guarantee the existence of zero-cost conversions (whose witness we can always produce by applying \( \text{elimIdDep} \)).

We will also use non-dependent identity function counterparts \( \text{Id}, \text{intrId}, \text{and elimId} \) (of \( \text{IdDep}, \text{intrIdDep}, \) and \( \text{elimIdDep}, \) respectively), where \( A : \star \), but also \( B : \star \) (rather than \( B : A \rightarrow \star \)). These are trivially derivable from the dependent versions, so we omit their definitions. Note that our derived non-dependent \( \text{Id} \) type is isomorphic to the \( \text{Id} \) type introduced by Firsov et al. [2018]. Also, note that the usage of the non-dependent \( \text{Id} \) type to zero-cost convert between values of different types is similar to how the \( \text{Coercible} \) type class [Breitner et al. 2016] is used in Haskell, and the relationship between them is further explained in our related work (Section 7.2).

Recall our informal notation of type \( \text{IdDep} \ A \ B \) as \((a : A) \leq B \ a\). The informal notation is inspired by Miquel [2001], who uses a non-dependent version of this notation \((A \leq B)\) for a subtyping judgement derivable in a Curry-style theory with implicit products. Indeed, our \( \text{Id} A B \) is inhabited when \( A \) is a subtype of \( B \), and correspondingly all of our combinators can also be understood as internalized subtyping inference rules (we discuss the relationship with subtyping further in our related work, Section 7.1). When there is an identity function from \( A \) to \( B \) (i.e. \( \text{Id} A B \)), and both \( A \) and \( B \) are functions, it becomes confusing to talk about domains and codomains (e.g. "domain" could refer to the identity function domain \( A \), or the domain of the non-identity function \( A \), or the domain of the non-identity function \( B \)). To avoid confusion, and inspired by the relationship with subtyping, we refer to the domain of an identity function as the \text{subtype} and the codomain as the \text{supertype} (thus, we can non-ambiguously refer to the domain and codomain of the subtype, and the same for the supertype).

4.2 Forgetful Reuse

Now we define generic solutions to the problem of forgetful program and proof reuse as \( \text{IdDep} \)-closed combinators for the non-dependent and dependent function types, respectively. As a demonstration of using our generic solution, we redo the \( \text{appV} \) reuse example from Section 3.2 in terms of our combinators (we also provide an additional example of reusing the proof of vector append associativity).

The examples in this section assume the existence of identity functions (i.e. values of type \( \text{IdDep} \)) to convert between lists and vectors:

\[
\begin{align*}
v2l &: \forall A : \star. \forall n : \text{Nat}. \text{Id} (\text{Vec} A n) (\text{List} A) \\
12v &: \forall A : \star. \text{IdDep} (\text{List} A) (\lambda xs. \text{Vec} A (\text{len} xs))
\end{align*}
\]

We delay the task of defining \( v2l \) and \( 12v \) to Section 5, where we define both identity functions as examples of using our generic data reuse combinators.

4.2.1 Program Reuse Combinator. All the names of our combinators are short descriptions of their return types. For example, below, \( \text{allArr2arr} \) has return type \( \text{Id} (\forall i : I. X i \rightarrow X' i) (Y \rightarrow Y') \). Mnemonically, \( \text{allArr2arr} \) returns an identity function from \( \text{allArr} (\forall \rightarrow) \) to \( (2 \ \text{arr} \rightarrow) \). We define \( \text{allArr2arr} \) as:

\[
\begin{align*}
\text{allArr2arr} &: \forall I : \star. \forall X,X' : I \rightarrow \star. \forall Y,Y' : \star. \\
& \quad \Pi r : Y \rightarrow I.
\end{align*}
\]
Larry Diehl, Denis Firsov, and Aaron Stump

\[ \Pi \ c1 : \text{IdDep} \ Y \ (\lambda \ y. \ X \ (r \ y)). \]
\[ \Pi \ c2 : \Pi \ y : \ Y. \ \text{Id} \ (X' \ (r \ y)) \ Y'. \]
\[ \text{Id} \ (\forall \ i : \ I. \ X \ i \to X' \ i) \ (\forall \ y : \ Y). \ \text{Id} \ (\forall \ i : \ I. \ X \ i \to X' \ i) \ Y' \ (\forall \ y : \ Y). \]

The combinator \text{allArr2arr} is a generic solution to forgetful \textit{non-dependent function} reuse (or, forgetful \textit{program} reuse). For example, it can solve a problem like the one below, where black boxes (■) represent arbitrary (not necessarily the same) types:

\[ \text{Id} \ (\forall \ n : \ \text{Nat}. \ \text{Vec} \ A \ n \to \ ■) \ (\text{List} \ A \to \ ■) \]

The domain of the subtype is an indexed type and the domain of the supertype is a non-indexed type. For example, if we were to solve the problem above with \text{allArr2arr}, we would set index type \(I\) to \(\text{Nat}\), the type family \(X\) to \(\text{Vec} \ A\), and the non-indexed type \(Y\) to \(\text{List} \ A\). The codomains of the subtype and supertype (i.e., the black boxes, or \(X'\) and \(Y'\), respectively) cannot depend on the explicit domain arguments (i.e., \(X\) and \(Y\)), which is why we say that \text{allArr2arr} solves the problem of non-dependent function reuse. However, the codomain of the subtype \((X')\) can depend on the implicit index argument (of type \(I\)). This covers all the implicit arguments of \text{allArr2arr}, and now we explain the explicit arguments:

- The argument \(r\) is the \textit{refinement function}, computing an index of type \(I\) from the non-indexed type \(Y\), e.g., \(\text{len} : \text{List} \ A \to \text{Nat}\).
- The argument \(c1\) is the \textit{contravariant dependent identity function}. It enriches the non-indexed supertype domain \(y : Y\), e.g., \(\text{xs} : \text{List} \ A\), to the indexed subtype domain \(X \ (r \ y)\), e.g., \(\text{Vec} \ A \ (\text{len} \ \text{xs})\).
- The argument \(c2\) is the \textit{covariant non-dependent identity function}. It forgets the indexed subtype codomain \(X' \ (r \ y)\) as the non-indexed supertype codomain \(Y'\).

Non-dependent \text{allArr2arr} is defined in terms of its dependent version, \text{allPi2Pi}, which we present subsequently.

4.2.2 Program Reuse Example. Now we demonstrate zero-cost forgetful program reuse of vector append to define list append, in terms of \text{allArr2arr}. We produce an identity function (Id) from \text{AppV} to \text{AppL}, called \text{appV2appl} below:

\[ \text{appV2appl} \triangleq \text{Id} \ \text{AppV} \ \text{AppL} \]
\[ = // \text{Id} \ (\forall \ A : \star. \ ■) \ (\forall \ A : \star. \ ■) \]
\[ \text{copyType} \ (\Lambda \ A. \)
\[ // \text{Id} \ (\forall \ n : \ \text{Nat}. \ \text{Vec} \ A \ n \to \ ■) \ (\text{List} \ A \to \ ■) \]
\[ \text{allArr2arr} \ \text{len} \ \text{v2l} \ (\lambda \ \text{xs}. \)
\[ // \text{Id} \ (\forall \ m : \ \text{Nat}. \ \text{Vec} \ A \ m \to \ ■) \ (\text{List} \ A \to \ ■) \]
\[ \text{allArr2arr} \ \text{len} \ \text{v2l} \ (\lambda \ \text{ys}. \)
\[ // \text{Id} \ (\text{Vec} \ A \ (\text{add} \ (\text{len} \ \text{xs}) \ (\text{len} \ \text{ys}))) \ (\text{List} \ A \)
\[ \text{v2l})))). \]

Our example includes goal types in comments, where each goal above illustrates the part of the problem solved by a combinator application below. The black boxes (■) hide the parts of the goals that are not relevant to what is being solved by the combinators below. We begin by handling the impredicative quantification \(\forall \ A : \star.,\) which is present in both \text{AppV} and \text{AppL}, using the easy to define auxiliary definition \text{copyType} from Figure 2. Next, we apply \text{allArr2arr} twice to handle contravariantly enriching both arguments from lists to vectors. In these applications, \(r\) is the length function (\text{len}), and \(c1\) is the enriching data reuse function 12v. Additionally, \(c2\) becomes the remainder of the \text{appV2appl} definition, giving us access to the list arguments \text{xs} and \text{ys}. Finally,
we covariantly forget the return type from a vector to a list using the forgetful data reuse function v2l.

Note that appV2appL, above, simultaneously captures the linear-time conversion function and the extensional identity proof from Section 3.2 (i.e. the former appV2appL and appV2appL1d). We can recover the actual zero-cost conversion by applying elimId to our identity function:

\[ \text{appV2appL!} \triangleq \text{AppV} \rightarrow \text{AppL} = \text{elimId appV2appL} \]

Previously, we used a bang (!) suffix as a syntactic convention for defining a zero-cost conversion. Now, we can also think of the elimination rule of identity functions (elimId) as a bang operator, because applying it to any Id results in a zero-cost conversion. From now on we omit defining the actual zero-cost conversions (like appV2appL!), because they can always be recovered by applying the elimination rule for the type of identity functions.

