A Formally Verified Compiler for Lustre

Timothy Bourke\textsuperscript{1,2}
Timothy.Bourke@inria.fr

Lélio Brun\textsuperscript{2,1}
Lelio.Brun@ens.fr

Pierre-Évariste Dagand\textsuperscript{3,4,1}
Pierre-Evariste.Dagand@lip6.fr

Xavier Leroy\textsuperscript{1}
Xavier.Leroy@inria.fr

Marc Pouzet\textsuperscript{3,2,1}
Marc.Pouzet@ens.fr

Lionel Rieg\textsuperscript{5,6}
Lionel.Rieg@yale.edu

\textsuperscript{1} Inria, Paris, France
\textsuperscript{2} Département d’Informatique, École normale supérieure, PSL Research University, Paris, France
\textsuperscript{3} Sorbonne Universités, UPMC Univ Paris 06, France
\textsuperscript{4} CNRS, LIP6 UMR 7606, Paris, France
\textsuperscript{5} Collège de France, Paris, France
\textsuperscript{6} Yale University, New Haven, Connecticut, USA

Abstract

The correct compilation of block diagram languages like Lustre, Scad, and a discrete subset of Simulink is important since they are used to program critical embedded control software. We describe the specification and verification in an Interactive Theorem Prover of a compilation chain that treats the key aspects of Lustre: sampling, nodes, and delays. Building on CompCert, we show that repeated execution of the generated assembly code faithfully implements the dataflow semantics of source programs.

We resolve two key technical challenges. The first is the change from a synchronous dataflow semantics, where programs manipulate streams of values, to an imperative one, where computations manipulate memory sequentially. The second is the verified compilation of an imperative language with encapsulated state to C code where the state is realized by nested records. We also treat a standard control optimization that eliminates unnecessary conditional statements.

1. Introduction

Lustre was introduced in 1987 as a programming language for embedded control and signal processing systems [13]. It gave rise to the industrial tool SCADE Suite\textsuperscript{1} and can serve as a target to compile a subset of Simulink\textsuperscript{2} to executable code [15, 61]. SCADE Suite is used to develop safety-critical applications like fly-by-wire controllers and power plant monitoring software. Several properties make Lustre-like languages suitable for such tasks: constructs for programming reactive controllers, execution in statically-bounded time and memory, a mathematically well-defined semantics based on dataflow streams [13], traceable and modular compilation schemes [8], and the practicability of automatic program verification [17, 25, 30, 38] and industrial certification. These languages allow engineers to develop and validate systems at the level of abstract block diagrams that are compiled directly to executable code.

Compilation transforms sets of equations that define streams of values into sequences of imperative instructions that manipulate the memory of a machine. Repeatedly executing the instructions is supposed to generate the successive values of the original streams: but how is this to be ensured? Existing approaches apply translation validation [49, 52, 60] or industrial certification based on development standards and coverage criteria [50]. We take a different approach by formally specifying the source, target, and intermediate languages, their corresponding models of computation, and the compilation algorithms in the Coq Interactive Theorem Prover (ITP) [63]. We state and prove a correctness relation between the source and target semantic models, and build on the CompCert project [10, 40] to integrate the semantics.

\textsuperscript{1} http://www.ansys.com/products/embedded-software/ansys-scade-suite
\textsuperscript{2} http://www.mathworks.com/products/simulink/
of machine-level types and operators and to extend the correctness relation to the level of assembly code for PowerPC, ARM and x86 processors.

In this paper, we make the following contributions:

- The Vélus compiler that turns synchronous dataflow equations into assembly instructions and a proof of its correctness in the Coq ITP (Section 2). Our compiler supports the key features of the Lustre [13] language. Initial experimental results are encouraging (Section 5).
- We identify the invariants, lemmas, and proof structures necessary to verify code generation, providing a blueprint for performing similar work in any ITP. Specifically, we (1) introduce a novel semantic model combining the infinite sequences of dataflow models with the incremental manipulation of memories that characterizes imperative models (Section 3.2), and (2) exploit separating assertions to concisely and modularly relate an idealized memory model to the lower-level model of CompCert’s Clight (Section 4).
- We follow the architecture of the industrial SCADE Suite compiler, including modular code generation (one sequential function per dataflow node) and an optimization to fuse conditionals (Section 3.3). The verification of the optimization uses properties of the source language and translation function and shows the utility of our formalization.
- Our compiler front-end is implemented and verified parametrically with respect to an abstract interface of operators (Section 4.1). This approach not only facilitates proof, it also allows our tool chain to be instantiated with different backends and their semantic models.

