Abstract

We present a library for expressing digital signal processing (DSP) algorithms using a deeply embedded domain-specific language (EDSL) in Haskell. The library supports definitions in functional programming style, reducing the gap between the mathematical description of streaming algorithms and their implementation. The deep embedding makes it possible to generate efficient C code without any overhead associated with the high-level programming model. The signal processing library is intended to be an extension of the Feldspar EDSL which, until now, has had a rather low-level interface for dealing with synchronous streams. However, the presented library is independent of the underlying expression language, and can be used to extend any pure EDSL for which a C code generator exists with efficient stream processing capabilities. The library is evaluated using example implementations of common DSP algorithms and the generated code is compared to its handwritten counterpart.

1. Introduction

In recent years, the amount of traffic passing through the global communications infrastructure has been increasing at a rapid pace. Worldwide, total Internet traffic is estimated to grow at an average rate of 32% annually, reaching approximately eighty million terabytes per month by the end of next year [16]. Mobile communications in particular have been growing at a phenomenal rate, which can be largely attributed to the rising popularity of mobile terminals.

For telecommunications infrastructure, the consequences of such a rapid growth rate have been a dramatic increase in the demand for network capacity and computational power [1]. At the same time, telecom carriers are faced with an increasing need to deliver new services faster, while simultaneously adapting the the recent diversification in available architecture. These factors, while positively influencing the available computational power, have also significantly increased the complexity of developing new solutions for telecommunication systems.

Today, digital signal processing software is typically implemented in low level C, which forces designers to focus on low-level implementation details rather than the mathematical specification of the algorithms. Our group is developing an embedded domain-specific language (EDSL), Feldspar [3], that aims to raise the abstraction level of signal processing software by expressing algorithms as pure functional programs.

However, signal processing is more than just pure computations – it is also about how to connect those functions in a network that operates on streaming data. A suitable programming model for reactive systems that process streams of data is synchronous dataflow (SDF) [19], which offers natural, high-level descriptions of streaming algorithms, while still permitting the generation of efficient code. Feldspar does have a library for programming with synchronous streams, but that library is quite low-level and tedious to use.

This paper describes a library for extending an existing Haskell EDSL with support for SDF. The underlying EDSL is used to represent pure functions (which, of course, can be arbitrarily complicated), and our library gives a means to connect such functions using an SDF programming model. If the underlying EDSL provides a C code generator with a given interface, our library is capable of emitting C code for SDF programs. We are interested in using Feldspar as the expression language; however, the library is not dependent on Feldspar, and so may be of interest to other EDSL developers.

This paper makes the following contributions:

• We present a simple EDSL for synchronous dataflow programming in Haskell. Practically, the result is a useful addition to Feldspar.

• We make use of observable sharing [7] to achieve a deep embedding without relying on combinators to express sharing or cycles. This technique has long been used in the hardware description EDSL Lava [6, 11], but our work now permits the combination not just of simple gates but of arbitrarily complex EDSL programs.

• We abstract away from the underlying expression language by establishing an interface for the underlying expression compiler.

• We show how to generate C code from our library, and evaluate the results on examples.

1.1 Synchronous dataflow programming

Dataflow programming is a paradigm which internally models an application as a directed graph [17, 21], similar to a
dataflow diagram. Nodes in the graph are then executable blocks, representing the different components of an application: they receive input, apply some transformation, and forward it to the other connected nodes. A dataflow application is then, simply stated, a composition of such blocks, with one or more source and sink blocks.

A later extension to dataflow programming is the introduction of synchronous dataflow. SDF is a subset of pure dataflow, in which the number of tokens produced or consumed by nodes during each step of evaluation is known at compile-time. Restricting the dataflow model in this way has the advantage that it can be statically scheduled [18], which, in turn, allows for generation of efficient code.

Lucid Synchrone [8, 20] is a member of the family of synchronous languages and is designed to model reactive systems. It was introduced as an extension of LUSTRE [13], and demonstrated that the language could be extended with new and powerful features. For instance, automatic clock and type inference was introduced, and a restricted form of higher-order functions was added. However, Lucid Synchrone is a standalone language which cannot easily be integrated with EDSLs such as Feldspar. For this reason, we chose to implement a library, partly inspired by Lucid Synchrone, that brings an SDF programming model to existing EDSLs, such as Feldspar.