4.2.3 Proof Reuse Combinator. The combinator allPi2pi is a generic solution to forgetful dependent function reuse (or, forgetful proof reuse). For example, it can solve a problem like the one below:

\[ \text{Id} \left( \forall n : \text{Nat}. \ \Pi \ x : \text{Vec} \ A \ n. \ 
\right) \left( \Pi \ x : \text{List} \ A. \ 
\right) \]

The subtype codomain may depend on the subtype (vector) domain, and the supertype codomain may depend on the supertype (list) domain. The definition of allPi2pi follows:

\[ \text{allPi2pi} \triangleq \forall I : \cdot. \ \forall X : I \rightarrow \cdot. \ \forall X' : \Pi \ i : I. \ X i \rightarrow \cdot. \]

\[ \forall Y : \cdot. \ \forall Y' : Y \rightarrow \cdot. \]

\[ \Pi r : Y \rightarrow I. \ 
\]

\[ \Pi c1 : \text{IdDep} Y (\lambda y. X (r y)). \]

\[ \Pi c2 : Y : Y. \ \text{Id} (X' (r y) (\text{elimIdDep} c1 y)) (Y' y). \]

\[ \text{Id} (\forall i : I. \ \Pi x : X i. \ X' i x) (\Pi y : Y. \ Y' y) = \lambda r,c1,c2,f. \ \text{pair} (\lambda y. \ \text{elimId} (c2 y) (f -(r y) (\text{elimIdDep} c1 y))) \beta. \]

Compared to allArr2arr, the I, Y, r, and c1 arguments are the same. However, now the subtype codomain X' may depend on its indexed domain X i, and the supertype codomain Y' may depend on its non-indexed domain Y. The explicit argument c2 is also different:

- The argument c2 is still the covariant non-dependent identity function. However, now it forgets the indexed subtype codomain X' (r y) (elimIdDep c1 y) as the non-indexed supertype codomain Y' y. In the X' application, y is zero-cost converted from Y to X (r y) via the contravariant argument c1.

Notice that allPi2pi is reminiscent of the subtyping rule between dependent functions. However, it is different in that it requires the additional r argument to handle the additional quantification (\forall i : I) in the subtype. Furthermore, the zero-cost conversion elimIdDep c1 y is explicitly performed in the type of argument c2, which depends on argument c1. In a subtyping rule the dependency and the zero-cost conversion would be hidden, because the subsumption rule would implicitly change the type of y by appealing to the subtyping premise represented by c1.

4.2.4 Proof Reuse Example. As an example of zero-cost proof reuse, we demonstrate how to prove associativity of list append from the associativity of vector append.\(^6\) First, we create type synonyms for the theorem of list append associativity (AssocL) and vector append associativity

\(^6\) The concept of proof reuse being "zero-cost" may seem odd, as systems like Coq erase proofs during program extraction. However, proofs in intentional type theory may sometimes have computational content we wish to preserve, and hence it can be valuable to zero-cost reuse such proofs. For example, a proof that an element exists in a list can be zero-cost reused as a natural number index into the list.
(Assoc\(\text{V}\)), parameterized by a definition of list append (App\(\text{L}\)) and vector append (App\(\text{V}\)), respectively:

\[
\text{Assoc}\(\text{L}\) \triangleright App\(\text{L}\) \rightarrow \star = \lambda \text{app}\(\text{L}\).
\]

\[
\forall A : \star. \ \Pi xs,ys,zs : \text{List} \ A. \\
\text{app}\(\text{L}\) (app\(\text{L}\) xs ys) zs \equiv \text{app}\(\text{L}\) xs (app\(\text{L}\) ys zs)).
\]

\[
\text{Assoc}\(\text{V}\) \triangleright App\(\text{V}\) \rightarrow \star = \lambda \text{app}\(\text{V}\).
\]

\[
\forall A : \star. \forall n : \text{Nat}. \ \Pi xs : \text{Vec} \ A n. \\
\forall m : \text{Nat}. \ \Pi ys : \text{Vec} \ A m. \forall o : \text{Nat}. \ \Pi zs : \text{Vec} \ A o. \\
\text{app}\(\text{V}\) (app\(\text{V}\) xs ys) zs \equiv \text{app}\(\text{V}\) xs (app\(\text{V}\) ys zs)).
\]

Next, we reuse any proof of Assoc\(\text{V}\) to prove Assoc\(\text{L}\) at zero-cost:

\[
\text{assoc}\(\text{V2assoc}\(\text{L}\) \triangleright \forall \text{app}\(\text{V}\) : \text{App}\(\text{V}\). \\
\text{Id} (\text{Assoc}\(\text{V}\) \text{app}\(\text{V}\)) (\text{Assoc}\(\text{L}\) (\text{elim}\text{Id appVappL appV})) \\
// \text{Id} (\forall A : \star. \star) (\forall A : \star. \star) \\
\text{= copyType} (\Lambda A. \star) \\
// \text{Id} (\forall n : \text{Nat}. \ \Pi xs : \text{Vec} \ A n. \star) (\Pi xs : \text{List} \ A. \star) \\
\text{allPi2pi len l2v (} \lambda xs. \\
// \text{Id} (\forall m : \text{Nat}. \ \Pi ys : \text{Vec} \ A m. \star) (\Pi ys : \text{List} \ A. \star) \\
\text{allPi2pi len l2v (} \lambda ys. \\
// \text{Id} (\forall o : \text{Nat}. \ \Pi zs : \text{Vec} \ A o. \star) (\Pi zs : \text{List} \ A. \star) \\
\text{allPi2pi len l2v (} \lambda zs. \\
// \text{Id} (\text{app}\(\text{V}\) (\text{app}\(\text{V}\) xs ys) zs \equiv \text{app}\(\text{V}\) xs (\text{app}\(\text{V}\) ys zs))) \\
// (\text{app}\(\text{V}\) (\text{app}\(\text{V}\) xs ys) zs \equiv \text{app}\(\text{V}\) xs (\text{app}\(\text{V}\) ys zs))) \\
\text{id)))).
\]

Notice that the identity function assoc\(\text{V2assoc}\(\text{L}\) is parameterized by any implementation (app\(\text{V}\)) of the type of vector append (App\(\text{V}\)). We apply the type synonym for vector append associativity (Assoc\(\text{V}\)) directly to vector append (app\(\text{V}\)), but the type synonym for list append associativity (Assoc\(\text{L}\)) expects an implementation of list append (i.e. a value of type App\(\text{L}\)). Hence, we apply Assoc\(\text{L}\) to the result of zero-cost converting app\(\text{V}\) to a list append, via elim\text{Id appVappL appV}, which uses our previously defined identity function appVappL.

Once again, we begin solving assoc\(\text{V2assoc}\(\text{L}\) by copying the type parameter \(\Lambda\) via copyType. Next, we apply allPi2pi to handle the 3 primary arguments to the theorem. The final goal is solvable by the auxiliary identity combinator for identity functions (id from Figure 2). Before erasure, the supertype of the final goal has instances of elim\text{Id appVappL appV}, instead of app\(\text{V}\). Similarly, before erasure, the subtype of the final goal has instances of elim\text{IdDep l2v xs}, instead of xs (and similarly for ys and zs). However, because these are zero-cost conversions, after erasure (as depicted in the comment above) the goal is simply solvable by id.

### 4.3 Enriching Reuse

Now we generically solve enriching program and proof reuse as IdDep-closed combinators for the non-dependent and dependent function types, respectively. Each forgetful program and proof reuse (Section 4.2) combinator returns a non-dependent identity function (Id). In contrast, each enriching version returns a dependent identity function (IdDep), where the dependency is used to define the premise necessary for enrichment. We demonstrate our enriching combinators by redoing the app\(\text{L}\) enriching program reuse example from Section 3.2.

#### 4.3.1 Program Reuse Combinator

The combinator arr2all\(\text{arrP}\) is a generic solution to enriching non-dependent function reuse (or, enriching program reuse). Recall (from Section 3.2) that (in
Generic Zero-Cost Reuse for Dependent Types

general) enriching program reuse must be performed modulo a premise required for the enrichment to be possible. For example, \texttt{arr2allArrP} can solve a problem like the one below:

\[
\text{IdDep (List } A \to \boxed{\_}\text{)} (\lambda \ f. (\Pi x : \text{List } A. \boxed{\_}) \Rightarrow \forall n : \text{Nat. Vec } A n \to \boxed{\_})
\]

Program enrichment returns a dependent identity function (\text{IdDep}). An additional implicit (erased) \textit{premise} argument (to the left of \(\Rightarrow\)) appears in the supertype, and the premise has a dependent domain whose type is equal to the subtype's domain (e.g. \text{List } A). The codomain of the premise can depend on the subtype (e.g. \(f\)), in addition to the domain of the premise (\(xs\)). The definition of \texttt{arr2allArrP} follows:

\[
\text{arr2allArrP} \triangleq \forall Y, Y' : \star. \forall P : Y \to Y' \to \star. \forall I : \star. \forall X, X' : I \to \star.
\]

\[
\Pi r : Y \to I.
\]

\[
\Pi c1 : \forall i : I. \text{Id} (X i) Y.
\]

\[
\Pi c1' : \forall i : I. \Pi x : X i. i \approx r \ (\text{elimId } c1 \ x).
\]

\[
\Pi c2 : \forall i : I. \Pi x : X i. \text{IdDep } Y' (\lambda y'. P \ (\text{elimId } c1 \ x)) \ y' \Rightarrow X' (r \ (\text{elimId } c1 \ x)).
\]

\[
\text{IdDep } (Y \to Y') (\lambda f. (\Pi y : Y. P y (f y)) \Rightarrow \forall i : I. X i \to X' i) = \lambda Y, Y', P, I, X, X'. \text{pi2allPiP} -Y -\langle \cdot, \cdot \rangle -P -I -X -\langle \cdot, \cdot, \cdot \rangle.
\]

Enriching \texttt{arr2allArrP} shares the following implicit arguments with forgetful \texttt{allArr2arr} (from Section 4.2): \(Y, Y', I, X, \) and \(X'\), as well as the explicit argument \(r\). However, the premise \(P\) appears as an additional implicit argument of the subtype, which may depend on both the domain \(Y\) and codomain \(Y'\). Now we explain the differing explicit arguments:

- The argument \(c1\) is the \textit{contravariant non-dependent identity function}. It forgets the indexed supertype domain \(X i\), e.g. \text{Vec } A n, as the non-indexed subtype domain \(Y\), e.g. \text{List } A.

- The argument \(c1'\) is the \textit{index preservation property}. It requires the index \(i\) of the supertype domain \((x : X i)\) to equal the refinement (using refinement function \(r\)) of zero-cost converting \(x\) (using the identity function \(c1\)), e.g. \(n \approx \text{len} \ (\text{elimId } v2l \ xs)\), where \(xs : \text{Vec } A n\) and \(c1\) is \(\text{v2l}\).