The outcome of our work is twofold. First, our Coq development extracts to OCaml, providing us with an executable compiler. Second, we have mechanically verified that for any node $f$ in a Lustre program $G$ accepted by our compiler, for any well-typed input stream $xs$ and corresponding output stream $ys$, the semantics of the generated assembly code is an infinite execution trace that is bisimilar to the trace that reads the stream of input values from, and writes the expected output values to, volatile variables:

$$G^\text{node} f (xs, ys)$$

$$\text{compile } G f = \text{OK asm}$$

$$\exists T, \text{asm} \Downarrow T \land \langle VLoad(xs(n)), VStore(ys(n)) \rangle_{n=0}^\infty \sim T$$

This proof is obtained by composing CompCert’s original correctness result, from Clight to assembly, and the novel correctness result from Lustre to Clight that we present in this paper. As in standard programming practice, we combine and extend existing systems to produce new functionality. By doing so in an ITP, we guarantee that the different parts function together correctly.

2. Lustre and its Compilation

A Lustre program comprises a list of nodes. A node is a named function between lists of input and output streams, defined by a set of equations. Consider this example adapted from the original paper [13]:

```latex
\text{node} \counter(\text{ini}, \text{inc}; \text{res}; \text{bool}) \text{returns} (n: \text{int}) \text{let}
\text{n} = \text{if} (\text{true fby} \text{false}) \text{or} \text{res} \text{then} \text{ini} \text{else} (0 \text{fby} n) + \text{inc}; \text{tel}
```

This node counter takes three input streams—two of integers and one of booleans—and returns a stream of integers. It calculates the cumulative sum of values on inc taking the value of ini initially or when res is true. In graphical editors, like those of SCADE Suite or Simulink, this node would be represented by a rectangle labeled with the node name and having three input ports and one output port.

The value of the output $n$ is defined by a single equation that freely mixes instantaneous operators (or and +), a multiplexer (if/then/else: both branches are active and the value of the guard selects one of the results), and the initialized delay operator ‘fby’ (“followed by”). The initialized delay corresponds to the flip-flop of sequential logic and the $z^{-1}$ of Digital Signal Processing. The subexpression true fby false defines the stream $T, F, F, F, \ldots$, and the subexpression 0 fby $n$ defines a stream $c$ such that $c(0) = 0$ and $\forall i > 0, c(i) = n(i - 1)$.

We describe the other features of Lustre, namely node instantiations and the when and merge sampling constructs, after outlining the compiler architecture.

2.1 Compiler Architecture and Implementation

In the classical approach to compiling Lustre [13, 28], all nodes are inlined and a control automaton is constructed to minimize the code executed in a given state. In the clock-directed modular approach [8], each dataflow node is compiled to a distinct sequential function and each equation is assigned a static clock expression that becomes a conditional in the imperative code. 3 This approach is used in the industrially certified SCADE Suite compiler and academic compilers [23, 35]. It is the one addressed in our work.

Figure 1 outlines the successive source-to-source transformations of a clock-directed modular Lustre compiler.

The first four stages deal with dataflow programs. Parsing turns a source file into an Abstract Syntax Tree (AST) without type or clock annotations. Elaboration rejects programs that are not well-typed and well-clocked, and otherwise yields an (annotated) Lustre program that can be assigned a semantics. Normalization ensures that every fby expression and node instantiation occurs in a dedicated equation and not nested arbitrarily within an expression. Scheduling sorts the dataflow equations by variable dependencies in anticipation of their translation one by one into imperative assignments: variables must be written before they are read, except those defined

\[ \text{Various trade-offs of inlining and modularity are possible [43, 55].} \]
Normalization and scheduling change the syntactic form of a program while preserving its dataflow semantics. Normalization is allowed because Lustre is referentially transparent: a variable can always be replaced by its defining expression and conversely. Scheduling is allowed because the meaning of a function is independent of their order. The normalized and scheduled example appears at left in Figure 3.