2. Signal

This library is based on the concept of a signal, which represents an infinite sequence of values in some pure expression language. Signals are constructed by the following interface:

\[
\begin{align*}
\text{map} & : (\text{exp } a \rightarrow \text{exp } b) \rightarrow \text{Signal } \text{exp } a \rightarrow \text{Signal } \text{exp } b \\
\text{repeat} & : \text{exp } a \rightarrow \text{Signal } \text{exp } a \\
\text{zip} & : \text{Signal } \text{exp } a \rightarrow \text{Signal } \text{exp } b \rightarrow \text{Signal } \text{exp } (a, b) \\
\text{fst} & : \text{Signal } \text{exp } (a, b) \rightarrow \text{Signal } \text{exp } a
\end{align*}
\]

where \( \text{exp} \) is the pure expression language. The \( \text{map} \) function promotes a pure function to operate over signals; \( \text{repeat} \) makes a constant-valued signal; \( \text{zip} \) and \( \text{fst} \) are used to make nodes with multiple incoming or outgoing signals.

Sequential operations are supported through the following functions, which manage a signal’s phase and frequency:

\[
\begin{align*}
\text{delay} & : \text{exp } a \rightarrow \text{Signal } \text{exp } a \\
\text{sample} & : \text{exp } \text{Int} \rightarrow \text{Signal } \text{exp } a \\
\end{align*}
\]

While few in number, these sequential functions are quite general and allow for arbitrary feedback networks to be expressed.

The need to implement particular signal functions may place demands on the underlying expression language, in that support for common data types or functionality may be required. For instance, in order to implement a signal version of Haskell’s \( \text{zipWith} \) function, the expression language needs to support tuples:

\[
\text{class TupExp exp where}
\begin{align*}
\text{tup} :& : \text{exp } a \rightarrow \text{exp } b \rightarrow \text{exp } c \\
\text{fst} :& : \text{exp } (a, b) \rightarrow \text{exp } a \\
\text{snd} :& : \text{exp } (a, b) \rightarrow \text{exp } b
\end{align*}
\]

\[
\text{zipWith} :: (\text{TupExp exp}, \text{Signal exp} \sim \text{sig}) \\
& \Rightarrow (\text{exp } a \rightarrow \text{exp } b \rightarrow \text{exp } c) \\
& \rightarrow \text{sig } a \rightarrow \text{sig } b \rightarrow \text{sig } c
\]

\[
\text{zipWith } f s u = \text{map } (\lambda p \rightarrow f (\text{fst } p) (\text{snd } p)) \$_{\text{zip }} s u
\]

Classes such as \( \text{TupExp} \) provide a suitable interface with the expression language, but without forcing a particular language to be hardwired into the system.

2.1 Example: FIR Filter

Consider the mathematical definition of a finite impulse response filter of rank \( N \):

\[
y_n = \sum_{i=0}^{N} b_i * x_{n-i}
\]

This description is convenient for software realization, as it can be deconstructed into three main components: a number of successive unit delays, multiplication with coefficients and a summation. We can represent the decomposed filter graphically, as in Figure 1.

Support for such numerical operations over signals is implemented by instantiating their corresponding classes in Haskell:

\[
\text{instance } (\text{TupleExp exp}, \text{Num } (\text{exp } a)) \Rightarrow \\
\text{Num } (\text{Signal exp } a)
\]

\[
\begin{align*}
\text{fromInteger} &= \text{repeat } . \text{fromInteger} \\
(+), (-) &= \text{zipWith } (+), (-)
\end{align*}
\]

where similar instance declarations can be made for fractional and floating point arithmetic.

Using Haskell’s standard classes in this way simplifies the construction of signals by providing a homogeneous user interface. Furthermore, as the pure Haskell code is separated from our signal library in this way, it introduces a meta-level of computation, helping us to reason about program
correctness. The ability to use pure Haskell in this way presents several benefits, as it improves the syntax and ease of programming signals significantly. For instance, in order to define complex networks, the user is only required to be versed in Haskell’s standard library operators, thereby further reducing the complexity of developing new networks.

Given this support for numerical operations, we now create helper functions, modeling the three main components of the FIR filter: summation and multiplication of signals and successive delaying. Using Haskell’s standard library functions, summation can be neatly expressed as a single fold operation:

```haskell
import qualified Prelude as P

sums :: ( TupleExp exp , Num ( exp a)) ⇒ [ Signal exp a] → Signal exp a
sums = P. foldr1 (+)
```

Similarly, both of the remaining components can be expressed using standard Haskell functions:

```haskell
muls :: ( TupleExp exp , Num ( exp a)) ⇒ [ exp a] → [ Signal exp a] → [ Signal exp a]
muls = P. zipWith (λ c s → repeat c * s)

delays :: [ exp a] → Signal exp a → [ Signal exp a]
delays as s = P. tail $ P. scanl (P. flip delay ) s as
```

The FIR filter can now be neatly expressed as:

```haskell
fir :: ( TupleExp exp , Num ( exp Float)) ⇒ [ exp Float ] → Signal exp Float → Signal exp Float
fir bs = sums . muls bs . delays ds
where
ds = P. replicate (P. length bs) 0
```

This description is close to the filter’s graphical representation, a beneficial attribute since domain experts in DSP tend to be comfortable with composing sub-components in this way.