- The argument \(c2\) is the \textit{covariant dependent identity function}. It enriches the non-indexed subtype codomain \(Y'\) as the indexed supertype codomain \(X' i\). The enrichment codomain also gets an additional implicit premise argument \(P\), which may depend on both the supertype domain and codomain.

Notice that \(X'\), in the supertype of argument \(c2\), is applied to the refinement \(r \ (\text{elimId } c1 x))\), rather than the index \(i\). This makes \texttt{arr2allArrP} easier to use, as the implementation automatically rewrites by \(c1'\) (the index preservation property)! We point out the consequence of this automatic rewrite in the following example.

4.3.2 Program Reuse Example. Below, we redo the enriching reuse of list append example from Section 3.2. While our forgetful identity function type combinators attack 2 pieces at a time (the domains of the supertype and subtype), the enriching function type combinators attack 3 (the additional piece being the premise, whose type is duplicated from the subtype domain).

\[
\text{appL2appV} \triangleq \text{IdDep AppL (\lambda appL. \text{LenDistAppL} \ appL \Rightarrow \text{AppV})}
\]

\[
= \text{copyTypeP} (\Lambda A. \text{IdDep (List } A \to \boxed{\_}) \Rightarrow \forall n : \text{Nat. Vec } A n \to \boxed{\_})
\]

\[
= (\lambda f. (\Pi x : \text{List } A. \boxed{\_}) \Rightarrow \forall n : \text{Nat. Vec } A n \to \boxed{\_})
\]

\[
\text{arr2allArrP} \text{len } v2l \text{v2lPresLen} (\Lambda n. \lambda xs. \text{IdDep (List } A \to \boxed{\_}) \Rightarrow \forall m : \text{Nat. Vec } A m \to \boxed{\_})
\]

\[
\text{arr2allArrP} \text{len } v2l \text{v2lPresLen} (\Lambda m. \lambda ys. \text{IdDep (List } A) \Rightarrow \forall zs. \text{len } zs \approx \text{add} (\text{len } xs) (\text{len } ys)
\]

Publication date: July 2018.
Larry Diehl, Denis Firsov, and Aaron Stump

// subst l2v)).

⇒ Vec A (add (len xs) (len ys)))

We begin with the auxiliary copyTypeP combinator (from Figure 2), which is a version of copyTypeP that also handles the premise (i.e. the 3rd piece). Next, we use our enriching combinator arr2allArrP to handle both inductive arguments of append, to which we supply a proof of length preservation (v2PresLen from Section 3.1) as an additional argument. This leaves us with the goal type (after erasure) in the final comment. Before erasure, as explained in the assocV2assoc example of Section 4.2, xs (in the supertype of the goal) is the zero-cost conversion elimId v21 xs (and similarly for ys and zs). Finally, we discharge the premise by rewriting, via the auxiliary combinator subst (from Figure 2). As an argument, subst takes the identity function to apply after rewriting, which is the enriching data reuse l2v (in this example).

The final goal type includes the sum of the lengths of both input vectors, rather than the sum of two vector indices. This convenience is a result of the automatic rewriting performed in the implementation of our arr2allArrP combinator! In contrast, the manual definition of appendV in Section 3.1 needed to manually rewrite by v2PresLen for both append inputs.

4.3.3 Proof Reuse Combinator. We include the definition of the enriching proof reuse combinator (pi2allPiP), for reference, below. We do not describe it in detail, as the extensions to handle the dependent arguments in codomains X' and Y' follow the same pattern as allPi2pi from Section 4.2.

\[
\begin{align*}
\forall Y : \star. \forall X : Y \rightarrow \star. \forall P : X \rightarrow Y. Y' \rightarrow \star. \\
\Pi r : Y \rightarrow I. \\
\Pi c1 : \forall i : I. Id (X i) Y. \\
\Pi c1' : \forall i : I. \Pi x : X i. i \equiv r (elimId c1 x). \\
\Pi c2 : \forall i : I. \Pi x : X i. IdDep (Y' (elimId c1 x)) (\lambda y'. \\
P (elimId c1 x) y') \Rightarrow Y' (r (elimId c1 x)) (\rho (c1'-i x) - x)). \\
IdDep (\Pi y : Y. Y' y) (\lambda f. (\Pi y : Y. P y (f y)) \Rightarrow \\
\forall i : I. \Pi x : X i. \lambda x. elimIdDep (\rho (c1'-i x) - c2 x) (f (elimId c1 x)) -(p (elimId c1 x))) \beta.
\end{align*}
\]

Note that in the type of c2, the second argument of the dependent supertype codomain X' must use c1' to rewrite r (elimId c1 x) to i, because the variable x has type X i. The \(\varsigma\) operator of Cedille takes an equality proof of type \(t \simeq t'\), and returns the symmetric version of type \(t' \simeq t\). The implementation also exposes that the index preservation property (c1') is indeed automatically rewritten (via \(\rho\)).

We omit the example of enriching proof reuse of list append associativity. It is very similar to the forgetful proof reuse example of vector append associativity, because v2PresLen becomes an additional argument, making the premise the trivial Unit type.

5 GENERIC DATA REUSE
In this section we give the generic zero-cost solution to the problem of linear-time data reuse presented in Section 3.1, and manually solved in Section 3.2. In Section 5.1, we review a type of least fixed points, used to generically encode datatypes. Section 5.3 covers generic forgetful data reuse, and Section 5.3 covers generic enriching data reuse.
Generic Zero-Cost Reuse for Dependent Types

\[
\begin{align*}
\text{id} & : \forall A : \star. \text{Id} A A = \lambda a. \text{pair} a \beta. \\
\text{copyType} & : \forall F : \star \to \star. \forall G : \star \to \star.
(\forall A : \star. \text{Id} (F A) (G A)) \to \text{Id} (\forall A : \star. F A) (\forall A : \star. G A) = \lambda c, x. \text{pair} (\lambda A. \text{elimId} (c -A) (x -A)) \beta.
\end{align*}
\]

\[
\begin{align*}
\text{copyTypeP} & : \forall F : \star \to \star. \forall P : \Pi A : \star. F A \to \star. \forall G : \star \to \star.
(\forall A : \star. \text{IdDep} (F A) (\lambda x. P A x \Rightarrow G A)) \to \text{IdDep} (\forall A : \star. F A) (\lambda x. (\forall A : \star. P A x \Rightarrow G A)) = \lambda c, x. \text{pair} (\lambda p, A. \text{elimIdDep} (c -A) (x -A) -(p -A)) \beta.
\end{align*}
\]

\[
\begin{align*}
\text{subst} & : \forall Y : \star. \forall I : \star. \forall X : I \to \star. \forall r : Y \to I.
\forall i : I. \text{IdDep} Y (\lambda y. X (r y)) \Rightarrow \text{IdDep} Y (\lambda y. r y \Rightarrow i) \Rightarrow X i = \lambda c, y. \text{pair} (\lambda q. \text{elimIdDep} c y) \beta.
\end{align*}
\]

\[
\begin{align*}
\text{supplyPrem} & : \forall Y : \star. \forall I : \star. \forall X : I \to \star. \forall i : I. \text{IdDep} Y (\lambda y. \forall i : I. X i y) \Rightarrow \forall i : I. \text{IdDep} Y (X i) = \lambda c, y. \text{pair} (\text{elimIdDep} c y) \beta.
\end{align*}
\]

Fig. 2. Auxiliary identity combinator, combinators to copy a shared impredicative quantification, combinator to rewrite by an equality constraint, and combinator to supply a premise.

5.1 Type of Least Fixed Points

Section 2.3 reviews the work by Stump [2018a] to manually derive induction principles for Church-encoded datatypes via intersecting (using \(\iota\)) with an inductivity predicate. Firsov and Stump [2018] solved the same problem generically, by deriving a least fixed point type for any functor, composed of 4 pieces:

1. An object mapping \(F : \star \to \star\).
2. An arrow mapping \((fmap : \forall X,Y : \star. (X \to Y) \to F X \to F Y)\).
3. A proof of the identity law for \(fmap\).
4. A proof of the composition law for \(fmap\).

Firsov et al. [2018] improved the solution by deriving a least fixed point type that only requires 2 pieces:

1. A type scheme \(F : \star \to \star\).
2. An identity mapping \((imap : \forall X,Y : \star. \text{Id} X Y \to \text{Id} (F X) (F Y))\).

In type theory, the type scheme \(F\) is the same as the object mapping of the functor. However, the identity mapping \((imap)\) is a restriction of the arrow map \((fmap)\), which only requires the user to lift an identity function \((\text{Id} from Section 4.1)\) between 2 types \((X and Y)\) to an identity function between the scheme \(F\) applied to the same 2 types. Deriving a concrete datatype in terms of the generic encoding of Firsov et al. [2018] takes less effort (compared to using the encoding of Firsov and Stump [2018]), because \(imap\) is less onerous to define, and no laws need to be proved.

Furthermore, the class of datatypes representable by the Firsov et al. [2018] encoding expands to include infinitary types and positive (not merely strictly-positive) types. Firsov et al. [2018] is an “efficient” lambda-encoding (using Mendler-style F-algebras, described in Section 5.4), in the sense that inductive types support a constant-time “predecessor” operation (e.g. \(\text{pred}\) for \(\text{Nat}\), and \(\text{tail}\) for \(\text{List}\)), using only linear space in the encoding. Expert readers may have noticed that the tail \(xs : \text{List}\) (where \(\text{List}\) is Church-encoded) in the cons case of \(\text{elimList}\) from Section 2.3 is erased (i.e. quantified using \(\forall\) rather than \(\Pi\)), hence computations cannot be defined with (unerased) access to the tail of the list. Deriving induction for a concrete \(\text{List}\) type encoded via
the work of Firsov et al. [2018], and using Mendler-style F-algebras, solves this problem (allowing unerased quantification over the tail via \( \Pi \), accessible in constant time).