Both normalization and scheduling were verified in Coq in prior work [2, 3]. We do not dwell on them here but rather focus on the problems that were not previously addressed. We elaborate directly to N-Lustre and apply a formally validated scheduling algorithm to produce SN-Lustre.

Returning to Figure 1, the first of these problems is the translation from SN-Lustre to the intermediate imperative language Obc [8, §4]. This pass changes both syntax and semantics: from dataflow equations defining functions on streams to sequences of imperative assignments defining transitions between memory states. Examples of Obc programs are given in Section 3, where we detail the translation pass along with the statement and proof of its correctness.

The translation pass produces a nesting of conditional statements for each equation in the source program. This facilitates compiler specification and proof but gives inefficient code. A fusion optimization is thus normally applied to merge adjacent conditionals. Its specification and verification are detailed in Section 3.3.

The second significant problem addressed in our work is the generation of compilable code, typically C, Java, or Ada, from Obc. In our work we generate Clight, a subset of C accepted by the CompCert compiler, whose compilation to assembly code has been modeled and verified in Coq [10, 40]. This requires incorporating a subset of Clight’s types and operators into SN-Lustre programs. We exploit the module system of Coq to abstract over the details until the generation pass where they are instantiated with Clight-specific definitions.

Figure 1. Clock-directed modular compiler architecture: the gray boxes and the solid arrows are addressed in this paper. The white boxes and the dashed arrows are addressed in previous work [2, 3] or by CompCert [40].

Figure 2. SN-Lustre : abstract syntax.

Figure 3. SN-Lustre: example program.
the representation of program memories. The state of Obc programs is a tree of variable environments. It must be encoded as nested records in the target Clight program, and the concomitant details of alignment, padding, and aliasing must be confronted. The generation pass is described in Section 4.

Our prototype compiler implements the languages in the gray boxes and the transformations marked by solid arrows in Figure 1. The lexer is generated by ocamlflex [42, §12]. The parser is generated in Coq by Menhir [53] which also produces proofs of correctness and completeness [36]. The elaborator is a mix of OCaml, mostly from CompCert for parsing constants, and Coq, it rejects programs that are not normalized. We prove using standard techniques that successful elaboration yields a well-typed and well-clocked N-Lustre program. Scheduling is implemented in OCaml and validated in Coq. The translation and generation algorithms are implemented as Coq functions and extracted to OCaml. We generate a Clight AST that CompCert’s formally verified or validated algorithms compile into assembly code.

2.2 SN-Lustre, Node instantiation, and Sampling

The abstract syntax of SN-Lustre is given in Figure 2. Many of the elements have already been presented. We now describe node instantiation and the constructs for sampling and merging streams via the extended example of Figure 3.

The counter node is the normalized and scheduled version of the earlier example. The d_integrator node [28, II(C)] instantiates counter twice to calculate speed and position values from a stream of accelerometer readings. It is itself instantiated in the tracker node that returns the position and a count of the number of times that the speed exceeds a given limit. The instants of overshoot are detected as rising edges (rising(in) = ((false fly not in) and in) on the expression s > limit. Consider the values of the streams in this node for a given acc and with limit a constant stream of 5s:

<table>
<thead>
<tr>
<th>acc</th>
<th>0</th>
<th>2</th>
<th>4</th>
<th>-2</th>
<th>0</th>
<th>3</th>
<th>-3</th>
<th>2</th>
<th>...</th>
</tr>
</thead>
<tbody>
<tr>
<td>limit</td>
<td>5</td>
<td>5</td>
<td>5</td>
<td>5</td>
<td>5</td>
<td>5</td>
<td>5</td>
<td>5</td>
<td>...</td>
</tr>
<tr>
<td>s</td>
<td>0</td>
<td>2</td>
<td>6</td>
<td>4</td>
<td>4</td>
<td>7</td>
<td>4</td>
<td>6</td>
<td>...</td>
</tr>
<tr>
<td>(s_c)</td>
<td>0</td>
<td>0</td>
<td>2</td>
<td>6</td>
<td>4</td>
<td>4</td>
<td>7</td>
<td>4</td>
<td>...</td>
</tr>
<tr>
<td>p</td>
<td>0</td>
<td>2</td>
<td>8</td>
<td>12</td>
<td>16</td>
<td>23</td>
<td>27</td>
<td>33</td>
<td>...</td>
</tr>
<tr>
<td>(p_c)</td>
<td>0</td>
<td>0</td>
<td>2</td>
<td>8</td>
<td>12</td>
<td>16</td>
<td>23</td>
<td>27</td>
<td>...</td>
</tr>
<tr>
<td>x</td>
<td>F</td>
<td>F</td>
<td>T</td>
<td>F</td>
<td>F</td>
<td>T</td>
<td>F</td>
<td>T</td>
<td>...</td>
</tr>
<tr>
<td>c</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>3</td>
<td>3</td>
<td>3</td>
<td>3</td>
<td>3</td>
<td>...</td>
</tr>
<tr>
<td>(c_c)</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>...</td>
</tr>
<tr>
<td>t</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>2</td>
<td>2</td>
<td>3</td>
<td>...</td>
</tr>
<tr>
<td>pt</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>...</td>
</tr>
</tbody>
</table>

