

Everything old is new again: Quoted Domain Specific Languages

Shayan Najd

The University of Edinburgh
sh.najd@ed.ac.uk

Sam Lindley

The University of Edinburgh
sam.lindley@ed.ac.uk

Josef Svenningsson

Chalmers University of Technology
josefs@chalmers.se

Philip Wadler

The University of Edinburgh
wadler@inf.ed.ac.uk

Abstract

We describe a new approach to domain specific languages (DSLs), called Quoted DSLs (QDSLs), that resurrects two old ideas: quotation, from McCarthy's Lisp of 1960, and the subformula property, from Gentzen's natural deduction of 1935. Quoted terms allow the DSL to share the syntax and type system of the host language. Normalising quoted terms ensures the subformula property, which guarantees that one can use higher-order or nested types in the source while guaranteeing first-order or flat types in the target, and enables using types to guide fusion. We test our ideas by re-implementing Feldspar, which was originally implemented as an Embedded DSL (EDSL), as a QDSL; and we compare the QDSL and EDSL variants.

Categories and Subject Descriptors D.1.1 [Applicative (Functional) Programming]; D.3.1 [Formal Definitions and Theory]; D.3.2 [Language Classifications]: Applicative (functional) languages

Keywords domain-specific language, DSL, EDSL, QDSL, embedded language, quotation, normalisation, subformula property

1. Introduction

Don't throw the past away
You might need it some rainy day
Dreams can come true again
When everything old is new again

— Peter Allen and Carole Sager

Implementing domain-specific languages (DSLs) via quotation is one of the oldest ideas in computing, going back at least to McCarthy's Lisp, which was introduced in 1960 and had macros as early as 1963. Today, a more fashionable technique is Embedded DSLs (EDSLs), which may use shallow embedding, deep embed-

ding, or a combination of the two. Our goal in this paper is to reinvigorate the idea of building DSLs via quotation, by introducing a new approach which we dub Quoted DSLs (QDSLs). A key feature of QDSLs is the use of normalisation to ensure the subformula property, first proposed by Gentzen in 1935.

Imitation is the sincerest of flattery.

— Charles Caleb Colton

Cheney et al. (2013) describe a DSL for language-integrated query in F# that translates into SQL. Their technique depends on quotation, normalisation of quoted terms, and the subformula property—an approach which we here dub QDSL. They conjecture that other DSLs might benefit from the same technique, particularly those that perform staged computation, where host code at generation-time computes target code to be executed at run-time.

Generality starts at two. Here we test the conjecture of Cheney et al. (2013) by reimplementing the EDSL Feldspar (Axelsson et al. 2010) as a QDSL. We describe the key features of the design, and show that the performance of the two versions is comparable. We compare the QDSL and EDSL variants of Feldspar, and assess the tradeoffs between the two approaches.

Davies and Pfenning (1996, 2001) also suggest quotation as a foundation for staged computation, and note a propositions-as-types connection between quotation and a modal logic; our type $Qt\ a$ corresponds to their type $\bigcirc a$. They also mention in passing the utility of normalising quoted terms, although they do not mention the subformula property.

The .NET Language-Integrated Query (LINQ) framework as used in C# and F# (Meijer et al. 2006; Syme 2006), and the Lightweight Modular Staging (LMS) framework as used in Scala (Rompf and Odersky 2010), exhibit considerable overlap with the techniques described here. Notably, they use quotation to represent staged DSL programs, and they make use to a greater or lesser extent of normalisation. In F# LINQ quotation is indicated in the normal way (by writing quoted programs inside special symbols), while in C# LINQ and Scala LMS quotation is indicated by type inference (quoted terms are given a special type).

In short, some researchers and developers are already exploiting this technique. We propose QDSL as a name to capture the commonalities among these approaches, and we observe the utility of the subformula property in this context.

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for components of this work owned by others than ACM must be honored. Abstracting with credit is permitted. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee. Request permissions from permissions@acm.org.

ICFP 2015, August 31–September 2, 2015, Vancouver, Canada.
Copyright © 2015 ACM \$15.00.
<http://dx.doi.org/10.1145/...>

shayan: Shouldn't this be box type rather than circle

Perhaps we may express the essential properties of such a normal proof by saying: it is not roundabout.

— Gerhard Gentzen

Our approach exploits the fact that normalised terms satisfy the subformula property, first introduced in the context of natural deduction by Gentzen (1935), and improved by Prawitz (1965).

The subformulas of a formula are its subparts; for instance, the subformulas of $A \rightarrow B$ are the formula itself and the subformulas of A and B . The subformula property states that every proof can be put into a normal form where the only propositions that appear in the proof are subformulas of the hypotheses and conclusion of the proof. Applying the principle of Propositions as Types (Howard 1980; Wadler 2015), the subformula property states that every lambda term can be put into a normal form where the only types that appear in the term are subformulas of the types of the free variables and the type of the term itself.

The subformula property provides users of the DSL with useful guarantees, such as the following:

- they may write higher-order terms while guaranteeing to generate first-order code;
- they may write a sequence of loops over arrays while guaranteeing to generate code that fuses those loops;
- they may write intermediate terms of nested type while guaranteeing to generate code that operates on flat data.

The first two of these are used in this paper, while the third is central to (Cheney et al. 2013).

The subformula property is closely related to conservativity. A conservativity result expresses that adding a feature to a system of logic, or to a programming language, does not make it more expressive. Consider intuitionistic logic with conjunction; conservativity states that adding implication to this logic proves no additional theorems that can be stated in the original logic.

shayan: Any reference?

Such a conservativity result is an immediate consequence of the subformula property; since the hypotheses and conjunction of the proof only mention conjunction, any proof, even if it uses implication, can be put into a normal form that only uses conjunction. Equivalently, any lambda calculus term that mentions only pair types in its free variables and result, even if it uses functions, can be put in a normal form that only uses pairs. Such a result is related to the first bullet point above; see Proposition 4.4 in Section 4.

As another example, the third bullet point above corresponds to a standard conservativity result for databases, namely that nested queries are no more expressive than flat queries (Wong 1993). This conservativity result, as implied by the subformula property, is used by Cheney et al. (2013) to show that queries that use intermediate nesting can be translated to SQL, which only queries flat tables and does not support nesting of data.

The subformula property holds only for terms in normal form. Previous work, such as (Cheney et al. 2013) uses a call-by-name normalisation algorithm that performs full β -reduction, which may cause computations to be repeated. Here we present call-by-value and call-by-need normalisation algorithms, which guarantee to preserve sharing of computations. We also present a sharpened version of the subformula property, which we apply to characterise the circumstances under which a QDSL may guarantee to generate first-order code.

Good artists copy, great artists steal.

— Picasso

EDSL is great in part because it steals the type system of its host language. Arguably, QDSL is greater because it steals the type system, the syntax, and the normalisation rules of its host language.

In theory, an EDSL should also steal the syntax of its host language, but in practice the theft is often only partial. For instance, an EDSL such as Feldspar or Nicola, when embedded in Haskell, can exploit the overloading of Haskell so that arithmetic operations in both languages appear identical, but the same is not true of comparison or conditionals.

sam: Comparisons and conditionals can be overloaded with the `RebindableSyntax` language option.

shayan: `RebindableSyntax` does not overload conditionals (nor comparisons), it overrides them. That is, conditional with `RebindableSyntax` can only have either a type proper for host language, or for an embedded language; once conditionals are overridden using `RebindableSyntax` to be used in an EDSL, they cannot be used in the host language (in the same module). Hence, one cannot use them to write the power function (if expression is used at both levels). It is important that we mention this observation here. There are fundamental limitations to what overriding mechanism can achieve in Haskell (and even in LMS).

shayan: comparison operator is just a name, which can be overridden using module system

In QDSL, of necessity the syntax of the host and embedded languages must be identical. For instance, this paper presents a QDSL variant of Feldspar, again in Haskell, where arithmetic, comparison, and conditionals are all represented by quoted terms, and hence identical to the host.

In theory, an EDSL also steals the normalisation rules of its host language, by using evaluation in the host to normalise terms of the target. In Section 5 we give several examples comparing our QDSL and EDSL versions of Feldspar. In the first of these, it is indeed the case that the EDSL achieves by evaluation of host terms what the QDSL achieves by normalisation of quoted terms. However, in other cases, the EDSL must perform normalisation of the deep embedding corresponding to what the QDSL achieves by normalisation of quoted terms.

Try to give all of the information to help others to judge the value of your contribution; not just the information that leads to judgment in one particular direction or another.

— Richard Feynman

The subformula property depends on normalisation, but normalisation may lead to an exponential blowup in the size of the normalised code.

sam: The blow-up can be much worse than exponential...

In particular, this occurs when there are nested conditional or case statements. We explain how the QDSL technique can offer the user control over where normalisation does and does not occur, while still maintaining the subformula property. Future work is required to consider the trade-offs between full normalisation as required for the subformula property and special-purpose normalisation as used in most DSLs; possibly a combination of both will prove fruitful.