2.2 Example: IIR Filter

Infinite impulse response (IIR) filters are digital filters with an infinite impulse response and, unlike FIR filters, contain feedback. These filters will therefore serve as an example of how the signal library handles recursively defined signals, that is, signals whose output depends on a combination of previous input and output values.

The IIR filter is often described and implemented in terms of a difference equation, which defines how the output signal is related to the input signal:

\[
y_n = \frac{1}{a_0} \left( \sum_{i=0}^{P} b_i x_{n-i} - \sum_{j=1}^{Q} a_j y_{n-j} \right)
\]

where \( P \) and \( Q \) are the feedforward and feedback filter orders, respectively, and \( a_j \) and \( b_i \) are the filter coefficients. We can represent the decomposed filter graphically, as in Figure 2.

![Figure 2. A direct form discrete-time IIR filter of order \( P \) and \( Q \)](image)

This description, besides the subtraction and division of signals, is quite similar to the previous FIR filter when deconstructed. This similarity seems to imply that they share computational components. As it turns out, the previously defined helper functions can indeed be reused to implement the IIR filter as well:

```haskell
iir :: ( TupleExp exp , Num ( exp Float) , Fractional ( exp Float)) ⇒ [ exp Float] → Signal exp Float → Signal exp Float
iir (a:as) bs s = repeat (1 / a) * (l - r)
where
  l = sums $ muls bs $ delays (inits bs) s
  r = sums $ muls as $ delays (inits as) r
  inits = P. flip P. replicate 0 . P. length
```

where the rightmost summation, here called \( r \), is defined in terms of itself rather than the input signal. Recursive definitions like this are made possible by the lazy nature of the delay operator. The general idea is that any recursion expressed using the signal library introduces feedback, while recursion introduced by pure Haskell code produces repetitive code instead.

3. Implementation

Signals are implemented on top of the following `Stream` data type:

```haskell
data Stream exp a = Stream (Prog exp (Prog exp a))
```

The `Prog exp` monad is a deep embedding of an imperative programming language that uses `exp` to represent pure expressions. In the above definition, the outer monad is used to initialize the stream, and the inner action is used to retrieve the next value of the stream. For example, the `head` function, which retrieves the first element of a stream is defined as follows:
head :: Stream exp a → Prog exp a
head (Stream init) = do
  next ← init
  next

The first line in the do block initializes the next action, and the second line uses next to get the first element.

Our model of streams is essentially the same as in Feldspar’s Stream library, except that the Prog exp monad used here is a standalone monad that adds imperative programming on top of any pure expression EDSL.

The problem with Stream is that is quite low-level and cumbersome to program with. For example, in order to define a filter that refers to previous values of some signal, one has to manually create a mutable buffer and update it on every iteration. Feldspar exports a few combinators that hide the details of creating such networks, but these combinators are rather ad hoc, and can only handle a few predefined cases.

Our Signal type can be seen as a front-end to Stream that offers a much more convenient interface. In the end, functions on signals are compiled to functions on streams, as seen in the type of the compile function:

\[
\text{compile} :: (\text{Typeable } a, \text{Typeable } b) \\
\Rightarrow (\text{Signal } exp \ a \rightarrow \text{Signal } exp \ b) \\
\rightarrow \text{IO} (\text{Stream } exp \ a \rightarrow \text{Stream } exp \ b)
\]

The IO in the result type comes from the data-reify package that performs observable sharing [10].

The basic way to create nodes in a signal network is by lifting a stream function to a signal function:

\[
\text{lift} :: (\text{Stream } exp \ a \rightarrow \text{Stream } exp \ b) \\
\rightarrow \text{Signal } exp \ a \\
\rightarrow \text{Signal } exp \ b
\]

We can, for example, use lift to define map from Section 2:

\[
\text{map } f = \text{lift} \,(\text{Stream.map } f)
\]

Lifting is however not enough to define generators, as those are supposed to create signals from nothing. Another signal construct is therefore introduced, modeling the signals identity morphism:

\[
\text{bot} :: \text{Signal } exp \ a
\]

which allows us to define repeat as:

\[
\text{repeat } e = \text{lift} \,(\text{const } \& \,\text{Stream.repeat } e) \,\text{bot}
\]

Here, Stream.map and Stream.repeat are the corresponding functions defined for streams.