The values of s and p are ultimately defined by counter instances with a value at every instant; the variables defined within the instances by c = 0 fly n are denoted (s_c) and (p_c).

The when construct samples a stream and controls the activation of program elements. The effect of sampling the constant streams in the definition of c is to activate the counter node at a slower rate than its context. In contrast, counter(0, 1, false) when x activates counter at every instant and samples the result. The value of c is the stream 1, 2, 3, ...; its synchronization with the other streams, giving the gaps in the table, is captured by the clock base on x, which is declared in the type/clock annotation ‘: int when x’ in Figure 3. Similarly, the expression pt when not x samples pt when x is false. Its clock is base not x.

The merge construct combines complementary streams according to the clock given as the first argument. Together with the fly defining pt, its effect here is to sustain the value of t between excess speed events. A set of clock typing rules [19] allows clocks to be inferred by the compiler and ensures that well-clocked programs can be executed synchronously, that is, without additional buffering. These details are not central to our approach; it is only important to note that every equation is associated with a clock having the hierarchical structure shown in Figure 2.

The normalization pass ensures that merges and ifs only occur at the top level of expressions that do not contain flies or node instantiations. The structural invariants given by normalization are captured in the abstract syntax by the three specific forms of equations, and the distinction between control expressions and expressions.

The simple example of Figure 3 was chosen for reasons of space and clarity. But just as the concision of the λ-calculus belies its expressiveness and utility, so too is the simplicity of Lustre misleading. It and the underlying formalism of difference equations forms the core of the Scade 6 language which is used to program complex and critical discrete controllers. It also suffices to model distributed embedded controllers [5, 14, 27], and to encode automata [20, 44, 45], temporal logic [31], regular expressions [58], and numerical simulations [6, 7]. We aim to provide a solid base to manipulate and reason about such applications in an ITP.

Differences with ‘classical’ Lustre. In addition to initialized registers (flies) [64], Lustre [13] provides distinct initialization (→) and delay operators (pre). Initialization is readily expressed in SN-Lustre: x → y becomes if (true fly false) then x else y. While pre is not fundamentally more expressive than fly [13, §3.2], its use sometimes gives more pleasing programs but requires a static analysis to check that it is used correctly (initialization analysis).

The merge operator [54] supersedes the current operator, whose value before the arrival of its input is undefined. It allows syntactic clock typing and is readily generalized for the compilation of case structures and automata [20]. Scade 6 uses merge and provides fly. Simulink provides counterparts for when (‘conditionally-executed subsystems’ [46, p.4-26]) and merge (‘Merge’ [47]).

For simplicity, we require all inputs and outputs of a node application to have the same clock; in Lustre the inputs and outputs may be on sub-clocks of the clock of the first input.
2.3 Intermediate Code: Translation and Generation

In the clock-directed modular approach, compilers typically generate imperative code in an object-based intermediate language. The idea is to encapsulate the memory required to implement a dataflow node with the code generated to manipulate it. The object style is a natural way of describing and reasoning about the translation, and the use of an intermediate language facilitates the code generation for different targets by separating generic transformations from language-specific details [8, §4].