Some researchers contend that an essential property of an embedded DSL which generates target code is that every term that is type-correct should successfully generate code in the target language. Neither the P-LINK of Cheney et al. (2013) nor the QDSL Feldspar of this paper satisfy this property, since the user is required to eyeball quoted code to ensure it mentions only permitted operators. If this is thought too onerous, it is possible to ensure the property with additional preprocessing.

sam: I find “QDSL Feldspar” a bit of a mouthful. Why not “QFeldspar” (and “QHaskell” and “EFeldspar”)?

This is the short and the long of it.

— Shakespeare

The contributions of this paper are:

- To introduce QDSLs as an approach to building DSLs based on quotation, normalisation of quoted terms, and the subformula property by presenting the design of a QDSL variant of Feldspar (Section 2).
- To compare QDSL and EDSL implementations of Feldspar, and show that they offer comparable performance (Section 3).
- To explain the role of the subformula property in formulating DSLs, to describe a normalisation algorithm suitable for call-by-value or call-by-need that does not lose sharing, and to formulate a sharpened version of the subformula property and apply it to characterise when higher-order terms normalise to first-order form (Section 4).
- To compare the QDSL variant of Feldspar with the deep and shallow embedding approach used in the EDSL variant of Feldspar (Section 5).

Section 6 describes related work, and Section 7 concludes.

sam: “cabal install qfeldspar” goes here?

2. Feldspar as a QDSL

Feldspar is an EDSL for writing signal-processing software, that generates code in C (Axelsson et al. 2010). We present a variant, QDSL Feldspar, that follows the structure of the previous design closely, but using the methods of QDSL rather than EDSL. We make a detailed comparison of the QDSL and EDSL designs in Section 5.

2.1 The top level

In QDSL Feldspar, our goal is to translate a quoted term to C code. The top-level function has the type:

$$qdsl :: (Rep\ a, Rep\ b) \Rightarrow Qt\ (a \rightarrow b) \rightarrow C$$

Here $Qt\ a$ represents a Haskell term of type a , its *quoted* representation, and type C represents code in C. The top-level function expects a quoted term representing a function from type a to type b , and returns C code that represents this function.

Not all types representable in Haskell are easily representable in the target language, C. For instance, we do not wish our target C code to manipulate higher-order functions. The argument type a and result type b of the main function must be representable, which is indicated by the type-class restrictions $Rep\ a$ and $Rep\ b$. Representable types include integers, floats, and pairs where the components are both representable.

```
instance Rep Int
instance Rep Float
instance (Rep a, Rep b) => Rep (a, b)
```

It is easy to add triples and larger tuples.

2.2 A first example

Let’s begin by considering the “hello world” of program generation, the power function, raising a float to a given integer. Since division by zero is undefined, we arbitrarily choose that raising zero to a negative power yields zero. Here is an optimised power function represented using QDSL:

```
power :: Int -> Qt (Float -> Float)
power n =
  if n < 0 then
    [[λx -> if x == 0 then 0
      else 1 / ($ (power (-n)) x)]]
  else if n == 0 then
    [[λx -> 1]]
  else if even n then
```

```
    [[λx -> $sqr ($ (power (n div 2)) x)]]
  else
    [[λx -> x * ($ (power (n - 1)) x)]]
sqr :: Qt (Float -> Float)
sqr = [[λy -> y * y]]
```

The typed quasi-quoting mechanism of Template Haskell is used to indicate which code executes at which time. Unquoted code executes at generation-time while quoted code executes at run-time. Quoting is indicated by $[[\dots]]$ and unquoting by $$(\dots)$.

Evaluating $power\ (-6)$ yields the following:

```
[[λx -> if x == 0 then 0 else
  1 / (λx -> (λy -> y * y)
    ((λx -> (x * ((λx -> (λy -> y * y)
      ((λx -> (x * ((λx -> 1) x)) x)) x)) x)) x)]]
```

Normalising using the technique of Section 4, with variables renamed for readability, yields the following:

```
[[λu -> if u == 0 then 0 else
  let v = u * 1 in
  let w = u * (v * v) in
  1 / (w * w)]]
```

With the exception of the top-level term, all of the overhead of lambda abstraction and function application has been removed; we explain below why this is guaranteed by the subformula property. From the normalised term it is easy to generate the final C code:

```
float prog (float u) {
  if (u == 0) {
    return 0;
  } else {
    float v = u * 1;
    float w = u * (v * v);
    return 1 / (w * w);
  }
}
```

sam: The above is illegal C. The main function must have return type int. We should use some other name for the function we generate.

By default, we always generate a routine called `prog`; it is easy to provide the name as an additional parameter if required.

Depending on your point of view, quotation in this form of QDSL is either desirable, because it makes manifest the staging, or undesirable because it is too noisy. QDSL enables us to “steal” the entire syntax of the host language for our DSL. The EDSL approach can use the same syntax for arithmetic operators, but must use a different syntax for equality tests and conditionals, as we explain in Section 5.

Within the quotation brackets there appear lambda abstractions and function applications, while our intention is to generate first-order code. How can the QDSL Feldspar user be certain that such function applications do not render transformation to first-order code impossible or introduce additional runtime overhead? The answer is the subformula property.

2.3 The subformula property

Gentzen’s subformula property guarantees that any proof can be normalised so that the only formulas that appear within it are subformulas of one of the hypotheses or of the conclusion of the proof. Viewed through the lens of Propositions as Types (Howard 1980; Wadler 2015), also known as the Curry-Howard isomorphism, Gentzen’s subformula property guarantees that any term can be normalised so that the type of each of its subterms is a subtype of either the type of one of its free variables (corresponding to hypotheses) or the term itself (corresponding to the conclusion). Here

the subtypes of a type are the type itself and the subtypes of its parts, where the parts of $a \rightarrow b$ are a and b , the parts of (a, b) are a and b , and that types *Int* and *Float* have no parts.

Further, it is easy to sharpen Gentzen's proof to guarantee a sharpened subformula property: any term can be normalised so that the type of each of its proper subterms is a proper subtype of either the type of one of its free variables (corresponding to hypotheses) or the term itself (corresponding to the conclusion). Here the proper subterms of a term are all subterms save for free variables and the term itself, and the proper subformulas of a type are all subformulas save for the type itself. In the example of the previous subsection, the sharpened subformula property guarantees that after normalisation a closed term of type $\text{float} \rightarrow \text{float}$ will only have proper subterms of type *float*, which is indeed true for the normalised term.

The subformula property depends on normalisation of terms, but complete normalisation is not always possible or desirable. The extent of normalisation may be controlled by introducing uninterpreted constants. In a context with recursion, we take $\text{fix} :: (a \rightarrow a) \rightarrow a$ as an uninterpreted constant. In a context where we wish to avoid unfolding a reduction $L M$, we take $\text{id} :: a \rightarrow a$ as an uninterpreted constant, and replace $L M$ by $\text{id } L M$.

2.4 A second example

In the previous code, we arbitrarily chose that raising zero to a negative power yields zero. Say that we wish to exploit the *Maybe* type of Haskell to refactor the code, separating identifying the exceptional case (negative exponent of zero) from choosing a value for this case (zero). We decompose *power* into two functions *power'* and *power''*, where the first returns *Nothing* in the exceptional case, and the second maps *Nothing* to a suitable default value.

The *Maybe* type is a part of the Haskell standard prelude.

```
data Maybe a = Nothing | Just a
maybe :: b → (a → b) → Maybe a → b
return :: a → Maybe a
(≡) :: Maybe a → (a → Maybe b) → Maybe b
```

Here is the refactored code.

```
power' :: Int → Qt (Float → Maybe Float)
power' n =
  if n < 0 then
    [|λx → if x == 0 then Nothing
              else do y ← $(power' (-n)) x
                   return (1 / y)|]
  else if n == 0 then
    [|λx → return 1|]
  else if even n then
    [|λx → do y ← $(power' (n div 2)) x
              return ($sqr y)|]
  else
    [|λx → do y ← $(power' (n - 1)) x
              return (x × y)|]
power'' :: Int → Qt (Float → Float)
power'' n =
  [|λx → maybe 0 (λy → y) ($ (power' n) x)|]
```

Here *sqr* is as before. Evaluation and normalisation of *power* (-6) and *power''* (-6) yield identical terms (up to renaming), and hence applying *qdsl* to these yields identical C code.

The subformula property is key: because the final type of the result does not involve *Maybe* it is certain that normalisation will remove all its occurrences. Occurrences of *do* notation are expanded to applications of (\equiv) , as usual. Rather than taking *return*, (\equiv) , and *maybe* as free variables (whose types have subformulas involv-

ing *Maybe*), we treat them as known definitions to be eliminated by the normaliser. The *Maybe* type is a sum type, and normalisation for sums is as described in Section 4.