In this work we generically solve zero-cost data reuse by defining combinators for the fixpoint type of Firsov et al. [2018], whose type is:

\[
\text{IdMapping} \triangleq (\star \to \star) \to \star = \lambda F. \forall X,Y : \star. \text{Id} X Y \to \text{Id} (F X) (F Y).
\]

\[
\text{Fix} \triangleq \Pi F : (\star \to \star) \to \Pi i : \star. \text{IdMapping} F \to \star
\]

This work derives the non-indexed fixpoint (Fix) in terms of an indexed fixpoint (IFixx), over indexed schemes and index-preserving identity mappings (IIdMapping). The non-indexed fixpoint is the trivial case where the index is the Unit type (having the single inhabitant unit). Below, we only give the type of the indexed fixpoint IFixx, and its implementation is a straightforward generalization of the non-indexed version by Firsov et al. [2018]:

\[
\text{IIdMapping} \triangleq \Pi I : \star. ((I \to \star) \to I \to \star) \to \star = \lambda I,F. \forall X,Y : \star. \text{Id} X Y \to \text{Id} (F X) (F Y).
\]

\[
\text{IFix} \triangleq \Pi I : \star. \Pi F : (I \to \star) \to I \to \star.
\]

\[
\Pi \text{imap} : \text{IIdMapping} I F. I \to \star
\]

### 5.2 Data Schemes and Identity Mappings

The examples in the remainder of this section will demonstrate how data reuse combinators reduce the problem of defining an identity function between fixpoints, to defining an identity function between schemes. This is a much simpler problem, because schemes are essentially sums-of-products, which do not have inductive arguments.

Our later examples will refer to the scheme for lists (ListF), and the scheme for vectors (VecF), whose Church-encodings appear below:

\[
\text{ListF} \triangleq \star \to \star \to \star = \lambda A,X. \forall C : \star. C \to (A \to X \to C) \to C.
\]

\[
\text{VecF} \triangleq \star \to (\text{Nat} \to \star) \to \text{Nat} \to \star = \lambda A,X,n.
\]

\[
\forall C : \text{Nat} \to \star. C \text{ zero} \to (\forall n : \text{Nat}. A \to X n \to C (\text{suc } n)) \to C n.
\]

Like Vec from Section 2.2 (the only difference is that the inductive argument is the abstract \( X n \) of the scheme, rather than an inductive \( C n \)), the cons case of VecF has the natural number as an implicit argument, so the constructors of ListF and VecF erase to the same underlying untyped terms. For reference, the type signatures for the ListF and VecF constructors and eliminators
appear in Figure 3. We assume an intersection-type encoding of \( \text{ListF} \) and \( \text{VecF} \), using the same technique as in Section 2.3, to make it possible to define the eliminators.\(^7\)

Next, we define the identity mappings \( \text{imapL} \) (for \( \text{ListF} \)) and \( \text{imapV} \) (for \( \text{VecF} \)), whose definitions only differ by which eliminator is used:

\[
\begin{align*}
\text{imapL} & : \forall A : \star. \, \text{IdMapping} \ (\text{ListF} \ A) = \lambda c. \, \text{elimListF} \ (\text{pair} \, \text{nilLF} \, \beta) 
\ (\lambda x, xs. \, \text{pair} \ (\text{consLF} \ x \ (\text{elimId} \ c \, xs) \, \beta) ). \\
\text{imapV} & : \forall A : \star. \, \text{IdIdMapping} \ (\text{Nat} \ (\text{VecF} \ A)) = \lambda c. \, \text{elimVecF} 
\ (\text{pair} \, \text{nilVF} \, \beta) 
\ (\lambda x, xs. \, \text{pair} \ (\text{consVF} \ x \ (\text{elimId} \ c \, xs) \, \beta) ). 
\end{align*}
\]

The returned value is the Sigma-type codomain of \( \text{Id} \) from Section 4.1, where the first component is the supertype and second component is the equality witness. For both \( \text{imapL} \) and \( \text{imapV} \), we mostly rebuild the term with constructors. The interesting subterm is the tail argument (\( \text{elimId} \ c \, xs \)) of the cons rebuilding (for both \( \text{consLF} \) and \( \text{consVF} \)). In the nil cases, the second component of the pair (constructing Sigma) is obviously reflexivity (\( \beta \)) when rebuilding \( \text{nilLF} \) with itself and \( \text{nilVF} \) with itself. However, in the cons cases, the second component is also \( \beta \). This is because the identity function being mapped (\( c \)) is erased when zero-cost converting (i.e. \( |\text{elimId} \ c \, xs| = xs \)). Hence, \( \beta \) is evidence of the cons rebuilding cases because \( |\text{consLF} \ x \ (\text{elimId} \ c \, xs)| = |\text{consLF} \ x \ xs \), and \( |\text{consVF} \ x \ (\text{elimId} \ c \, xs)| = |\text{consVF} \ x \ xs \).

### 5.3 Forgetful Reuse

#### 5.3.1 Data Reuse Combinator

The combinator \( \text{ifix2fix} \) is a generic solution to forgetful fixpoint reuse (or, forgetful data reuse). For example, it can solve a problem like the one below:

\[
\text{Id} \ (\text{IFix Nat} \ (\text{VecF} \ A) \, \text{imapV} \, n) \ (\text{Fix} \ (\text{ListF} \ A) \, \text{imapL})
\]

Above, the subtype is an indexed fixpoint and the supertype is a non-indexed fixpoint (hence, this the forgetful direction of data reuse). The type of \( \text{ifix2fix} \) follows:

\[
\begin{align*}
\text{ifix2fix} & : \forall I : \star. \, \forall F : (I \rightarrow \star) \rightarrow I \rightarrow \star. \, \forall G : \star \rightarrow \star. \\
& \quad \Pi \, \text{imapF} : I\text{IdMapping} \ I \ F. \\
& \quad \Pi \, \text{imapG} : I\text{IdMapping} \ G. \\
& \quad \Pi \ c : \forall X : I \rightarrow \star. \, \forall Y : \star. \\
& \quad \quad (\forall i : I. \, \text{Id} \ (X \ i) \ Y) \rightarrow \forall i : I. \, \text{Id} \ (F \ X \ i) \ (G \ Y). \\
& \quad \forall i : I. \, \text{Id} \ (\text{IFix} \ F \ \text{imapF} \ i) \ (\text{Fix} \ G \ \text{imapG})
\end{align*}
\]

If we were to solve the problem above with \( \text{ifix2fix} \), we would set index type \( I \) to \( \text{Nat} \), the indexed scheme \( X \) to \( \text{VecF} \ A \), and the non-indexed scheme \( Y \) to \( \text{ListF} \ A \). This covers all the implicit arguments of \( \text{ifix2fix} \), and now we explain the explicit arguments:

- The argument \( \text{imapF} \) is the index-preserving identity mapping for the indexed scheme \( F \), e.g. \( \text{imapV} \ A \) for \( \text{VecF} \ A \).
- The argument \( \text{imapG} \) is the identity mapping for the non-indexed scheme \( G \), e.g. \( \text{imapL} \ A \) for \( \text{ListF} \ A \).
- The argument \( c \) is the identity algebra. It forgets the indexed subtype scheme \( (F \ X \ i) \) as the non-indexed supertype scheme \( (G \ Y) \), while assuming how to forget the abstract indexed subtype \( (X) \) as the abstract non-indexed supertype \( (Y) \).

The type of \( \text{ifix2fix} \) is reminiscent of standard patterns appearing in generic programming using fixpoint encodings of datatypes. If you define a non-recursive identity function between schemes, where the "recursive" positions \( X \ i \) are abstract, and you have access to an abstract

\(^7\) There is no predecessor problem to worry about when deriving induction principles (or, eliminators) for schemes, because schemes do not contain inductive occurrences.
forgetful identity function (from $X \cdot i$ to $Y$), you are rewarded with a recursive identity function between fixpoints of those schemes.

We omit the implementation of $\text{ifix2fix}$ combinator since the exact details depend on a particular encoding of Mendler-style fixed points. Intuitively, the identity function from $\text{IFix}$ to $\text{Fix}$ is developed by using the generic dependent elimination of $\text{IFix}$ to apply the $c$ argument on each inductive level of the value. The premise of $c$, namely $(\forall i : I. \text{Id}(X \cdot i) Y)$, is the inductive hypothesis of the dependent elimination.

5.3.2 Data Reuse Example. Now we demonstrate zero-cost forgetful reuse of vector data as list data. First, we establish type synonyms for the list and vector types, derived generically as the fixpoints of their schemes and identity mappings:

\[
\text{List} \triangleleft \star \rightarrow \star \equiv \lambda A. \text{Fix} (\text{ListF} A) \text{imapL}.
\]

\[
\text{Vec} \triangleleft \star \rightarrow \text{Nat} \rightarrow \star \equiv \lambda A,n. \text{IFix Nat} (\text{VecF} A) \text{imapV} n.
\]

Next, we define an identity function ($v2l$) from $\text{Vec} A n$ to $\text{List} A$ by applying $\text{ifix2fix}$ to the identity mappings and an identity algebra. For legibility, we provide the identity algebra ($\text{vf2lf}$) as a standalone definition:

\[
v2l \triangleleft \forall A : \star. \forall n : \text{Nat}. \text{Id} (\text{VecF} A n) (\text{ListF} A Y) = \lambda c. \text{elimVecF} (\text{pair} \; \text{nilLF} \beta) (\lambda x,xs. \text{pair} (\text{consLF} x (\text{elimId} c xs) \beta)).
\]

\[
vf2lf \triangleleft \forall A : \star. \forall n : \text{Nat}. \text{Id} (\text{VecF} A n) (\text{ListF} A Y) = \text{ifix2fix} \text{imapV} \text{imapL} v2l.
\]