It turns out that this approach is also advantageous for formal verification because it separates the task of reasoning about the translation between the dataflow and imperative models from that of reasoning about the packing of bytes into machine memory. The invariants needed for each task differ in purpose and form. Translating directly from SN-Lustre to Clight would eliminate the need to define and verify the invariants needed for each task differ in purpose and form. Translating directly from SN-Lustre to Clight would eliminate the need to define and verify the invariants needed for each task differ in purpose and form. In the clock-directed modular approach, compilers typically generate imperative code in an object-based intermediate language. The idea is to encapsulate the memory required to implement a dataflow node with the code generated to manipulate it. The object style is a natural way of describing and reasoning about the translation, and the use of an intermediate language facilitates the code generation for different targets by separating generic transformations from language-specific details [8, §4].

It turns out that this approach is also advantageous for formal verification because it separates the task of reasoning about the translation between the dataflow and imperative models from that of reasoning about the packing of bytes into machine memory. The invariants needed for each task differ in purpose and form. Translating directly from SN-Lustre to Clight would eliminate the need to define and verify the invariants needed for each task differ in purpose and form.

2.3 Intermediate Code: Translation and Generation

In the clock-directed modular approach, compilers typically generate imperative code in an object-based intermediate language. The idea is to encapsulate the memory required to implement a dataflow node with the code generated to manipulate it. The object style is a natural way of describing and reasoning about the translation, and the use of an intermediate language facilitates the code generation for different targets by separating generic transformations from language-specific details [8, §4].

It turns out that this approach is also advantageous for formal verification because it separates the task of reasoning about the translation between the dataflow and imperative models from that of reasoning about the packing of bytes into machine memory. The invariants needed for each task differ in purpose and form. Translating directly from SN-Lustre to Clight would eliminate the need to define and verify the invariants needed for each task differ in purpose and form. In the clock-directed modular approach, compilers typically generate imperative code in an object-based intermediate language. The idea is to encapsulate the memory required to implement a dataflow node with the code generated to manipulate it. The object style is a natural way of describing and reasoning about the translation, and the use of an intermediate language facilitates the code generation for different targets by separating generic transformations from language-specific details [8, §4].

It turns out that this approach is also advantageous for formal verification because it separates the task of reasoning about the translation between the dataflow and imperative models from that of reasoning about the packing of bytes into machine memory. The invariants needed for each task differ in purpose and form. Translating directly from SN-Lustre to Clight would eliminate the need to define and verify the invariants needed for each task differ in purpose and form.

It turns out that this approach is also advantageous for formal verification because it separates the task of reasoning about the translation between the dataflow and imperative models from that of reasoning about the packing of bytes into machine memory. The invariants needed for each task differ in purpose and form. Translating directly from SN-Lustre to Clight would eliminate the need to define and verify the invariants needed for each task differ in purpose and form. In the clock-directed modular approach, compilers typically generate imperative code in an object-based intermediate language. The idea is to encapsulate the memory required to implement a dataflow node with the code generated to manipulate it. The object style is a natural way of describing and reasoning about the translation, and the use of an intermediate language facilitates the code generation for different targets by separating generic transformations from language-specific details [8, §4].

It turns out that this approach is also advantageous for formal verification because it separates the task of reasoning about the translation between the dataflow and imperative models from that of reasoning about the packing of bytes into machine memory. The invariants needed for each task differ in purpose and form. Translating directly from SN-Lustre to Clight would eliminate the need to define and verify the invariants needed for each task differ in purpose and form. In the clock-directed modular approach, compilers typically generate imperative code in an object-based intermediate language. The idea is to encapsulate the memory required to implement a dataflow node with the code generated to manipulate it. The object style is a natural way of describing and reasoning about the translation, and the use of an intermediate language facilitates the code generation for different targets by separating generic transformations from language-specific details [8, §4].

It turns out that this approach is also advantageous for formal verification because it separates the task of reasoning about the translation between the dataflow and imperative models from that of reasoning about the packing of bytes into machine memory. The invariants needed for each task differ in purpose and form. Translating directly from SN-Lustre to Clight would eliminate the need to define and verify the invariants needed for each task differ in purpose and form. In the clock-directed modular approach, compilers typically generate imperative code in an object-based intermediate language. The idea is to encapsulate the memory required to implement a dataflow node with the code generated to manipulate it. The object style is a natural way of describing and reasoning about the translation, and the use of an intermediate language facilitates the code generation for different targets by separating generic transformations from language-specific details [8, §4].