2.5 While

Code that is intended to compile to a *while* loop in C is indicated in QDSL Feldspar by application of *while*.

$$\text{while} :: (\text{Rep } s) \Rightarrow (s \rightarrow \text{Bool}) \rightarrow (s \rightarrow s) \rightarrow s \rightarrow s$$

Rather than using side-effects, *while* takes three arguments: a predicate over the current state, of type $s \rightarrow \text{Bool}$; a function from current state to new state, of type $s \rightarrow s$; and an initial state of type s ; and it returns a final state of type s . Since we intend to compile the *while* loop to C, the type of the state is constrained to representable types.

We can define a *for* loop in terms of a *while* loop.

$$\begin{aligned} \text{for} :: (\text{Rep } s) \Rightarrow \text{Qt } (\text{Int} \rightarrow s \rightarrow (\text{Int} \rightarrow s \rightarrow s) \rightarrow s) \\ \text{for} = [| \lambda n \ s_0 \ b \rightarrow \text{snd } (\text{while } (\lambda(i, s) \rightarrow i < n) \\ \quad (\lambda(i, s) \rightarrow (i + 1, b \ i \ s)) \\ \quad (0, s_0)) |] \end{aligned}$$

The state of the *while* loop is a pair consisting of a counter and the state of the *for* loop. The body b of the *for* loop is a function that expects both the counter and the state of the *for* loop. The counter is discarded when the loop is complete, and the final state of the *for* loop returned.

As an example, we can define Fibonacci using a *for* loop.

$$\begin{aligned} \text{fib} :: \text{Qt } (\text{Int} \rightarrow \text{Int}) \\ \text{fib} = [| \lambda n \rightarrow \text{fst } (\$ \text{for } n \ (0, 1) \ (\lambda i \ (a, b) \rightarrow (b, a + b))) |] \end{aligned}$$

Again, the subformula property plays a key role. As explained in Section 2.3, primitives of the language to be compiled, such as (\times) and *while*, are treated as free variables of a given arity. As described in Section 4, we can ensure that after normalisation every occurrence of *while* has the form

$$\text{while } (\lambda s \rightarrow \dots) (\lambda s \rightarrow \dots) (\dots)$$

where the first ellipses has type *Bool*, and both occurrences of the bound variable s and the second and third ellipses all have the same type.

Unsurprisingly, and in accord with the subformula property, each occurrence of *while* in the normalised code will contain subterms with the type of its state. The restriction of state to representable types increases the utility of the subformula property. For instance, since we have chosen that *Maybe* is not a representable type, we can ensure that any top-level function without *Maybe* in its type will normalise to code not containing *Maybe* in the type of any subterm. An alternative choice is possible, as we will see in the next section.

2.6 Arrays

A key feature of Feldspar is its distinction between two types of arrays, manifest arrays, *Arr*, which may appear at run-time, and “pull arrays”, *Vec*, which are eliminated by fusion at generation-time. Again, we exploit the subformula property to ensure no subterms of type *Vec* remain in the final program.

The type *Arr* of manifest arrays is simply Haskell's array type, specialised to arrays with integer indices and zero-based indexing. The type *Vec* of pull arrays is defined in terms of existing types, as a pair consisting of the length of the array and a function that given an index returns the array element at that index.

$$\text{type Arr } a = \text{Array Int } a$$

$$\text{data Vec } a = \text{Vec Int } (\text{Int} \rightarrow a)$$

Values of type *Arr* are representable, assuming that the element type is representable, while values of type *Vec* are not representable.

instance (*Rep a*) \Rightarrow *Rep (Arr a)*

For arrays, we assume the following primitive operations.

```
makeArr :: (Rep a)  $\Rightarrow$  Int  $\rightarrow$  (Int  $\rightarrow$  a)  $\rightarrow$  Arr a
lenArr  :: (Rep a)  $\Rightarrow$  Arr a  $\rightarrow$  Int
ixArr   :: (Rep a)  $\Rightarrow$  Arr a  $\rightarrow$  Int  $\rightarrow$  a
```

The first populates a manifest array of the given size using the given indexing function, the second returns the length of the array, and the third returns the array element at the given index.

We define functions to convert between the two representations in the obvious way.

```
toVec    :: Rep a  $\Rightarrow$  Qt (Arr a  $\rightarrow$  Vec a)
toVec    = [| $\lambda a \rightarrow$  Vec (lenArr a) ( $\lambda i \rightarrow$  ixArr a i)|]
fromVec  :: Rep a  $\Rightarrow$  Qt (Vec a  $\rightarrow$  Arr a)
fromVec  = [| $\lambda (Vec n g) \rightarrow$  makeArr n g|]
```

It is straightforward to define operations on vectors, including combining corresponding elements of two vectors, summing the elements of a vector, dot product of two vectors, and norm of a vector.

```
zipVec   :: Qt ((a  $\rightarrow$  b  $\rightarrow$  c)  $\rightarrow$  Vec a  $\rightarrow$  Vec b  $\rightarrow$  Vec c)
zipVec   = [| $\lambda f (Vec m g) (Vec n h) \rightarrow$ 
              Vec ( $\$ \$ \min m n$ ) ( $\lambda i \rightarrow f (g i) (h i)$ )|]
sumVec   :: (Rep a, Num a)  $\Rightarrow$  Qt (Vec a  $\rightarrow$  a)
sumVec   = [| $\lambda (Vec n g) \rightarrow$   $\$ \$ \text{for } n 0 (\lambda i x \rightarrow x + g i)$ |]
dotVec   :: (Rep a, Num a)  $\Rightarrow$  Qt (Vec a  $\rightarrow$  Vec a  $\rightarrow$  a)
dotVec   = [| $\lambda u v \rightarrow$   $\$ \$ \text{sumVec } (\$ \$ \text{zipVec } (\times) u v)$ |]
normVec  :: Qt (Vec Float  $\rightarrow$  Float)
normVec  = [| $\lambda v \rightarrow \text{sqrt } (\$ \$ \text{dotVec } v v)$ |]
```

The second of these uses the *for* loop defined in Section 2.5, the third is defined using the first two, and the fourth is defined using the third. Recall that our sharpened subformula property required that (\times) be fully applied, so before normalisation (\times) is expanded to $\lambda x y \rightarrow x \times y$.

sam: We haven't yet mentioned anything about constants being fully applied in the sharpened subformula property.

Our final function cannot accept *Vec* as input, since the *Vec* type is not representable, but it can accept *Arr* as input. Invoking *qdsl* [| $\lambda v \rightarrow$ $\$ \$ \text{normVec } (\$ \$ \text{toVec } v)$ |] produces the following C code.

```
float prog (float[] a) {
  float x = 0;
  int i = 0;
  while (i < lenArr a) {
    x = x + a[i] * a[i];
    i = i+1;
  }
  return sqrt(x);
}
```

Types and the subformula property help us to guarantee fusion. The subformula property guarantees that all occurrences of *Vec* must be eliminated, while occurrences of *Arr* will remain. There are some situations where fusion is not beneficial, notably when an intermediate vector is accessed many times, in which case fusion will cause the elements to be recomputed. An alternative is to materialise the vector as an array with the following function.

```
memorise :: Rep a  $\Rightarrow$  Qt (Vec a  $\rightarrow$  Vec a)
```

For example, if

```
blur :: Rep a  $\Rightarrow$  Qt (Vec a  $\rightarrow$  Vec a)
```

averages adjacent elements of a vector, then one may choose to compute either

```
[| $\$ \$ \text{blur} \circ \$ \$ \text{blur}$ |] or [| $\$ \$ \text{blur} \circ \$ \$ \text{memorise} \circ \$ \$ \text{blur}$ |]
```

with different trade-offs between recomputation and memory usage. Strong guarantees for fusion in combination with *memorise* give the programmer a simple interface which provides powerful optimisations combined with fine control over memory usage.

We have described the application of the subformula to array fusion as based on “pull arrays” (Svenningsson and Axelsson 2012), but the same technique should also apply to other techniques that support array fusion, such as “push arrays” (Claessen et al. 2012).

3. Implementation

The original EDSL Feldspar generates values of an algebraic type (called *Dp a* in Section 5), with constructs that represent *while* and manifest arrays similar to those above. A backend then compiles values of type *Dp a* to C code. QDSL Feldspar provides a transformer from *Qt a* to *Dp a*, and shares the EDSL Feldspar backend.

The transformer from *Qt* to *Dp* performs the following steps.

- In any context where a constant *c* is not fully applied, it replaces *c* with $\lambda \bar{x}. c \ \bar{x}$. It replaces identifiers connected to the type *Maybe*, such as *return*, $(\gg=)$, and *maybe*, by their definitions.
- It normalises the term to ensure the subformula property, using the rules of Section 4. The normaliser does not yet support all Haskell data types, but does support tuples, and the types *Maybe* and *Vec*.
- It traverses the term, converting *Qt* to *Dp*. It checks that only permitted primitives appear in *Qt*, and translates these to their corresponding representation in *Dp*. Permitted primitives include: $(=)$, $(<)$, $(+)$, (\times) , and similar, plus *while*, *makeArr*, *lenArr*, and *ixArr*.