The identity algebra $vf2lf$ is defined by constructing an identity function, and the construction is very similar to how we defined the identity mappings $\text{imapL}$ and $\text{imapV}$ in Section 5.2. This time, the conversion changes the types (by going from indexed scheme $\text{VecF}$ to scheme $\text{ListF}$), but $\beta$ still suffices as equality in both cases because the constructors of both schemes erase to the same untyped terms. More concretely, $|\text{nilVF}| = |\text{nilLF}|$ and $|\text{consVF} \times xs| = |\text{consLF} \times (\text{elimId} c xs)|$. In the cons case, $xs$ has (abstract vector) type $X n$, but this is zero-cost converted via $c$ to (abstract list) type $Y$. Hence, because we know that $|\text{consVF}| = |\text{consLF}|$, it follows that:

\[
|\text{consVF} \times xs| = |\text{consVF}| \times xs = |\text{consLF}| \times xs = |\text{consLF} \times (\text{elimId} c xs)|
\]

5.4 Mendler-Style Algebras

In generic developments using fixpoint-encodings of datatypes, it is common to define non-dependent functions as the fold of an algebra. Our generic enriching data reuse combinator (in Section 5.5) requires an algebra argument (which is folded in the dependent type signature of the combinator). However, because our fixpoint type is defined using a Mendler-style encoding [Firsov et al. 2018], our enriching combinator must take a Mendler-style algebra. Below, we give the definition of a Mendler-style algebra ($\text{AlgM}$), and we include the more familiar Church-style algebra ($\text{AlgC}$) for reference:

\[
\text{AlgC} \triangleleft (\star \rightarrow \star) \rightarrow \star \rightarrow \star \equiv \lambda F,X. F X \rightarrow X.
\]

\[
\text{AlgM} \triangleleft (\star \rightarrow \star) \rightarrow \star \rightarrow \star \equiv \lambda F,X. \forall R : \star. \Pi \text{rec} : R \rightarrow X. F R \rightarrow X.
\]

Mendler algebras ($\text{AlgM}$) exploit parametricity to abstractly hide inductive data via impredicative quantification ($\forall R : \star$). However, a recursion function ($\Pi \text{rec} : R \rightarrow X$) is provided to explicitly make recursive calls on the hidden data.\(^8\)

\(^8\) Mendler-style data hiding and explicit recursion is one of the ingredients used by [Firsov et al. 2018] to define constant-time predecessor functions.
Below, we give an example of defining the list length function \(\text{len}\) as the fold of a Mendler-style length algebra \(\text{lenAlgM}\). We also provide the type of the Mendler-style \(\text{foldM}\) function for reference.

\[
\text{lenAlgM} \\
\forall X : \star. \text{AlgM} (\text{ListF} X) \text{Nat} \\
= \lambda \text{rec. elimListF zero (}\lambda x, x \cdot \text{suc (}\text{rec } xs))
\]

\[
\text{len} \\
\forall A : \star. \text{List} A \rightarrow \text{Nat} = \text{foldM} \text{lenAlgM}.
\]

The length algebra \(\text{lenAlgM}\) case-splits (using \(\text{elimListF}\)) on the scheme \(\text{ListF}\) of the generically encoded list. The nil case returns \(\text{zero}\), and the cons case returns the \(\text{suc}\) (sessor) of the result of applying the recursion function \(\text{rec}\) to the abstract recursive data \(x \cdot \text{r}\). Our example in Section 5.5 uses both the length algebra \(\text{lenAlgM}\) and the length function \(\text{len}\) defined as its fold.

### 5.5 Enriching Reuse

The combinator we define in this section \(\text{fix2ifix}\) generically solves data enrichment, going from a non-indexed to an indexed type, when the index can be computed as a total function from the non-indexed type (e.g. going from \(\text{List}\) to \(\text{Vec}\) via the total function \(\text{len : List} A \rightarrow \text{Nat}\)).

#### 5.5.1 Data Reuse Combinator

Next, we define the combinator \(\text{fix2ifix}\), which is a generic solution to enriching \text{fixpoint} reuse (or, enriching \text{data} reuse). For example, it can solve a problem like the one below:

\[
\text{IdDep } (\text{Fix} (\text{ListF} A) \text{imapL}) \\
(\lambda x. \text{IFix} \text{Nat} (\text{VecF} A) \text{imapV} (\text{foldM} \text{lenAlgM} x))
\]

Notice that \(\text{fix2ifix}\) must return a \text{dependent} identity function, because the index of the output vector is computed as the length \(\text{len}\) of the input list \(x\). The type of \(\text{fix2ifix}\) follows:

\[
\text{fix2ifix} \\
\forall G : \star \rightarrow \star. \forall I : \star. \forall F : (I \rightarrow \star) \rightarrow I \rightarrow \star.
\]

Both \(\text{fix2ifix}\) and \(\text{ifix2fix}\) (from Section 5.3) share the same implicit arguments, namely \(I\), \(F\), and \(G\), and they also share the explicit \text{imapF} and \text{imapG} arguments. However, \(\text{fix2ifix}\) has the following differing explicit arguments:

- The argument \(\text{ralg}\) is the \text{refinement algebra} for the non-indexed scheme \(G\), e.g. \(\text{lenAlgM} A\) for \(\text{ListF} A\).

- The argument \(c\) is the \text{dependent identity algebra}. It enriches the non-indexed subtype scheme \((y : G Y)\) to the indexed supertype scheme \((F X (\text{ralg} r y))\), while assuming how to enrich the abstract non-indexed subtype \((y : Y)\) as the abstract indexed supertype \((X (r y))\).

Similar to the \(c\) of forgetful \(\text{ifix2fix}\), the \(c\) of enriching \(\text{fix2ifix}\) requires a \text{non-recursive} identity function between schemes, while assuming access to an identity function between abstract “recursive” positions. However, the identity function in \(c\) for \(\text{fix2ifix}\) are \text{dependent}. Hence, the index of the assumed supertype \((X)\) is computed from the non-indexed subtype \((y : Y)\) by applying an abstract \text{refinement function} \(r\). Correspondingly, the index of the produced supertype \((F X)\) is
computed from the non-indexed subtype \((ys : G Y)\) by applying the refinement algebra \((ralg \ r)\), while using \(r\) for the rec(ursive) function of the Mendler-style algebra.

The implementation of \(fix2ifix\), just like \(ifix2fix\), applies \(c\) to each inductive level. The outcome is also similar, as \(fix2ifix\) allows the user to define a non-recursive identity algebra, and it produces a recursive identity function between fixpoints. The primary difference is that \(fix2ifix\) results in a dependent identity function. Hence, the index in the dependent result is computed by folding the Mendler-style algebra \((ralg)\) over the inductive input \(x\).

5.5.2 Data Reuse Example. Now we demonstrate zero-cost enriching reuse of list data as vector data. The dependent identity function \((l2v)\) from \(xs : List A\) to \(Vec A (len \ xs)\) is defined by applying \(fix2ifix\) to the identity mappings and the dependent identity algebra \(lf2vf\):

\[
lf2vf \triangleq \forall A : \star. \forall Y : \star. \forall X : Nat \rightarrow \star.
\Pi r : Y \rightarrow Nat.
\Pi c : IdDep Y (\lambda y. X (r y)).
IdDep (ListF A Y ) (\lambda xs. VecF A X (lenAlgM r xs))
= \lambda c. elimListF (pair nilVF \beta)
(\lambda x,xs. pair (consVF x (elimIdDep c xs) \beta)).
\]

\[
l2v \triangleq \forall A : \star. IdDep (List A) (\lambda xs. Vec A (len \ xs))
= fix2ifix imapL imapV lenAlgM lf2vf.
\]

The Mendler-style algebra used by \(lf2vf\) is our previously defined length algebra \((lenAlgM)\). The definition of \(lf2vf\) is essentially the same as \(vf2lf\) from Section 5.3, but now we eliminate a list scheme and produce vector scheme constructors. Because the vector and list scheme constructors erase to the same terms, the argument for why reflexivity (\(\beta\)) suffices as identity evidence stays the same. Another difference is that the abstract tail is computed as a dependent elimination \((elimIdDep\ c c\ xs)\), rather than \(elimId\). However, the dependent elimination is also erased \(|elimIdDep\ c c\ xs| = xs\).

6 GENERIC RELATIONAL REUSE

In this section we demonstrate that the techniques of our paper scale to more complex dependently typed programs, rather than the pedagogical list and vector running example used thus far. A common example (e.g. as used by Norell [2008]) of dependently typed programming, due to McBride and McKinna [2004], is writing a correct-by-construction \(infer\) function from unchecked STLC terms \((Raw)\) to checked STLC terms \((Term)\), where \(Raw\) is an unindexed type and \(Term\) is a type indexed by the input context \((Ctx)\) and output type \((Tp)\) of \(infer\). The examples of this section will use the STLC datatypes from this problem domain:

\[
Raw \triangleq \star = Fix RawF imapR.
\]

\[
var \triangleq \forall n : Nat. in (varF n).
\]

\[
lam \triangleq \forall A,B : Tp. in (lamF A b).
\]

\[
app \triangleq \forall f,a : Raw. in (appF f a).
\]

\[
Term \triangleq CtxTp \rightarrow \star = IFix CtxTp TermF imapT.
\]

\[
ivar \triangleq \forall G : Ctx. \forall A : Tp. Mem Tp A G \rightarrow Term (pair G A)
= \lambda i. iin (ivarF i).
\]

\[
ilam \triangleq \forall G : Ctx. \Pi A : Tp. \forall B : Tp. Term (pair (ext G A) B) \rightarrow Term (pair G (Arr A B)) = \lambda A,b. iin (ilamF A b).
\]

\[
iapp \triangleq \forall G : Ctx. \forall A,B : Tp. Term (pair G (Arr A B)) \rightarrow Term (pair G A) \rightarrow Term (pair G B) = \lambda f,a. iin (iappF f a).
\]
Generic Zero-Cost Reuse for Dependent Types

The Raw terms use de Bruijn indexing, hence the variable constructor (\texttt{var}) takes a natural number (\texttt{Nat}). The lambda terms are annotated so that type inference is possible, as seen in the \texttt{lam} and \texttt{ilam} constructors, which use the unerased $\Pi$ quantifier for the domain type $A$. The checked Term’s are indexed by \texttt{CtxTp}, which is simply the non-dependent pair (Prod, derived from Sigma) of a context (Ctx, which is a List of types $Tp$) and a type ($Tp$, which can be a base type Base, or a function type $Arr$). The indexed \texttt{ivar} constructor contains a membership proof (\texttt{Mem}, a standard list membership relation), ensuring that the type $A$ appears in the context $G$. Finally, the underlying schemes (\texttt{RawF} and \texttt{TermF}), their identity mappings (\texttt{imapR} and \texttt{imapT}), their constructors (e.g. \texttt{lamF} and \texttt{ilamF}), and their eliminators (\texttt{elimRawF} and \texttt{elimTermF}) are defined as in Section 5.2, without any surprises, and have been omitted for space reasons.