The generation of Clight from Obc involves introducing a record type for each class to store memories and instances, producing a function for each class/method pair, encoding multiple return values using records and pointers, and dealing with various technicalities of the detailed machine model exposed to C programs. We describe this pass and its correctness proof in Section 4.

We obtain the full compilation chain, from N-Lustre to assembly, by composing our compiler with CompCert and, similarly, we prove the end-to-end correctness theorem by composing the respective correctness lemmas.

3. Translation to Obc

The translation pass is specified as a functional program in Coq. It maps a list of SN-Lustre nodes into a list of Obc classes. Most of the translate function is shown in Figure 5. The tnode function is not shown for want of space. It calculates a set mems of variables defined by fby, uses it to partition declarations of node variables into declarations of class memories and the local variables of the step method, and defines the bodies of the step and reset methods using treqs and treqsr, respectively. We show that if the source program is well-typed then so is the resulting program.

Consider now the other auxiliary functions. The function var maps a dataflow variable into a local or state variable based on its membership in mems. Expressions are otherwise translated by trexp, which propagates constants and operators, and removes whens. The cases for e when not x and ⊕ e are not shown being easily deducible. Unlike trexp which maps SN-Lustre expression into an Obc one, texp maps a variable and expression into an update statement. Only the case for merge is shown, the case for if/then/else is identical.

The ctrl function embodies the principle that clocks in the source language are transformed into control structures in the target language [8], by following the structure of a clock ck to nest a statement s in conditional tests. It is applied for
each of the three cases in treqs: definitions become updates, delays become state updates, and node calls become calls to the (previously generated) step method for the associated node. The last case requires a way to identify this particular step to the (previously generated) delays become state updates, and node calls become calls of equations.

The step method is generated by folding treqs over the list of equations. Similarly, reset is generated with treqr mapping delays into state updates, and node calls into invocations of the reset method for the associated instances.

The step and reset statements for the example tracker node are shown below at left and right:

\[
\begin{align*}
\text{s, p} & \triangleq \text{d integrator, step(acc)}; \quad \text{d integrator, reset();}
\text{x} & \triangleq \text{rising, step(s > limit);} \quad \text{rising, reset();}
\text{if then c} & \triangleq \text{counter, step(0, 1, false) counter, reset();}
\text{else skip;}
\text{if then t} & \triangleq \text{c else t} \triangleq \text{state(pt);}
\text{skip state(pt)} & \triangleq t;
\end{align*}
\]

While the translation function is succinct, the formal justification of its correctness is not! We must define semantic models for SN-Lustre and Obc, and show a theorem describing how the translate function preserves the observable behavior of source programs. We also find it necessary to introduce a third semantic model to facilitate the proof. We outline the models and proof before describing the optimization of translated step methods.

### 3.1 Semantic Models

#### Dataflow.

The semantic model of SN-Lustre is a specification for the compilation chain. It must correspond with existing formal but unmechanized definitions of the language [13, §3.2][19, §3] and permit effective reasoning within an ITP. The first important choice is the representation of streams. Earlier work [2, 3] used lists, but their extrapolation to infinite objects is not simple and they engender uninteresting proof obligations—for instance, for the empty list or to argue that two streams are of equal length. Coinductive streams are a more natural choice but their use in ITPs can involve nontrivial technicalities [18, §5]. We thus model streams as functions from natural numbers to a value domain. The \(n\)th value of a stream \(s\) is denoted by \(s(n)\). This choice worked well; time was treated by induction on the index argument \((n)\) and our proofs encountered the expected obligations in the expected forms, without tedious additional technicalities.

The next choice is how to handle the sampling operators; that is, how to model the gaps in the table shown earlier. In a Kahn semantics [37, 51], these gaps are simply omitted, but in a synchronous semantics they express the timing of calculations relative to the underlying iterations of the system. We model them by explicitly encoding presence \((c)\) and absence \((\text{abs})\) in the value domain. On one hand, this complicates the definition of fby, see Figure 6, which can no longer be encoded by simply prefixing the initial value to the argument stream. Rather the argument stream must ‘slide to the right’ over the absent gaps which remain fixed in place; this is what the auxiliary hold operator encodes. On the other hand, it allows the explicit manipulation of presence and absence in invariants and proofs.