An unfortunate feature of typed quasiquotation in GHC is that the implementation discards all type information when creating the representation of a term. Type *Qt a* is equivalent to *TH.Q (TH.TExp a)*, where *TH* denotes the library for Template Haskell, *TH.Q* is the quotation monad of Template Haskell (used to look up identifiers and generate fresh names), and *TH.TExp a* is the parse tree for a quoted expression returning a value of type *a*. Type *TH.TExp a* is just a wrapper for *TH.Exp*, the (untyped) parse tree of an expression in Template Haskell, where *a* is a phantom type variable. Hence, the translator from *Qt a* to *Dp a* is forced to re-infer all the type information for the subterms of the term of type *Qt a*. This is why we currently translate the *Maybe* monad as a special case, rather than supporting overloading for monad operations in general. Moreover, overloaded arithmetic operations are treated as fully polymorphic operations, at first; their type constraints are ignored. Just before normalisation, their inferred types are used to translate them to their corresponding monomorphic primitives.

shayan: I have added above two lines about how QFeldspar treats overloaded arithmetic operations.

We measured the behaviour of five benchmark programs.

IPGray	Image Processing (Grayscale)
IPBW	Image Processing (Black and White)
FFT	Fast Fourier Transform
CRC	Cyclic Redundancy Check
Windowing	Average array in a sliding window

sam: The key should appear in the table: I have no idea what Hc, Hr, etc. mean.

	QDSL				EDSL without CSE				EDSL with CSE			
	Hc	Hr	Cc	Cr	Hc	Hr	Cc	Cr	Hc	Hr	Cc	Cr
IPGray	11.40	0.00	0.08	0.42	6.22	0.00	0.08	0.45	6.55	0.00	0.07	0.45
IPBW	11.26	0.00	0.08	0.20	6.22	0.00	0.07	0.20	6.55	0.00	0.07	0.20
FFT	11.36	0.24	0.08	5.08	6.15	0.47	0.13	—	6.65	0.16	0.09	6.15
CRC	11.29	0.01	0.08	0.14	6.13	0.00	0.08	0.15	6.49	0.00	0.08	0.15
Windowing	11.19	0.01	0.08	0.31	6.21	0.01	0.08	0.33	6.51	0.00	0.08	0.33

Figure 1. Comparison of QDSL and EDSL MiniFeldspar (all times in seconds)

Figure 1 lists the results. Columns Hc and Hr list compile-time and run-time in Haskell, and Cc and Cr list compile-time and run-time in C. Runs for EDSL are shown both with and without common subexpression elimination (CSE), which is supported by a simple form of observable sharing (Claessen and Sands 1999; Gill 2009). QDSL does not require CSE, since the normalisation algorithm preserves sharing. One benchmark, FFT, exhausts memory without CSE. All benchmarks produce essentially the same C for both QDSL and EDSL, which run in essentially the same time. The one exception is FFT, where Feldspar appears to introduce spurious conversions that increase the runtime.

Measurements were done on a quad-core Intel i7-2640M CPU running at 2.80 GHz and 3.7 GiB of RAM, with GHC Version 7.8.3 and GCC version 4.8.2, running on Ubuntu 14.04 (64-bit).

4. The subformula property

This section introduces a collection of reduction rules for normalising terms that enforces the subformula property while ensuring sharing is preserved. The rules adapt to both call-by-need and call-by-value.

We work with simple types. The only polymorphism in our examples corresponds to instantiating constants (such as *while*) at different types.

Types, terms, and values are presented in Figure 2. Let A, B, C range over types, including base types (ι), functions ($A \rightarrow B$), products ($A \times B$), and sums ($A + B$). Let L, M, N range over terms, and x, y, z range over variables. Let c range over constants, which are fully applied according to their arity, as discussed below. As usual, terms are taken as equivalent up to renaming of bound variables. Write $FV(M)$ for the set of free variables of M , and $N[x := M]$ for capture-avoiding substitution of M for x in N . Let V, W range over values, and P range over terms that are not values.

Let Γ range over type environments, which are sets of pairs of variables with types $x : A$. Write $\Gamma \vdash M : A$ to indicate that term M has type A under type environment Γ . Typing rules are standard.

Reduction rules for normalisation are presented in Figure 3. The rules are confluent, so order of application is irrelevant to the final answer, but we break them into three phases to ease the proof of strong normalisation. It is easy to confirm that all of the reduction rules preserve sharing and preserve order of evaluation.

Write $M \mapsto_i N$ to indicate that M reduces to N in phase i . Only Phases 1 and 2 are required to normalise terms, but if the semantics is call-by-need then Phase 3 may also be applied. Let F and G range over two different forms of evaluation frame used in Phases 1 and 2 respectively. We write $FV(F)$ for the set of free variables of F , and similarly for G . The reduction relation is closed under compatible closure.

The normalisation procedure consists of exhaustively applying the reductions of Phase 1 until no more apply, then similarly for Phase 2, and finally for Phase 3. Phase 1 performs let-insertion,

naming subterms that are not values, along the lines of a translation to A-normal form (Flanagan et al. 1993) or reductions (let.1) and (let.2) in Moggi’s metalanguage for monads (Moggi 1991). Phase 2 performs two kinds of reduction: β rules apply when an introduction (construction) is immediately followed by an elimination (deconstruction), and κ rules push eliminators closer to introducers to enable β rules. Phase 3 “garbage collects” unused terms as in the call-by-need lambda calculus (Maraist et al. 1998). Phase 3 should be omitted if the intended semantics of the target language is call-by-value rather than call-by-need.

Every term has a normal form.

PROPOSITION 4.1 (Strong normalisation). *Each of the reduction relations \mapsto_i is confluent and strongly normalising: all \mapsto_i reduction sequences on well-typed terms are finite.*

The only non-trivial proof is for \mapsto_2 , which can be proved via a standard reducibility argument (see, for example, (Lindley 2007)). If the target language includes general recursion, normalisation should treat the fixpoint operator as an uninterpreted constant.

The *subformulas* of a type are the type itself and its components. For instance, the subformulas of $A \rightarrow B$ are itself and the subformulas of A and B . The *proper subformulas* of a type are all its subformulas other than the type itself.

The *subterms* of term are the term itself and its components. For instance, the subterms of $\lambda x. N$ are itself and the subterms of N and the subterms of $L M$ are itself and the subterms of L and M . The *proper subterms* of a term are all its subterms other than the term itself.

Constants are always fully applied; they are introduced as a separate construct to avoid consideration of irrelevant subformulas and subterms. The type of a constant c of arity k is written

$$c : A_1 \rightarrow \dots \rightarrow A_k \rightarrow B$$

An application of a constant c of arity k is written

$$c M_1 \dots M_k$$

and its subterms are itself and M_1, \dots, M_k (but not $c M_1 \dots M_j$ for $j < k$). Free variables are equivalent to constants of arity zero.

Terms in normal form satisfy the subformula property.

PROPOSITION 4.2 (Subformula property). *If $\Gamma \vdash M : A$ and M is in normal form, then every subterm of M has a type that is either a subformula of A , a subformula of a type in Γ , or a subformula of the type of a constant in M .*

The proof follows the lines of Prawitz (1965). The differences are that we have introduced fully applied constants (to enable the sharpened subformula property, below), and that our reduction rules introduce *let*, in order to ensure sharing is preserved.

Normalisation may lead to an exponential increase in the size of a term, for instance when there are nested *case* expressions. This was not a problem for the examples we considered in Section 3,

Types	$A, B, C ::= \iota \mid A \rightarrow B \mid A \times B \mid A + B$
Terms	$L, M, N ::= x \mid c \overline{M} \mid \lambda x. N \mid L M \mid \text{let } x = M \text{ in } N \mid (M, N) \mid \text{fst } L \mid \text{snd } L \mid$ $\text{inl } M \mid \text{inr } N \mid \text{case } L \text{ of } \{\text{inl } x. M; \text{inr } y. N\}$
Values	$V, W ::= x \mid \lambda x. N \mid (V, W) \mid \text{inl } V \mid \text{inr } W$

sam: Perhaps P should appear here?