\section{Forgetful Data Reuse}

Forgetful data reuse from Term to Raw uses ifix2fix (from Section 5.3) to forget TermF as RawF via the helper function tf2rf.

\begin{align*}
\text{tf2rf} &
\equiv 
\forall X : \text{CtxTp} \rightarrow \star. \forall Y : \star. \\
&\quad \Pi c : \forall GA : \text{CtxTp}. \text{Id} (X GA) Y. \\
&\quad \forall GA : \text{CtxTp}. \text{Id} (\text{TermF} X GA) (\text{RawF} Y) \\
&\quad = \lambda c. \text{elimTermF} \\
&\quad (\lambda i. \text{pair} (\text{varF} (\text{elimId} i2n i)) \beta) \\
&\quad (\lambda b. \text{pair} (\text{lamF} A (\text{elimId} c b)) \beta) \\
&\quad (\lambda f,a. \text{pair} (\text{appF} (\text{elimId} c f) (\text{elimId} c a)) \beta).
\end{align*}

\begin{align*}
\text{t2r} &
\equiv 
\forall GA : \text{CtxTp}. \text{Id} (\text{Term} GA) \text{Raw} \\
&\quad = \text{ifix2fix imapT imapR tf2rf}.
\end{align*}

The lambda and application cases use the zero-cost conversion $c$ to translate each $X GA$ to a $Y$. The variable case uses the omitted forgetful data reuse $i2n$ of type $\text{Id} (\text{Mem} A x xs) \text{Nat}$, which similarly forgets membership proofs as natural numbers.

\section{Enriching Data Reuse}

The enriching data reuse combinator ifix2fix (from Section 5.5) can be used when the index of the enriched type can be computed as a total function (via the refinement algebra) from the non-indexed type. When enriching a Raw term to a Term, we are faced with the problem below:

\begin{align*}
\text{IdDep} (\text{Fix} \text{RawF} \text{imapR}) \\
(\lambda t. \forall GA : \text{CtxTp}. \text{Typed} GA t \Rightarrow \text{IFix} \text{CtxTp} \text{TermF} \text{imapT})
\end{align*}

\subsection{Data Reuse Combinator}

Above, \text{Typed} $GA$ $t$ is an erased premise necessary for the enrichment to be possible. To solve such a problem, we define the generalized $\text{fix2ifixP}$ combinator, which allows for data enrichment with a \textit{premise}:

\begin{align*}
\text{fix2ifixP} &
\equiv 
\forall G : \star \rightarrow \star. \forall I : \star. \forall F : (I \rightarrow \star) \rightarrow I \rightarrow \star. \\
&\quad \Pi \text{imapG} : \text{IdMapping} G. \Pi \text{imapF} : \text{IIDMapping} I F. \\
&\quad \forall P : I \rightarrow \text{FixIndM} G \text{imapG} \rightarrow \star. \\
&\quad \Pi c : \forall Y : \star. \forall X : I \rightarrow \star. \\
&\quad \Pi c1 : \text{Id} Y (\text{FixIndM} G \text{imapG}). \\
&\quad \Pi c2 : \text{IdDep} Y (\lambda y. \forall i : I. \text{P} i (\text{elimId} c1 y) \Rightarrow X i). \\
&\quad \text{IdDep} (G Y) (\lambda ys. \forall i : I. \text{P} i (\text{in} (\text{elimId} (\text{imapG} c1) ys)) \Rightarrow F X i). \\
&\quad \text{IdDep} (\text{FixIndM} G \text{imapG}) (\lambda y. \forall i : I. \text{P} i y \Rightarrow \text{IFix} I F \text{imapF} i)
\end{align*}
Besides the implicit arguments that fix2ifixP shares with fix2ifix, the premise \( P \) is an additional implicit argument, which may depend on the index (also implicitly quantified as part of the premise) and the non-indexed data.

The explicit refinement algebra \( r \) no longer appears. The identity algebra \( c \) generalizes to enrich the non-indexed subtype scheme (\( Ys : G Y \)) to the indexed supertype scheme (\( FXi \)), once provided the index (\( i : I \)) and the premise (of type \( P iYs \)) as erased arguments. These arguments must be erased for the zero-cost conversion to indeed be an identity function from non-indexed to indexed data. As expected in \( c \), \( c2 \) is used to convert a non-indexed inductive occurrence (\( Y \)) to an indexed version (\( Xi \)), under an erased premise (\( P i(\text{elimId}c1y) \)). However, to type the premise we must convert the abstract \( Y \) to a concrete non-indexed fixpoint (\( \text{FixIndM}G \text{imapG} \)), and for this we have the additional \( c1 \) argument.

### 6.2.2 Data Reuse Example

To enrich a Raw term as a typed Term, under a well-typedness premise (Typing), below we apply our new combinator fix2ifixP to the helper enriching dependent identity algebra rf2tfP. In many places of this paper we omit implicit arguments, as there is a nice coincidence between when arguments are erased and inferrable. In the example below, rf2tfP is much larger than previous examples because the erased index (\( GA : \text{CtxTp} \)) and the erased premise (of type \( \text{Typed} \ GA \ (\text{elimId}c1y) \)) are not inferrable. As previously, we use dash (\( - \)) to explicitly supply implicit arguments, but now we use the new syntax of a variable name along with an equal sign to only supply relevant (non-inferrable) implicit arguments (e.g. for explicitly supplying an index and a premise, below we use \( GA = \bullet \) and \( e = \bullet \), respectively).

\[
\text{rf2tfP} \ ▶ ∀ Y : \bullet. ∀ X : \text{CtxTp} → ☐.
\]

\[
\begin{align*}
\Pi c1 & : \text{Id} \ Y \ \text{Raw}. \\
\Pi c2 & : \text{IdDep} \ Y (\lambda y. \forall GA : \text{CtxTp}. \text{Typed} \ GA \ (\text{elimId}c1y) ⇒ XGA). \\
\text{IdDep} \ (\text{RawF} \ Y) (\lambda ys. \forall GA : \text{CtxTp}.
\text{Typed} \ GA \ (\text{in} \ (\text{imapRaw} \ c1 \ ys)) ⇒ \text{TermF} \ XGA)
= \lambda c1,c2. \text{elimRawF}
\end{align*}
\]

\[
\begin{align*}
(\lambda n. \text{pair} \ (\Lambda GA,p. \rho \ (\text{etaSigma} \ GA) - \text{ivarF}) \ (\text{elimId}n2ipn -\left(\begin{array}{l}
(\text{proj1} \ GA) \\
(\text{proj2} \ GA)
\end{array}\right)n(\rho\varsigma(\text{etaSigma} \ GA) - p))))  β)
\end{align*}
\]

\[
\begin{align*}
(\lambda A,y. \text{pair} \ (\Lambda GC,p. \rho \ (\text{etaSigma} \ GC) - \\
(\text{invLamEq} \ GC \ A \ (\text{elimId}c1y) p) - \text{ilamF})
\end{align*}
\]

\[
\begin{align*}
A -\left(\begin{array}{l}
(\text{elimIdDep} \ c2 \ y -\left(\begin{array}{l}
(\text{elimIdCod} \ GC \ A \ (\text{elimId}c1y) p)
\end{array}\right)\right) \\
(\text{elimIdDep} \ c2 \ y -\left(\begin{array}{l}
(\text{elimIdCod} \ GC \ A \ (\text{elimId}c1y) p)
\end{array}\right)\right) β)
\end{align*}
\]

\[
\begin{align*}
(\lambda y1,y2. \text{pair} \ (\Lambda GB,p. \rho \ (\text{etaSigma} \ GB) - \text{iappF}) \\
(\text{invAppDom} \ GB \ (\text{elimId}c1y1) \ (\text{elimId}c1y2) p) -\left(\begin{array}{l}
(\text{proj2} \ GB)
\end{array}\right)
\end{align*}
\]

\[
\begin{align*}
(\text{elimIdDep} \ c2 \ y -\left(\begin{array}{l}
(\text{elimIdCod} \ GB \ (\text{elimId}c1y1) \ (\text{elimId}c1y2) p)
\end{array}\right)\right) β)
\end{align*}
\]

\[
\begin{align*}
\text{rf2tfP} \ ▶ \text{IdDep} \ \text{Raw} \ (\lambda t. \forall GA : \text{CtxTp}. \text{Typed} \ GA \ t ⇒ \text{Term} \ GA)
= \text{fix2ifixP} \ \text{TermF} \ \text{imapR} \ \text{imapT} \ \text{rf2tfP}.
\end{align*}
\]
Generic Zero-Cost Reuse for Dependent Types

(Typed ◁ CtxTp → Raw → * )
(invVarLookup ◁ II G : Ctx. II A : Tp. II n : Nat. Typed (pair G A) (var n) → Lookup A G n)
(invLamCod ◁ II GC : CtxTp. II A : Tp. II b : Raw. Typed GC (lam A b) → Tp)
(invLamEq ◁ II GC : CtxTp. II A : Tp. II b : Raw. II p : Typed GC (lam A b). proj2 GC ≃ Arr A (invLamCod GC A b p))
(invLamBod ◁ II GC : CtxTp. II A : Tp. II b : Raw. II p : Typed GC (lam A b). Typed (pair (consL A (proj1 GC)) (invLamCod GC A b p)) b)
(invAppDom ◁ II GB : CtxTp. II f : Raw. II a : Raw. Typed GB (app f a) → Tp)
(invAppFun ◁ II GB : CtxTp. II f : Raw. II a : Raw. II p : Typed GB (app f a).
Typed (pair (proj1 GB) (Arr (invAppDom GB f a p) (proj2 GB))) f)
(invAppArg ◁ II GB : CtxTp. II f : Raw. II a : Raw. II p : Typed GB (app f a).
Typed (pair (proj1 GB) (invAppDom GB f a p) a))
(termTyped ◁ II GA : CtxTp. II t : Term GA. Typed GA (elimId t2r t))

Fig. 4. An abstract typing relation, Typed, specified in terms of its inversion lemmas, and used as the premise of STLC term enrichment.

elimRawF, using c2 to enrich abstract inductive Raw terms (y : Y) to their typed Term equivalents (X GA), once given a suitable index (GA : CtxTp) and premise Typed GA (elimId c1 y).