In this extended abstract, we will not show the definition of Signal, lift and compile, but here is an outline of how it all works:

- **Signal** is a tree type, whose nodes consist of lifted stream functions, delay and sample. Observable sharing is used to turn this tree into a DAG.
- **The compiler assigns a mutable reference (supported by the imperative Prog monad) for each node in the graph and creates a program that executes all nodes in sequence, reading and writing data to the corresponding references.
- **By analyzing the graph, certain references can be eliminated**, and chains of delay nodes can be turned into efficient cyclic buffers.

While the stream type is kept abstract in signals, basing them on the co-iterative approach [5] allows us to ensure that no unused computations are performed. Co-iteration is a concept for reasoning about infinite streams and allows one to handle such streams in a strict manner. This strictness in turn enables stream transformers to pick and remove parts of streams as they please – ensuring that no unused part of a stream is ever computed. This property is kept for signals, as lifted nodes are fused during compilation.

Furthermore, signals offer optimization for a common concept in DSP: feedback networks. As the recursively defined streams in feedback networks may reference a number of previous values, memory efficient buffers are introduced for storing the delayed values in these signals. These buffers have memory proportional in size to the number of delays and are used to minimize the number of read/write operations used during execution. Detecting such feedback in signal networks is made possible by using observable sharing [7], which allows us to reify signal networks into graphs were the back-edges between nodes are visible.

4. Evaluation

In order to evaluate the signal library we will look at both its expressiveness and the code it generates. As the actual library is being finished at the time of writing, we delay most of the evaluation until the final report and only include a comparison between the generated code of an example and its hand-written counterpart. Note, too, that the compiler has some notable limitations: it currently only handles pure signal functions, that is, we can only compile types that hide the details of creating such networks, but these combinators are rather ad hoc, and can only handle a few predefined cases.

Disregarding the current state of our compiler, the produced code has room for several potential improvements. Firstly, too many references are used during the compilation process, which then spills into the compiled code and produces a number of unused variables. There are also some leftover pairs found in the produced code, remnants from the zipping constructor. Both the numerous variables and the pairs can however be optimized away, and should not be present in the final report. There is also room for more subtle improvements; for example, a dedicated initialization function could remove the need for some if-statements. Another interesting improvement is to make use of C specific memory management functions, such as memset or memcpy, for cases when the data is neatly ordered – as it is in the case of the FIR filter, for example.

For the actual comparison, we generated code from the previous FIR filter and compared it to a hand-written version. However, as the compiler only accepts signal functions, we feed the FIR filter a one-element list of ones before compiling it – effectively making it a rank-1 FIR filter. This produces the following code:
typedef struct {
    float first;
    float second;
} pair;

int main()
{
    FILE *v0;
    v0 = fopen("test", "r");
    FILE *v1;
    v1 = fopen("test2", "w");

    int i = 0;
    while (i < 3)
    {
        float v3;
        float v4;
        fscanf(v0,"%f",v4);
        float v5 = v4;
        float v6;
        pair v7;
        float v8;
        float v9;
        bool v10 = true;
        float v11 = v5;
        bool v12;
        if (v12 = v10)
        {
            v10 = false;
            v8 = 0.0;
        }
        else
        {
            float v13 = v9;
            v8 = v13;
        }
        v9 = v11;
        float v14;
        v14 = 1.0;
        float v15 = v14;
        float v16 = v8;
        v7.first = v15;
        v7.second = v16;
        pair v17 = v7;
        v6 = v17.first * v17.second;
        float v18 = v6;
        v3 = v18;
        float v19 = v3;
        fprintf(v1,"%f", v19);
        i++;
    }

    return 0;
}

While the hand-written code is for a more general FIR filter, where the number of coefficients hasn’t been fixed yet, the comparison should however still be valid as the underlying ideas have not changed.

double insamp[ ... ];

void firInit( void )
{
    memset( insamp, 0, sizeof( insamp ) );
}

void fir( double *coeffs,
        double *input, double *output,
        int length, int filterLength )
{
    double acc; // accumulator for MACs
    double *coeffp; // pointer to coefficients
    double *inputp; // pointer to input samples
    int n;
    int k;