Figure 2. Types, Terms, and Values

Phase 1 (let-insertion)

$$F ::= c(\overline{V}, [], \overline{N}) \mid [] M \mid V [] \mid ([], M) \mid (V, []) \mid \text{fst } [] \mid \text{snd } [] \mid \text{inl } [] \mid \text{inr } [] \mid \text{case } [] \text{ of } \{\text{inl } x. M; \text{inr } y. N\}$$

$$(\text{let}) \quad F[P] \mapsto_1 \text{let } x = P \text{ in } F[x], \quad x \text{ fresh}$$

Phase 2 (symbolic evaluation)

$$G ::= [] V \mid \text{let } x = [] \text{ in } N$$

$(\kappa.\text{let})$	$G[\text{let } x = P \text{ in } N]$	\mapsto_2	$\text{let } x = P \text{ in } G[N],$	$x \notin FV(G)$
$(\kappa.\text{case})$	$G[\text{case } V \text{ of } \{\text{inl } x. M; \text{inr } y. N\}]$	\mapsto_2	$\text{case } V \text{ of } \{\text{inl } x. G[M]; \text{inr } y. G[N]\},$	$x, y \notin FV(G)$
$(\beta.\rightarrow)$	$(\lambda x. N) V$	\mapsto_2	$N[x := V]$	
$(\beta.\times_1)$	$\text{fst } (V, W)$	\mapsto_2	V	
$(\beta.\times_2)$	$\text{snd } (V, W)$	\mapsto_2	W	
$(\beta.+_1)$	$\text{case } (\text{inl } V) \text{ of } \{\text{inl } x. M; \text{inr } y. N\}$	\mapsto_2	$M[x := V]$	
$(\beta.+_2)$	$\text{case } (\text{inr } W) \text{ of } \{\text{inl } x. M; \text{inr } y. N\}$	\mapsto_2	$N[y := W]$	
$(\beta.\text{let})$	$\text{let } x = V \text{ in } N$	\mapsto_2	$N[x := V]$	

Phase 3 (garbage collection)

$$(\text{need}) \quad \text{let } x = P \text{ in } N \mapsto_3 N, \quad x \notin FV(N)$$

sam: If $[] M$ appears in F then $[] V$ is redundant in G . Shayan's original presentation includes $M []$ in F but not $[] M$ or $V []$. This was an in order to avoid naming every intermediate partial application.

Figure 3. Normalisation Rules

but may be a problem in some contexts. Normalisation may be controlled by introduction of uninterpreted constants (see Section 2.3), but further work is needed to understand the contexts in which complete normalisation is desirable and the contexts in which it is problematic.

Examination of the proof in Prawitz (1965) shows that in fact normalisation achieves a sharper property.

PROPOSITION 4.3 (Sharpened subformula property). *If $\Gamma \vdash M : A$ and M is in normal form, then every proper subterm of M that is not a free variable or a subterm of a constant application has a type that is a proper subformula of A or a proper subformula of a type in Γ .*

The sharpened subformula property says nothing about the type of subterms of constant applications, but this is immediately apparent by recursive application of the sharpened subformula property. Given a subterm that is a constant application $c \overline{M}$, where c has type $\overline{A} \rightarrow B$, then the subterm itself has type B , each subterm M_i has type A_i , and every proper subterm of M_i that is not a free variable of M_i or a subterm of a constant application has a type that is a proper subformula of A_i or a proper subformula of the type of one of its free variables.

In Section 2, an important property is that every top-level term passed to *qdsl* is suitable for translation to C after normalisation. Here we are interested in C as a *first-order* language. The exact property required is somewhat subtle. One might at first guess the required property is that every subterm is representable, in the sense introduced in Section 2.1, but this is not quite right. The top-level

term is a function from a representable type to a representable type. And the constant *while* expects subterms of type $s \rightarrow \text{Bool}$ and $s \rightarrow s$, where the state s is representable. Fortunately, the exact property required is not hard to formulate in a general way, and is easy to ensure by applying the sharpened subformula property.

We introduce a variant of the usual notion of *rank* of a type, with respect to a notion of representability. A term of type $A \rightarrow B$ has rank $\min(m+1, n)$ where m is the rank of A and n is the rank of B , while a term of representable type has rank 0.

The property we need to ensure ease of translation to C (or any other first-order language) is as follows.

PROPOSITION 4.4 (Rank and representability). *Consider a term M of rank 1, where every free variable of M has rank 0 and every constant in M has rank at most 2. Then M normalises to a form where every subterm either is of representable type or is of the form $\lambda \overline{x}. N$ where each of the bound variables x_i and the body N has representable type.*

The property follows immediately by observing that any term L with type of rank 1 can be rewritten in the form $\lambda \overline{x}. L \overline{x}$ where each bound variable and the body has representable type, and then normalising and applying the sharpened subformula property.

sam: Is it worth saying something here about in future generalising to higher-order target languages?

5. Feldspar as an EDSL

This section reviews the combination of deep and shallow embeddings required to implement Feldspar as an EDSL, and considers

the trade-offs between the QDSL and EDSL approaches. Much of this section reprises Svenningsson and Axelsson (2012).

The top-level function of EFeldspar has the type:

$$qdsl :: (Rep\ a, Rep\ b) \Rightarrow (Dp\ a \rightarrow Dp\ b) \rightarrow C$$

Here $Dp\ a$ is the deep representation of a term of type a . The deep representation is described in detail in Section 5.3 below, and is chosen to be easy to translate to C. As before, type C represents code in C, and type class Rep restricts to representable types.

5.1 A first example

Here is the power function of Section 2.2, now represented using EDSL:

```
power :: Int → Dp Float → Dp Float
power n x =
  if n < 0 then
    x ==. 0 ? (0, 1 / power (-n) x)
  else if n == 0 then
    1
  else if even n then
    sqr (power (n div 2) x)
  else
    x × power (n - 1) x
sqr :: Dp Float → Dp Float
sqr y = y × y
```

Type $Q\ (Float \rightarrow Float)$ in the QDSL variant becomes $Dp\ Float \rightarrow Dp\ Float$ in the EDSL variant, meaning that $power\ n$ accepts a representation of the argument and returns a representation of that argument raised to the n 'th power.

In EDSL, no quotation is required, and the code looks almost—but not quite!—like an unstaged version of power, but with different types. Clever encoding tricks, explained later, permit declarations, function calls, arithmetic operations, and numbers to appear the same whether they are to be executed at generation-time or run-time. However, as explained later, comparison and conditionals appear differently depending on whether they are to be executed at generation-time or run-time, using $M == N$ and $if\ L\ then\ M\ else\ N$ for the former but $M ==. N$ and $L?(M, N)$ for the latter.

Evaluating $power\ (-6)$ yields the following:

$$(\lambda u \rightarrow (u ==. 0) ? (0, \\ 1 / ((u \times ((u \times 1) \times (u \times 1))) \times \\ (u \times ((u \times 1) \times (u \times 1))))))$$

Applying common-subexpression elimination, or using a technique such as observable sharing, permits recovering the sharing structure.

$$\begin{array}{l|l} v & (u \times 1) \\ w & u \times (v \times v) \\ \text{top} & (u ==. 0) ? (0, 1 / (w \times w)) \end{array}$$

From the above, it is easy to generate the final C code, which is identical to that in Section 2.2.

Here are points of comparison between the two approaches.

- A function $a \rightarrow b$ is embedded in QDSL as $Qt\ (a \rightarrow b)$, a representation of a function, and in EDSL as $Dp\ a \rightarrow Dp\ b$, a function between representations.
- QDSL enables the host and embedded languages to appear identical. In contrast, EDSL requires some term forms, such as comparison and conditionals, to differ between the host and embedded languages.
- QDSL requires syntax to separate quoted and unquoted terms. In contrast, EDSL permits the host and embedded languages to

intermingle seamlessly. Depending on your point of view, explicit quotation syntax may be considered as an unnecessary distraction or as drawing a useful distinction between generation-time and run-time. If one takes the former view, the type-based approach to quotation found in C# and Scala might be preferred.

- QDSL may share the same representation for quoted terms across a range of applications; the quoted language is the host language, and does not vary with the specific domain. In contrast, EDSL typically develops custom shallow and deep embeddings for each application, although these may follow a fairly standard pattern.
- QDSL yields an unwieldy term that requires normalisation. In contrast, EDSL yields the term in normalised form in this case, though there are other situations where a normaliser is required (see Section 5.2).
- QDSL requires traversing the quoted term to ensure it only mentions permitted identifiers. In contrast, EDSL guarantees that if a term has the right type it will translate to the target. If the requirement to eyeball code to ensure only permitted identifiers are used is considered too onerous, it should be easy to build a preprocessor that checks this property. In Haskell, it is possible to incorporate such a preprocessor using MetaHaskell (Mainland 2012).
- Since QDSLs may share the same quoted terms across a range of applications, the cost of building a normaliser or a preprocessor might be amortised. In the conclusion, we consider the design of a tool for building QDSLs that may perform both these functions.
- QDSL preserves sharing. In contrast, EDSL loses sharing, which must later be recovered either by common subexpression elimination or by applying a technique such as observable sharing.
- Once the deep embedding or the normalised quoted term is produced, generating the domain-specific code is similar for both approaches.

sam: A point not discussed here is code reuse between host and embedded languages. Links allows exactly the same code to be written once and used both in host and embedded code. QDSLs in Haskell do not (currently). F# supports an attribute which can be used to reflect an arbitrary module as a quoted AST. What about other EDSL approaches?

5.2 A second example

In Section 2.4, we exploited the *Maybe* type to refactor the code.

In EDSL, we must use a new type, where *Maybe*, *Nothing*, *Just*, and *maybe* become *Opt*, *none*, *some*, and *option*, and *return* and $(\gg=)$ are similar to before.

```
type Opt a
none  :: Undef a ⇒ Opt a
some  :: a → Opt a
return :: a → Opt a
(≫=)  :: Opt a → (a → Opt b) → Opt b
option :: (Undef a, Undef b) ⇒
  b → (a → b) → Opt a → b
```

Type class *Undef* is explained in Section 5.6, and details of type *Opt* are given in Section 5.7.