For the lambda (ilamF) and application (iappF) cases, the index (GA : CtxTp) and Typing premise are additional arguments we must supply to get an identity conversion back from c2, which functions as an inductive hypothesis for the purpose of this combinator. We explicitly supply both of these implicit arguments by appealing to inversion lemmas from our abstractly specified Typing premise in Figure 4. For example, in the ilamF case we get the lambda codomain via invLamCod and the inductive Typed premise of the lambda body via invLamBod.

Note that any implementation of Typed works for our example, so long as the inversion lemmas are provable. Obvious choices are an indexed datatype version of Typed (i.e. a binary relation), or Typed GA t as infer (proj1 GA) t ≃ just (proj2 GA), making the premise state that type inference succeeds. In our Cedille formalization, we use a third appealing “free” definition, exploiting Cedille’s heterogeneous equality, namely Typed GA t as Sigma (Term GA) (λ t ’. t ’ ≃ t).

Finally, the variable case (ivarF) appeals to the omitted enrichment n2iP from natural numbers to membership proofs (IdDep Nat (λ n. ∀ xs : List A. Lookup A x xs n ⇒ Mem A x xs)). The premise is an abstract Lookup relation, of type II A : *. A → List A → Nat → *, stating that the element appears at the natural number position of the (zero-indexed) list, where Lookup is also abstractly specified in terms of its obvious inversion lemmas. The definition of n2iP is similar to r2tP, but simpler because it only has two cases and they have less-complicated arguments.

6.3 Enriching Program Reuse

For our example, we show how to enrich a one-step β-reduction function (stepR of type StepR) between Raw terms to one between typed Term’s (of type StepT), provided the premise that the raw step function preserves types (TpPres).

StepR ◁ * → Raw → Raw.
TpPres ◁ StepR → * = λ stepR.
II t : Raw. II GA : CtxTp. Typed GA t → Typed GA (stepR t).
StepT ◁ * = ∀ GA : CtxTp. Term GA → Term GA.
6.3.1 Data Reuse Combinator. Recall from Section 4.3 that the enriching program reuse combinator $\text{arr2allArrP}$ takes the total refinement function $r$ as an argument, in addition to the index preservation property argument $c_1'$, which is used to automatically perform a rewrite by the property. Below we define $\text{arr2allArrP2}$, which is a version that works with data that must be reused with a premise.

$$\text{arr2allArrP2} \triangleq \forall Y,Y',I : \star. \forall P : I \rightarrow Y \rightarrow \star. \forall P' : I \rightarrow Y' \rightarrow \star. \forall X,X' : I. \mathrm{Id} (X i) Y.$$ 

Notice that the premise has been broken into two pieces (compared to $\text{arr2allArrP}$), where $P$ now is the premise on the non-indexed data (of type $Y$), and $P'$ is a premise on the output of the function being enriched (as before). Now $c_1'$ is a proof that the indexed data ($x : X i$) implies that the premise holds for its forgetful variant ($P i (\text{elimId} c_1 x)$). The automation performed by $\text{arr2allArrP2}$ is automatically supplying this evidence to the $P$ argument of the erased premise. Also notice that in the erased premise, the user may assume a non-erased (explicitly quantified) index argument ($\Pi i : I$), because the entire functional premise is erased anyway.

We omit the obvious dependent version ($\text{pi2allPiP2}$). For space reasons, we also omit the definition of $\text{arr2allArrP2}$, which is an obvious alteration of $\text{arr2allArrP}$. In fact, in our implementation both $\text{arr2allArrP}$ and $\text{arr2allArrP2}$ are defined in terms of a more general version that does not perform any automation (i.e. a combinator without a $c_1'$ argument).

6.3.2 Data Reuse Example. We enrich the type-preserving one-step $\beta$-reduction function by applying $\text{arr2allArrP2}$ once for the single argument of the unary function:

$$\text{stepR2stepT} \triangleq \mathrm{IdDep} \text{StepR} (\lambda f. \text{TpPres} f \Rightarrow \text{StepT})$$

After applying $\text{arr2allArrP2}$, the goal has the erased index part of the argument ($\forall GA : \text{CtxTp}$) on the outside of the dependent identity function type. We solve the goal by applying the auxiliary combinator $\text{supplyPrem}$ (from Figure 2) to our enriching data reuse result for terms ($\text{r2tP}$ from Section 6.2), which expects $\forall GA : \text{CtxTp}$ as part of the supertype premise.

Our work scales to similar but more complicated program enrichments. See our formalization for an example of enriching a substitution function for Raw terms to a version for typed Term's, under the premise that substitution is type-preserving. This requires more zero-cost conversions, to go between a List of Raw terms and a type-safe environment of Term's, encoded as the All type, representing that all elements of a list satisfy some predicate.

7 RELATED WORK

7.1 Subtyping

Miquel [2001] shows that in a Curry-style type theory with implicit products, the subtyping judgement can be derived as follows:

$$\Gamma \vdash X \subseteq Y \iff \Gamma, x : X \vdash x : Y$$
The \textit{Id} type can be seen as the internalization of this judgement, with \textit{IdDep} corresponding to a dependent version (i.e. our informal syntax \((x : X) \leq Y x\), not covered by Miquel). Miquel also showed that the subsumption rule of subtyping is admissible in the theory with the derived judgement, and our elimination rule \(\text{elimId}\) corresponds to its internalization. Finally, all of our combinators can also be translated to admissible subtyping rules in his theory. Miquel covers several admissible subtyping rules, but not ones corresponding to our primary forgetful and enriching combinators for program, proof, and data reuse. Our data reuse combinators may be of particular interest to the subtyping community, as they corresponds to Mendler-style datatype-generic subtyping rules.

Inspired by the internalized subtyping judgement of Miquel, Barras and Bernardo [2008] show how to derive zero-cost forgetful data reuse conversions for Church-encoded datatypes. This work was extended by Diehl and Stump [2018] to the enriching direction. In Section 3.2, we derive zero-cost data reuse in terms of a linear-time conversion and its extensional identity proof, using \(\phi\). In contrast, the zero-cost conversions of Barras and Bernardo [2008] and Diehl and Stump [2018] require no extensional identity proof, as the conversions erase to the identity function by a clever exploitation of \(\eta\)-equality, without needing a rule like \(\phi\). Our work can be seen as the generic version of their manual zero-cost reuse. When working generically with abstract combinator definitions, an abstraction like \textit{IdDep} is necessary, and hence also a rule like \(\phi\) (used to eliminate it).

### 7.2 Coercible in Haskell

Breitner et al. describe a GHC extension to Haskell (available starting with GHC 7.8) for a type class \textit{Coercible} \(a \ b\), which allows casting from \(a\) to \(b\) when such a cast is indeed the identity function [Breitner et al. 2016]. The motivation is to support retyping of data defined using Haskell’s newtype statement, which is designed to give programmers the power to erect abstraction barriers that cannot be crossed outside of the module defining the newtype. Within such a module, however, \textit{Coercible} \(a \ b\) and the associated function \textit{coerce} : \(a \rightarrow b\) allow programmers to apply zero-cost casts to change between a newtype and its definition.

\textit{Coercible} had to be added as primitive to GHC, along with a rather complex system of \textit{roles} specifying how coercibility of application of type constructors follows from coercibility of arguments to those constructors. In contrast, in the present work, we have shown how to derive zero-cost coercions within the existing type theory of Cedille (via \textit{IdDep}, also derived in Cedille, which is the dependent equivalent of \textit{Coercible}). On the other hand, much of the complexity of \textit{Coercible} in GHC arises from (1) how it interoperates with programmer-specified abstraction (via \textit{newtype}) and (2) the need to resolve \textit{Coercible} \(a \ b\) class constraints automatically, similarly to other class constraints in Haskell. The present work does not address either issue. However, the present work does allow for dependent casts to indexed variants of datatypes, which \textit{Coercible} does not cover (because \textit{Coercible} is only equivalent to our non-dependent \textit{Id}).

### 7.3 Dependent Interoperability

The field of dependent interoperability is concerned with reusing code between non-dependent and dependent implementations of datatypes and functions. The goal is to support interaction between non-dependent and dependent languages, like extracted OCaml and Coq. The most similar work to ours in this field is that of Dagand et al. [2016]. Inspired by Homotopy Type Theory (HoTT), Dagand et al. [2016] formalize partial equivalence types, simultaneously representing the forgetful and enriching directions of reuse.

They also develop combinators that are closed with respect to their partial equivalence type. For example, their \textit{HODepEquiv} combinator is quite like our forgetful program reuse combinator.
a11Arr2arr. However, their work primarily focuses on the forgetful direction of reuse for total functions, as partial functions can be reused by inserting dynamic checks and failures using their partial equivalence type. In contrast, we emphasize the total reuse of functions in the enriching direction (like arr2allArrP), using premises to make the total functions possible. Because they are primarily interested in program reuse, not proof reuse, they do not provide dependent versions of their reuse combinators. Additionally, they only provide combinators for function types, not fixpoint types, as their work assumes manual solutions to the problem of data reuse.