    // put the new samples at the high
    // end of the buffer
    memcpy( &insamp[filterLength - 1], input,
            length * sizeof( double ) );

    // apply the filter to each input sample
    for ( n = 0; n < length; n++ ) {
        // calculate output n
        coeffp = coeffs;
        inputp = &insamp[filterLength - 1 + n];
        acc = 0;
        for ( k = 0; k < filterLength; k++ ) {
            acc += (*coeffp++) * (*inputp--);
        }
        output[n] = acc;
    }

    // shift input samples back in time
    // for next time
    memmove( &insamp[0], &insamp[length],
             (filterLength - 1) * sizeof(double) );
}

While benchmarking would clearly be of interest here, we delay it to the final paper.

5. Related Work

Lava is a family of simple hardware description languages embedded in Haskell [6, 11]. What look like circuit descriptions are actually circuit generators that are run to produce internal representations, which can be used to produce further useful artifacts such as netlists or formulas for use in formal verification. Later versions of Lava have used observable sharing to provide a more user-friendly approach to capturing feedback in these circuit descriptions. There is a close link between this approach to hardware description and SDF languages like LUSTRE. One way to view the present work is as a beefed up Lava in which the building blocks are general data-processing functions rather than just Boolean gates.

Lucid Synchrone is another functional language for SDF and is hosted in OCaml, importing every ground type from the host language and lifting them to corresponding stream versions [8, 20]. Sequential operations over the imported stream are also offered, similar to those from our signal library. Lucid Synchrone incorporates several type systems
(including clock inference) that guarantee safety properties of the generated code. In addition, special syntax for defining automata is provided. Our work is inspired by Lucid Synchrone and we will investigate the inclusion of these features in our signal library.

Functional reactive programming is another common paradigm for modeling continuous signals, and its libraries are typically implemented using Haskell’s arrow or applicative classes, see for example Yampa [9, 12], reactive-banana [2] or Sodium [4]. Although we cannot support the promotion of abstract functions required by these otherwise attractive interfaces, we have still drawn inspiration from the lifting functions of FRP.

The use of pure Haskell for modeling signals [22] has also been investigated, demonstrating that functional programming and lazy evaluation can directly model common signal problems quite satisfactorily. Other related work includes Atom [14], Ivory and Tower [15], but we delay the discussion of these to the final paper.

6. Discussion

Functional programming encourages a style of programming in which combinators or higher order functions capture common patterns of computation. It is when we combine SDF with a sufficiently general value type that the question of how to design an appropriate set of combinators becomes an interesting one. For instance, the ability to pass arrays, or any similar data type, as values over signals means that the programmer is concerned with processing chunks of data rather than just individual values. This change of perspective is necessary if we are to implement key functions like FFT on signals. This generality does however come at a cost, as ill-defined signals could be expressed and type-checked due to a lack of constraints on sequential signals.

The benefits of a general lifting constructor come into full effect when the underlying expression language has powerful features of its own. In the case of Feldspar, an expression language which already supports a plethora of different combinators, several complex signal functions can often be obtained by simply lifting the regular ones. For instance, as Feldspar already contains an FFT algorithm over vectors, a signal version can be obtained by simply lifting the existing one.

The simplicity of the current signal library does mean that some ill-defined signals can be expressed. Sequential operations in particular are quite susceptible to grammatical errors, as one can easily create signals with an undefined behavior when using delay/sample. For instance, consider the following function:

\[
\begin{align*}
f &:: \text{Signal} \times a \to \text{Signal} \times a \\
f \ \text{sig} &= \text{sample \ 2 \ sig} + \ \text{sig}
\end{align*}
\]

The clocks of these two streams are obviously not equal, but there are at present no constraints in place to keep such signals for being defined. Recursively defined signals suffer from a similar problem: there is no constraint in place to check whether the recursive signal has been delayed or not before it is used. Signals with undefined initial values can therefore be expressed by, for example, writing:

\[
\begin{align*}
g &:: \text{Num} \times (\exp \times a) \Rightarrow \text{Signal} \times (\exp \times a) \\
g \ \text{o} &\text{ where o = map \ (+1) \ o}
\end{align*}
\]

Such ill-defined signals could however be identified during the compilation process, by inspecting the signal’s reified syntax tree.

Our plan is to develop this library further, based on ideas from Lucid Synchrone and FRP.

Acknowledgments

This research was funded by the Swedish Foundation for Strategic Research (in the RAW FP project) and by the Swedish Research Agency (Vetenskapsrådet).

References