Here is the refactored code.

```
power' :: Int → Dp Float → Opt (Dp Float)
power' n x =
  if n < 0 then
    (x ==. 0) ? (none,
```



```

do y ← power' (-n) x
  return (1 / y))
else if n == 0 then
  return 1
else if even n then
  do y ← power' (n div 2) x
    return (sqr y)
else
  do y ← power' (n - 1) x
    return (x × y)
power'' :: Int → Dp Float → Dp Float
power'' n x = option 0 (λy → y) (power' n x)

```

Here *sqr* is as before.

The term of type *Dp Float* generated by evaluating *power* (-6) *x* is large and unscrutable:

```

(((fst ((x == 0.0) ? (((False ? (True, False)), (False ?
(undef, undef))), (True, (1.0 / ((x × ((x × 1.0) × (x ×
1.0))) × (x × ((x × 1.0) × (x × 1.0)))))))) ? (True,
False)) ? (((fst ((x == 0.0) ? (((False ? (True, False)),
(False ? (undef, undef))), (True, (1.0 / ((x × ((x × 1.0)
× (x × 1.0))) × (x × ((x × 1.0) × (x × 1.0)))))))) ? ((snd
((x == 0.0) ? (((False ? (True, False)), (False ? (undef,
undef))), (True, (1.0 / ((x × ((x × 1.0) × (x × 1.0))) ×
(x × ((x × 1.0) × (x × 1.0))))))))), undef)), 0.0))

```

Before, evaluating *power* yielded a term essentially in normal form. However, here rewrite rules need to be repeatedly applied, similar to those required to enforce the subformula property as described in Section 4. After applying these rules, common subexpression elimination yields the same structure as in the previous subsection, from which the same C code is generated.

Hence, an advantages of the EDSL approach—that it generates terms essentially in normal form—turns out to be restricted to a limited set of types, including functions and products, but excluding sums. If one wishes to deal with sum types, separate normalisation is required. This is one reason why we do not consider normalisation as required by QDSL to be particularly onerous.

Here are points of comparison between the two approaches.

- Both QDSL and EDSL can exploit notational conveniences in the host language. The example here exploits Haskell **do** notation; the embedding of SQL in F# by Cheney et al. (2013) exploits F# sequence notation. For EDSL, exploiting **do** notation just requires instantiating *return* and (\gg) correctly. For QDSL, it is also necessary for the normaliser to recognise and expand **do** notation and to substitute appropriate instances of *return* and (\gg).
- As this example shows, sometimes both QDSL and EDSL may require normalisation. As mentioned previously, for QDSLs the cost of building a normaliser might be amortised across several applications. In contrast, each EDSL usually has a distinct deep representation and so requires a distinct normaliser.

5.3 The deep embedding

Recall that a value of type *Dp a* represents a term of type *a*, and is called a deep embedding.

```

data Dp a where
  LitB :: Bool → Dp Bool
  LitI :: Int → Dp Int
  LitF :: Float → Dp Float
  If :: Dp Bool → Dp a → Dp a → Dp a
  While :: (Dp a → Dp Bool) →
    (Dp a → Dp a) → Dp a → Dp a

```

```

Pair :: Dp a → Dp b → Dp (a, b)
Fst :: Dp (a, b) → Dp a
Snd :: Dp (a, b) → Dp b
Prim1 :: String → Dp a → Dp b
Prim2 :: String → Dp a → Dp b → Dp c
Arr :: Dp Int → (Dp Int → Dp a) → Dp (Arr a)
ArrLen :: Dp (Arr a) → Dp Int
ArrIx :: Dp (Arr a) → Dp Int → Dp a
Variable :: String → Dp a

```

The type above represents a low level, pure functional language with a straightforward translation to C. It uses higher-order abstract syntax (HOAS) to represent constructs with variable binding Pfenning and Elliot (1988).

sam: Technically this isn't true HOAS because it allows exotic terms.

The deep embedding has boolean, integer, and floating point literals, conditionals, while loops, pairs, primitives, arrays, and special-purpose constructs for variables and values. Constructs *LitB*, *LitI*, *LitF* build literals. Construct *If* builds a conditional. Construct *While* resembles *while* in Section 2.5. Constructs *Pair*, *Fst*, and *Snd* build pairs and extract the first and second component. Constructs *Prim1* and *Prim2* represent primitive operations, the string is the name of the operation. Construct *ArrIx* creates a new array from a length and a body that computes the array element for each index, construct *ArrLen* extracts the length from an array, and construct *ArrIx* fetches the element at a given index. Construct *Variable* represents a variable.

5.4 Class Syn

We introduce a type class *Syn* that allows us to convert shallow embeddings to and from deep embeddings.

```

class Syn a where
  type Internal a
  toDp :: a → Dp (Internal a)
  fromDp :: Dp (Internal a) → a

```

Type *Internal* is a GHC type family (Chakravarty et al. 2005). Functions *toDp* and *fromDp* translate between the shallow embedding *a* and the deep embedding *Dp (Internal a)*.

The first instance of *Syn* is *Dp* itself, and is straightforward.

```

instance Syn (Dp a) where
  type Internal (Dp a) = a
  toDp = id
  fromDp = id

```

Our representation of a run-time *Bool* will have type *Dp Bool* in both the deep and shallow embeddings, and similarly for *Int* and *Float*.

We do not code the target language using its constructs directly. Instead, for each constructor we define a corresponding “smart constructor” using class *Syn*.

```

true, false :: Dp Bool
true = LitB True
false = LitB False

(?) :: Syn a ⇒ Dp Bool → (a, a) → a
c ? (t, e) = fromDp (If c (toDp t) (toDp e))

while :: Syn a ⇒ (a → Dp Bool) → (a → a) → a → a
while c b i = fromDp (While (c ∘ fromDp)
  (toDp ∘ b ∘ fromDp)
  (toDp i))

```

Numbers are made convenient to manipulate via overloading.

```

instance Num (Dp Int) where
  a + b = Prim2 "(+)" a b
  a - b = Prim2 "(-)" a b
  a × b = Prim2 "(*)" a b
  fromInteger a = LitI (fromInteger a)

```

With this declaration, $1 + 2 :: Dp\ Int$ evaluates to

```
Prim2 "(+)" (LitI 1) (LitI 2),
```

permitting code executed at generation-time and run-time to appear identical. A similar declaration works for *Float*.

Comparison also benefits from smart constructors.

```

(==.) :: (Syn a, Eq (Internal a)) => a -> a -> Dp Bool
a ==. b = Prim2 "(=)" (toDp a) (toDp b)

(<.) :: (Syn a, Ord (Internal a)) => a -> a -> Dp Bool
a <. b = Prim2 "<" (toDp a) (toDp b)

```

Overloading cannot apply here, because Haskell requires $(==)$ return a result of type *Bool*, while $(==.)$ returns a result of type *Dp Bool*, and similarly for $(<.)$.

5.5 Embedding pairs

We set up a correspondence between host language pairs in the shallow embedding and target language pairs in the deep embedding.

```

instance (Syn a, Syn b) => Syn (a, b) where
  type Internal (a, b) = (Internal a, Internal b)
  toDp (a, b) = Pair (toDp a) (toDp b)
  fromDp p = (fromDp (Fst p), fromDp (Snd p))

```

This permits us to manipulate pairs as normal, with (a, b) , *fst* a , and *snd* a . Argument p is duplicated in the definition of *fromDp*, which may require common subexpression elimination as discussed in Section 5.1.

We have now developed sufficient machinery to define a *for* loop in terms of a *while* loop.

```

for :: Syn a => Dp Int -> a -> (Dp Int -> a -> a) -> a
for n s0 b = snd (while (\(i, s) -> i <. n)
                      (\(i, s) -> (i + 1, b i s))
                      (0, s0))

```

The state of the *while* loop is a pair consisting of a counter and the state of the *for* loop. The body b of the *for* loop is a function that expects both the counter and the state of the *for* loop. The counter is discarded when the loop is complete, and the final state of the *for* loop returned.

Thanks to our machinery, the above definition uses only ordinary Haskell pairs. The condition and body of the *while* loop pattern match on the state using ordinary pair syntax, and the initial state is constructed as an ordinary pair.

5.6 Embedding undefined

For the next section, which defines an analogue of the *Maybe* type, it will prove convenient to work with types which have a distinguished value at each type, which we call *undef*.

It is straightforward to define a type class *Undef*, where type a belongs to *Undef* if it belongs to *Syn* and it has an undefined value.

```

class Syn a => Undef a where
  undef :: a

instance Undef (Dp Bool) where
  undef = false

instance Undef (Dp Int) where

```

```
  undef = 0
```

```

instance Undef (Dp Float) where
  undef = 0

```

```

instance (Undef a, Undef b) => Undef (a, b) where
  undef = (undef, undef)

```

For example,

```

(/#) :: Dp Float -> Dp Float -> Dp Float
x /# y = (y ==. 0) ? (undef, x / y)

```

behaves as division, save that when the divisor is zero it returns the undefined value of type *Float*, which is also zero.