The class of datatypes reusable in their setting is larger, because isomorphic datatypes, with different representations, can be related. In contrast, our work requires the erasures of the constructors of related types to be the same untyped terms. However, for the price of a smaller class of reusable types, we gain the ability to perform conversions at zero-cost.

Note that dependent interoperability [Dagand et al. 2018] and zero-cost conversions solve related, but different, problems. Consider the headV function for vectors, headV : ∀ A : ⋆. ∀ n : Nat. Vec A (suc n) → A. Dependent interoperability could use this to define a partial head function for lists, whereas we cannot perform forgetful function when the domain is partial. At best, it would be easy to define a forgetful function reuse combinator with a premise, that would allow us to reuse headV to define headL : ∀ A : ⋆. ∀ n : Nat. Π xs : List A. NonEmpty xs ⇒ A (this example would also use our data reuse combinator with a premise, fix2ifixP, from Section 6.2).

Finally, Dagand et al. [2016] automates the assembly of combinators to reuse programs by registering them as instances of Coq’s type class mechanism. Cedille does not currently have type classes, but we could employ the same automation strategy if type classes get added to Cedille in the future.

7.4 Ornaments

Ornaments [McBride 2011] are used to define refined version of types (e.g. Vec) from unrefined types (e.g. List) by “ornamenting” the unrefined type with extra index information. In contrast, our work establishes a relationship between Vec and List after-the-fact, by defining forgetful and enriching IdDep values between the types. By defining vectors as natural-number-ornamented lists, ornaments can be used to calculate the “patch” type necessary to adapt a function from one type to another type [Dagand and McBride 2012]. For example, ornaments could calculate that LenDistAppL is the premise necessary to adapt appL from lists to vectors (appV).

Although ornaments can be used to derive conversions between types in an ornamental relationship [Ko and Gibbons 2013; McBride 2011], they take linear time, rather than constant time (i.e. the conversions are not zero-cost). Besides refining the indices of existing datatypes, ornaments also allow data to be added to existing datatypes. For example, vectors can be index-refined lists, but lists can also be natural numbers with elements added. Our work only covers the index refinement aspect of ornaments. It would be interesting to explore adapting a restriction of ornaments, where only erased data can be added, to deriving zero-cost coercions.

7.5 Type Theory in Color

Type Theory in Color (TTC) [Bernardy and Guilhem 2013] generalizes the concept of erased arguments of types to various colors, which may be erased optionally and independently according to modalities in the type theory. In the vector datatype declaration, the index data can be colored. If a vector is passed to a function expecting a list (whose modality enforces the lack of the index data color), then a forgetful zero-cost conversion (using our parlance) is performed.
Lists can also be used as vectors, via an enriching zero-cost conversion in the other direction. This works due to a mechanism to interpret lists as a predicate on natural numbers. The list predicate is generated as the erasure of its colored elements (like ornaments, colors can add data in addition to refining indices), which results in refining lists by the length function.

Our work can be used to define an enriching zero-cost conversion from natural numbers to the datatype of finite sets (Fin). This is not possible with colors, because Fin is indexed by successor (suc) in both of its constructors, which would require generating a predicate on the natural numbers from a non-deterministic function (or relation). Colors allow zero-cost conversions to be generated and implicitly applied because colors erase types, as well as values, whereas implicit products only erase values (e.g. Λ is erased, but not ∀). Thus, while zero-cost conversions need to be explicitly crafted and applied in our setting, we are able to define zero-cost conversions (like taking natural numbers to finite sets) for which there is no unique solution.

8 EXTENSIONS AND FUTURE WORK

8.1 Auxiliary Combinators

Our program and proof reuse combinators expect index arguments to appear next to their indexed types in type signatures. For this reason, our combinators would not be directly applicable if we wrote the type signature of vector append with the natural number indices of both vector arguments at the beginning, followed by both vectors. However, it is straightforward to define an auxiliary combinator that flips argument order, which Dagand et al. [2016] do for their partial type equivalence abstraction.

If a subsequent indexed type argument depends on the same index as a previous argument, rather than a new one, it also prevents our combinators from being applicable. Consider the artificial example of vector append where both input vectors must be the same length. This can be solved via a straightforward auxiliary combinator that introduces a new index quantification, along with an equality that constrains the new index to equal the old index.

8.2 Differently Indexed Combinators

In this paper we only considered relating non-indexed types (e.g. list) to indexed types (e.g. vector). In general, we may want to relate an indexed type to a less indexed type, like relating vectors to ordered vectors in the introduction. Our combinators straightforwardly generalize to two indexed types, X : I → * and Y : J → *, along with a function that translates the more refined index to the less refined index (of type I → J). Our work could then be used in the common scenario where data remains the same but only the index changes, e.g. zero-cost converting a list of elements less than \( n \), to a list of elements less than \( n + m \).

9 CONCLUSION

We have demonstrated how to reuse programs, proofs, and types at zero-cost, in both the forgetful and enriching directions. We achieve this generically via combinators over the type of dependent identity functions (IdDep). Because partially applying the elimination rule of IdDep results in the term erasing to an identity function, any conversion making use of the result of \( e \) has no runtime overhead.

ACKNOWLEDGMENTS

We gratefully acknowledge feedback from anonymous reviewers, NSF support under award 1524519, and DoD support under award FA9550-16-1-0082 (MURI program).
REFERENCES

Bruno Barras and Bruno Bernardo. 2008. The implicit calculus of constructions as a programming language with dependent types. Foundations of Software Science and Computational Structures (2008), 365–379.

Jean-Philippe Bernardy and Moulin Guilhem. 2013. Type-theory in Color. In Proceedings of the 18th ACM SIGPLAN International Conference on Functional Programming (ICFP ’13). ACM, New York, NY, USA, 61–72. https://doi.org/10.1145/2500365.2500577

Edwin Brady. 2013. Idris, a general-purpose dependently typed programming language: Design and implementation. Journal of Functional Programming 23, 05 (2013), 552–593.

Joachim Breitner, Richard A. Eisenberg, Simon Peyton Jones, and Stephanie Weirich. 2016. Safe zero-cost coercions for Haskell. J. Funct. Program. 26 (2016), e15.

Pierre-Evariste Dagand and Conor McBride. 2012. Transporting Functions Across Ornaments. In Proceedings of the 17th ACM SIGPLAN International Conference on Functional Programming (ICFP ’12). ACM, New York, NY, USA, 103–114. https://doi.org/10.1145/2364527.2364544

Pierre-Evariste Dagand, Nicolas Tabareau, and Éric Tanter. 2016. Partial Type Equivalences for Verified Dependent Inter-operability. In Proceedings of the 21st ACM SIGPLAN International Conference on Functional Programming (ICFP 2016). ACM, New York, NY, USA, 298–310. https://doi.org/10.1145/2951913.2951933

Pierre-Evariste Dagand, Nicolas Tabareau, and Éric Tanter. 2018. Foundations of dependent interoperability. Journal of Functional Programming 28 (2018).

Leonardo de Moura, Soonho Kong, Jeremy Avigad, Floris Van Doorn, and Jakob von Raumer. 2015. The Lean theorem prover (system description). In International Conference on Automated Deduction. Springer, 378–388.

Larry Diehl and Aaron Stump. 2018. Zero-Cost Coercions for Program and Proof Reuse. (2018). arXiv:1802.00787

Denis Firsov, Richard Blair, and Aaron Stump. 2018. Efficient Mendler-Style Lambda-Encodings in Cedille. In Interactive Theorem Proving - 9th International Conference, ITP 2018, Held as Part of the Federated Logic Conference, FloC 2018, Oxford, UK, July 9-12, 2018, Proceedings. 235–252. https://doi.org/10.1007/978-3-319-94821-8_14

Denis Firsov and Aaron Stump. 2018. Generic Derivation of Induction for Impredicative Encodings in Cedille. In Certified Programs and Proofs (CPP), June Andronick and Amy Felty (Eds.).

Matthew Flatt and PLT. 2010. Reference: Racket. Technical Report PLT-TR-2010-1. PLT Design Inc. https://racket-lang.org/tr1/.

Herman Geuvers. 2001. Induction Is Not Derivable in Second Order Dependent Type Theory. In Typed Lambda Calculi and Applications (TLCA). 166–181.

Hsiang-Shang Ko and Jeremy Gibbons. 2013. Relational algebraic ornaments. In Proceedings of the 2013 ACM SIGPLAN workshop on Dependently-typed programming. ACM, 37–48.

Alexei Kopylov. 2003. Dependent Intersection: A New Way of Defining Records in Type Theory. In 18th IEEE Symposium on Logic in Computer Science (LICS). 86–95.

Conor McBride. 2011. Ornamental algebras, algebraic ornaments. (2011).

Conor McBride and James McKinna. 2004. The view from the left. Journal of functional programming 14, 1 (2004), 69–111.

Alexandre Miquel. 2001. The Implicit Calculus of Constructions Extending Pure Type Systems with an Intersection Type Binder and Subtyping. In Typed Lambda Calculi and Applications (TLCA), Samson Abramsky (Ed.). 344–359.

Ulf Norell. 2007. Towards a practical programming language based on dependent type theory. Ph.D. Dissertation. Chalmers University of Technology.

Ulf Norell. 2008. Dependently typed programming in Agda. In International School on Advanced Functional Programming. Springer, 230–266.

Aaron Stump. 2017. The calculus of dependent lambda eliminations. J. Funct. Program. 27 (2017), e14.

Aaron Stump. 2018a. From Realizability to Induction via Dependent Intersection. Ann. Pure Appl. Logic (2018). to appear.

Aaron Stump. 2018b. Syntax and Semantics of Cedille. (2018). arXiv:1806.04709

The Coq Development Team. 2008. The Coq Proof Assistant Reference Manual. http://coq.inria.fr

J. B. Wells. 1999. Typability and Type Checking in System F are Equivalent and Undecidable. Ann. Pure Appl. Logic 98, 1-3 (1999), 111–156.