Svenningsson and Axelsson (2012) claim that it is not possible to support *undef* without changing the deep embedding, but here we have defined *undef* entirely as a shallow embedding. (It appears they underestimated the power of their own technique!)

5.7 Embedding option

We now explain in detail the *Opt* type seen in Section 2.4.

The deep-and-shallow technique cleverly represents deep embedding $Dp\ (a, b)$ by shallow embedding $(Dp\ a, Dp\ b)$. Hence, it is tempting to represent $Dp\ (Maybe\ a)$ by $Maybe\ (Dp\ a)$, but this cannot work, because *fromDp* would have to decide at generation-time whether to return *Just* or *Nothing*, but which to use is not known until run-time.

Instead, Svenningsson and Axelsson (2012) represent values of type *Maybe a* by the type *Opt' a*, which pairs a boolean with a value of type a . For a value corresponding to *Just x*, the boolean is true and the value is x , while for one corresponding to *Nothing*, the boolean is false and the value is *undef*. We define *some'*, *none'*, and *opt'* as the analogues of *Just*, *Nothing*, and *maybe*. The *Syn* instance is straightforward, mapping options to and from the pairs already defined for *Dp*.

```

data Opt' a = Opt' { def :: Dp Bool, val :: a }

instance Syn a => Syn (Opt' a) where
  type Internal (Opt' a) = (Bool, Internal a)
  toDp (Opt' b x) = Pair b (toDp x)
  fromDp p = Opt' (Fst p) (fromDp (Snd p))

some' :: a -> Opt' a
some' x = Opt' true x

none' :: Opt' a
none' = Opt' false undef

option' :: Syn b => b -> (a -> b) -> Opt' a -> b
option' d f o = def o ? (f (val o), d)

```

The next obvious step is to define a suitable monad over the type *Opt'*. The natural definitions to use are as follows:

```

return :: a -> Opt' a
return x = some' x

(≫) :: (Undef b) => Opt' a -> (a -> Opt' b) -> Opt' b
o ≫ g = Opt' (def o ? (def (g (val o)), false))
         (def o ? (val (g (val o)), undef))

```

However, this adds type constraint *Undef b* to the type of (\gg) , which is not permitted. The need to add such constraints often arises, and has been dubbed the constrained-monad problem (Hughes 1999; Sculthorpe et al. 2013; Svenningsson and Svensson 2013). We solve it with a trick due to Persson et al. (2011).

sam: I think Jeremy Yallop may have also covered this kind of issue somewhere in his thesis.

We introduce a second continuation-passing style (CPS) type *Opt*, defined in terms of the representation type *Opt'*. It is straightforward to define *Monad* and *Syn* instances for the CPS type, op-

erations to lift the representation type to CPS and to lower CPS to the representation type, and to lift *some*, *none*, and *option* from the representation type to the CPS type. The *lift* operation is closely related to the (\gg) operation we could not define above; it is properly typed, thanks to the type constraint on *b* in the definition of *Opt a*.

```

newtype Opt a =
  O { unO :: ∀ b. Undef b ⇒ ((a → Opt' b) → Opt' b) }

instance Monad Opt where
  return x = O (λg → g x)
  m >> k = O (λg → unO m (λx → unO (k x) g))

instance Undef a ⇒ Syn (Opt a) where
  type Internal (Opt a) = (Bool, Internal a)
  fromDp = lift ∘ fromDp
  toDp = toDp ∘ lower

lift :: Opt' a → Opt a
lift o = O (λg → Opt' (def o ? (def (g (val o)), false))
              (def o ? (val (g (val o))), undef)))

lower :: Undef a ⇒ Opt a → Opt' a
lower m = unO m some'

none :: Undef a ⇒ Opt a
none = lift none'

some :: a → Opt a
some a = lift (some' a)

option :: (Undef a, Undef b) ⇒
          b → (a → b) → Opt a → b
option d f o = option' d f (lower o)

```

These definitions support the EDSL code presented in Section 5.2.

5.8 Embedding vector

Recall that values of type *Array* are created by construct *Arr*, while *ArrLen* extracts the length and *ArrIx* fetches the element at the given index. Corresponding to the deep embedding *Array* is a shallow embedding *Vec*.

```

data Vec a = Vec (Dp Int) (Dp Int → a)

instance Syn a ⇒ Syn (Vec a) where
  type Internal (Vec a) = Array Int (Internal a)
  toDp (Vec n g) = Arr n (toDp ∘ g)
  fromDp a = Vec (ArrLen a)
              (λi → fromDp (ArrIx a i))

instance Functor Vec where
  fmap f (Vec n g) = Vec n (f ∘ g)

```

The constructor *Vec* resembles the constructor *Arr*, but the former constructs a high-level representation of the array and the latter an actual array. It is straightforward to make *Vec* an instance of *Functor*.

It is straightforward to define operations on vectors, including combining corresponding elements of two vectors, summing the elements of a vector, dot product of two vectors, and norm of a vector.

```

zipVec :: (Syn a, Syn b) ⇒
  (a → b → c) → Vec a → Vec b → Vec c
zipVec f (Vec m g) (Vec n h)
  = Vec (m 'min' n) (λi → f (g i) (h i))

sumVec :: (Syn a, Num a) ⇒ Vec a → a
sumVec (Vec n g)
  = for n 0 (λi x → x + g i)

dotVec :: (Syn a, Num a) ⇒ Vec a → Vec a → a

```

```

dotVec u v = sumVec (zipVec (×) u v)
normVec :: (Syn a, Num a) ⇒ Vec a → Dp Float
normVec v = sqrt (dotVec v v)

```

The vector representation makes it easy to define any functions when each vector element is computed independently, including *drop*, *take*, *reverse*, vector concatenation, and the like, but is less well suited to functions with dependencies between elements, such as computing a running sum.

An important consequence of the style of definition we have adopted is that it provides lightweight fusion. The definition of *dotVec* would not produce good C code if it first computed *zipVec (×) u v*, put the result into an intermediate vector *w*, and then computed *sumVec w*. Fortunately, it does not. Assume *u* is *Vec m g* and *v* is *Vec n h*. Then we can simplify *dotVec u v* as follows:

```

dotVec u v
  ~> sumVec (zipVec (×) u v)
  ~> sumVec (zipVec (×) (Vec m g) (Vec n h))
  ~> sumVec (Vec (m 'min' n) (λi → g i × h i))
  ~> for (m 'min' n) (λi x → x + g i × h i)

```

Indeed, we can see that by construction that whenever we combine two primitives the intermediate vector is always eliminated, a stronger guarantee than provided by conventional optimising compilers.

The type class *Syn* enables conversion between types *Arr* and *Vec*. Hence for EDSL, unlike QDSL, explicit calls *toVec* and *fromVec* are not required. Invoking *edsl normVec* produces the same C code as in Section 2.6.

As with QDSL, there are some situations where fusion is not beneficial. We may materialise a vector as an array with the following function.

```

memorise :: Syn a ⇒ Vec a → Vec a
memorise (Vec n g)
  = Vec n (λi → fromDp (ArrIx (Arr n (toDp ∘ g)) i))

```

The above definition depends on common subexpression elimination to ensure *Arr n (toDp ∘ g)* is computed once, rather than once for each element of the resulting vector. If

```

blur :: Syn a ⇒ Vec a → Vec a

```

averages adjacent elements of a vector, then one may choose to compute either

```

blur ∘ blur   or   blur ∘ memorise ∘ blur

```

with different trade-offs between recomputation and memory usage.

EDSL silently converts between representations, while QDSL forces all conversions to be written out; following the pattern that EDSL is more compact, while QDSL is more explicit. For QDSL it is the subformula property which guarantees that all intermediate uses of *Vec* are eliminated, while for EDSL this is established by operational reasoning on the behaviour of the type *Vec*.

6. Related work

DSLs have a long and rich history (Bentley 1986). An early use of quotation in programming is Lisp (McCarthy 1960), and perhaps the first application of quotation to domain-specific languages is Lisp macros (Hart 1963).

This paper uses Haskell, which has been widely used for EDSLs Hudak (1997); Reid et al. (1999); Bjesse et al. (1998); Gill (2014). We contrast QDSL with an EDSL technique that combines deep and shallow embedding, as described by Svenningsson and Axelsson (2012), and as used in several Haskell EDSLs including

Feldspar (Axelsson et al. 2010), Obsidian (Svensson et al. 2011), Nikola (Mainland and Morrisett 2010), Hydra (Giorgidze and Nilsson 2011), and Meta-Repa (Ankner and Svenningsson 2013).

ODonnell (1993) identified loss of sharing in the context of embedded circuit descriptions. Claessen and Sands (1999) and Gill (2009) propose solutions to this problem via extensions to Haskell that permit observable sharing.

A proposition-as-types principle for quotation as a modal logic was proposed by Davies and Pfenning (1996, 2001). As they note, their technique has close connections to two-level languages (Nielson and Nielson 2005).

Other approaches to DSL that make use of quotation include C# and F# versions of LINQ under .NET (Meijer et al. 2006; Syme 2006) and Scala Lightweight Modular Staging (Rompf and Odersky 2010). The underlying idea for QDSLs was established for F# LINQ by Cheney et al. (2013).

TODO: Mention the work on normalisation with effects: Cooper (DBLP 2006), Cheney and Lindley (TLDI 2012), Cheney, Lindley, Radanne, Wadler (PEPM 2014)

TODO: Mention the shredding paper: Cheney, Lindley, and Wadler (SIGMOD 2014)

TODO: We should say a little more about what novel features LMS brings to the table that are not available / harder to achieve in Haskell.

7. Conclusion

A good idea can be much better than a new one.

– Gerard Berry

We have compared EDSLs and QDSLs, arguing that QDSLs offer competing expressiveness and efficiency. EDSLs often (but not always) mimic the syntax of the host language, and often (but not always) perform normalisation in the host languages, while QDSLs (always) steal the syntax of the host language, and (always) ensure the subformula property, at the cost of requiring a normaliser, one per host language.

The subformula property may have applications in DSLs other than QDSLs. For instance, after Section 5.7 of this paper was drafted, it occurred to us that a different approach to options in EDSL would be to extend type *Dp* with constructs for type *Maybe*. So long as type *Maybe* does not appear in the input or output of the program, a normaliser that ensures the subformula property could guarantee that C code for such constructs need never be generated.

As we noted in the introduction, rather than build a special-purpose tool for each QDSL, it should be possible to design a single tool for each host language. Our next step is to design Haskell QDSL, with the following features.

- Type inference for the terms returned from typed quasi-quotations, restoring type information currently discarded by GHC.
- Based on the above, full support for type classes and overloading within quasi-quotation.
- The user may choose either an ADT or GADT representation of the term returned by typed quasi-quotation, whichever is more convenient.
- A normaliser to ensure the subformula property, which works with any datatype declared in Haskell.
- The user may supply a type environment indicating which constants (or free variables) may appear in typed quasi-quotations.

Such a tool could easily subsume the special-purpose translator from *Qt* to *Dp* described at the beginning of Section 3, and lift most of its restrictions. For instance, the current prototype is restricted to the *Maybe* monad, while the envisioned tool will work with any monad.

sam: Say something about adding support for observable sharing hacks?

Molière’s Monsieur Jourdain was bemused to discover he had been speaking prose his whole life. Similarly, many of us have used QDSLs for years, if not by that name. DSL via quotation is the heart of Lisp macros, Microsoft LINQ, and Scala LMS, to name but three. We hope that by naming the concept and drawing attention to the central benefits of normalisation and the subformula property, we may help the concept to flower further for years to come.

Acknowledgement Najd is a recipient of the Google Europe Fellowship in Programming Technology, and this research is supported in part by this Google Fellowship. Svenningsson is a SICS Visiting Fellow and is funded by a HiPEAC collaboration grant and by the Swedish Foundation for Strategic Research, under grant RawFP. Lindley and Wadler are funded by EPSRC Grant EP/K034413/1.

8. References

- J. Ankner and J. Svenningsson. An EDSL approach to high performance Haskell programming. In *Haskell*. ACM, 2013.
- E. Axelsson, K. Claessen, G. Dévai, Z. Horváth, K. Keijzer, B. Lyckegård, A. Persson, M. Sheeran, J. Svenningsson, and A. Vajdax. Feldspar: A Domain Specific Language for Digital Signal Processing Algorithms. In *MEMOCODE*, 2010.
- J. Bentley. Programming pearls: little languages. *Communications of the ACM*, 29(8):711–721, 1986.
- P. Bjesse, K. Claessen, M. Sheeran, and S. Singh. Lava: Hardware design in Haskell. In *ICFP*. ACM, 1998.
- M. Chakravarty, G. Keller, S. Peyton Jones, and S. Marlow. Associated types with class. In *POPL*. ACM, 2005.
- J. Cheney, S. Lindley, and P. Wadler. A practical theory of language-integrated query. In *ICFP*. ACM, 2013.
- K. Claessen and D. Sands. Observable sharing for functional circuit description. In *ASIAN*. Springer, 1999.
- K. Claessen, M. Sheeran, and B. Svensson. Expressive array constructs in an embedded GPU kernel programming language. In *DAMP*. ACM, 2012.
- Rowan Davies and Frank Pfenning. A modal analysis of staged computation. In Jr. Guy Steele, editor, *Principles of Programming Languages (POPL)*, pages 258–270. St. Petersburg Beach, Florida, 1996. ACM.
- Rowan Davies and Frank Pfenning. A modal analysis of staged computation. *Journal of the ACM*, 48(3):555–604, 2001.
- Cormac Flanagan, Amr Sabry and Bruce F Duba, and Matthias Felleisen. The essence of compiling with continuations. In *Programming Language Design and Implementation (PLDI)*, pages 237–247. ACM, 1993.
- Gerhard Gentzen. Untersuchungen über das logische schließen. i. *Mathematische zeitschrift*, 39(1):176–210, 1935.
- A. Gill. Type-safe observable sharing in Haskell. In *Haskell*. ACM, 2009.
- A. Gill. Domain-specific languages and code synthesis using Haskell. *Commun. ACM*, 57(6), 2014.
- G. Giorgidze and H. Nilsson. Embedding a functional hybrid modelling language in Haskell. In *IFL*. Springer, 2011.
- Timothy P. Hart. MACRO definitions for LISP. Technical Report AIM-057, MIT, 1963.
- W. A. Howard. The formulae-as-types notion of construction. In *To H. B. Curry: Essays on Combinatory Logic, Lambda Calculus, and Formalism*, pages 479–491. Academic Press, 1980. The original version was circulated privately in 1969.
- P. Hudak. Domain-specific languages. *Handbook of Programming Languages*, 3, 1997.
- J. Hughes. Restricted data types in Haskell. In *Haskell*. ACM, 1999.
- S. Lindley. Extensional rewriting with sums. In *TLCA*. Springer, 2007.

- G. Mainland. Explicitly heterogeneous metaprogramming with Meta-Haskell. In *International Conference on Functional Programming (ICFP)*. ACM, 2012.
- G. Mainland and G. Morrisett. Nikola: embedding compiled GPU functions in Haskell. In *Haskell*. ACM, 2010.
- J. Maraist, M. Odersky, and P. Wadler. The call-by-need lambda calculus. *JFP*, 8(03):275–317, 1998.
- John McCarthy. Recursive functions of symbolic expressions and their computation by machine, part i. *Communications of the ACM*, 3(4):184–195, 1960.
- E. Meijer, B. Beckman, and G. Bierman. LINQ: reconciling object, relations and XML in the .NET framework. In *SIGMOD*. ACM, 2006.
- Eugenio Moggi. Notions of computation and monads. *Information and Computation*, 93(1):55–92, 1991.
- Flemming Nielson and Hanne Riis Nielson. *Two-level functional languages*, volume 34. Cambridge university press, 2005.
- J. O'Donnell. Generating netlists from executable circuit specifications in a pure functional language. In *Functional Programming*. Springer, 1993.
- A. Persson, E. Axelsson, and J. Svenningsson. Generic monadic constructs for embedded languages. In *IFL*, 2011.
- F. Pfenning and C. Elliot. Higher-order abstract syntax. In *PLDI*, 1988.
- Dag Prawitz. *Natural Deduction: A Proof-Theoretical Study*. Almqvist and Wiksell, Stockholm, 1965.
- A. Reid, J. Peterson, G. Hager, and P. Hudak. Prototyping real-time vision systems: An experiment in DSL design. In *ICSE*. ACM, 1999.
- T. Rompf and M. Odersky. Lightweight modular staging: a pragmatic approach to runtime code generation and compiled DSLs. In *GPCE*. ACM, 2010.
- N. Sculthorpe, J. Bracker, G. Giorgidze, and A. Gill. The constrained-monad problem. In *ICFP*. ACM, 2013.
- J. Svenningsson and E. Axelsson. Combining deep and shallow embedding for EDSL. In *TFP*. Springer, 2012.
- J. Svenningsson and B. Svensson. Simple and compositional reification of monadic embedded languages. In *ICFP*. ACM, 2013.
- J. Svensson, M. Sheeran, and K. Claessen. Obsidian: A domain specific embedded language for parallel programming of graphics processors. In *IFL*. Springer, 2011.
- Don Syme. Leveraging .NET meta-programming components from F#: integrated queries and interoperable heterogeneous execution. In *ML*. ACM, 2006.
- Philip Wadler. Propositions as types. *Communications of the ACM*, 2015. To appear.
- Limsoon Wong. Normal forms and conservative properties for query languages over collection types. In *Principles of Database Systems (PODS)*, pages 26–36. ACM, 1993.