\[
\begin{align*}
(c \text{ hold}^* x)_{(0)} & = c
\text{x}_{(n)} & = (c \text{ hold}^* x)_{(n)}
(c \text{ fby}^* x)_{(n)} & = \text{abs}
(c \text{ hold}^* x)_{(n+1)} & = (c \text{ hold}^* x)_{(n)}
\end{align*}
\]

Figure 6. Definition of the fby\(^*\) operator.
The dataflow semantics is standard [13] and defined in two parts: a combinational fragment over values at an instant and a sequential fragment over streams. The combinational judgments are relative to a pair \((R, b)\) of an instantaneous environment \(R\) mapping variables to present or absent values, and a boolean \(b\) that indicates whether the enclosing node is active in the instant. There are judgments for variables \((R|_{\text{var}} x \downarrow v)\), clocks \((R|_{\text{ck}} c k \downarrow b')\), expressions \((R|_{\text{ex}} x \downarrow v)\), clocked expressions \((R|_{\text{cex}} c e \downarrow v)\), control expressions \((R|_{\text{ce}} e :: c k \downarrow v)\), and clocked control expressions \((R|_{\text{cek}} e c :: c k \downarrow v)\).

The sequential judgments are defined relative to a history environment \(H\) that maps each variable to a stream of values, and a base clock \(bk\) that is a stream of boolean values representing the base clock of the enclosing node. History environments essentially encode the kind of semantic table shown earlier, each variable name is mapped to a row giving its values over time. Since nodes can be conditionally activated, the semantics of equations is specified with respect to a base clock that indicates when the enclosing node is active.

A mutually inductive definition simultaneously defines the semantics of equations and nodes. The semantic judgment for delay equations integrates the semantics of equations and nodes. The semantic judgment for a statement \(s\) in the context of a program \(prog\) relates initial and updated memory pairs. For example, the rule for updates is:

\[
\text{memory} V \triangleq \left\{ \begin{array}{l}
\text{values} : \text{ident} \rightarrow V \\
\text{instances} : \text{ident} \rightarrow \text{memory} V
\end{array} \right.
\]

The memory of a program compiled from SN-Lustre reflects the tree of nodes in the source code: there is an entry in values for each \(\text{fby}\) and one in instances for each node call. The subscript in the step and reset calls associates an instance of a node with its sub-memory.

We define a big-step semantics for Obc. The semantic judgment for a statement \(s\) in the context of a program \(prog\) relates initial and updated memory pairs. For example, the rule for updates is:

\[
\text{memory, env} \downarrow v e \downarrow v
\]

That is, after executing \(x := e\) the global memory is unchanged and the \(x\) is updated in the local memory with the value \(v\) given by the expression semantics. The rule for state updates is the same but for updating \(\text{mem}\) values.

A method call evaluates the argument expressions in the initial memory \(\text{mem}\), looks up the class name in \(\text{prog}\), and executes its step statement relative to a smaller \(\text{prog}'\), a global (sub-)memory retrieved from \(\text{mem}\).instances, and a local memory that associates inputs to their values. The instance memory and result variables are then updated.

Executing an Obc program reads inputs from, and calculates results in, a top-level environment \(\text{env}\), and updates internal memories \(\text{mem}\) for subsequent re-executions.

### 3.2 Correctness

The correctness of the translation pass is simple to state: if a node \(f\) of a dataflow program \(G\) maps the streams of \(\bar{x}\) to the streams of outputs \(\bar{y}\), then, in translate \(G\), first executing the associated \(\text{reset}\) method before repeatedly executing the associated \(\text{step}\) method with the successive values of \(\bar{x}\) generates the successive values of \(\bar{y}\).

The formal statement of this property requires predicates expressing that the equations within each node of \(G\) are correctly scheduled and that nodes are not applied circularly. These properties are assured by elaboration and scheduling.

The resulting statement is too weak to prove directly. The semantic judgment \(G|_{\text{node}} f(\bar{x}, \bar{y})\) only describes the input/output behavior of a node. The values of internal streams, like \(c_c\), \(s_c\), and \(p_c\) in the example, are hidden and, unlike Obc memories, these streams may not necessarily have a value at each instant. These facts prevent the formulation and proof of the invariant needed to show correctness. We solve this problem by introducing a new semantic model that combines aspects from the dataflow and imperative models.
An intermediate model with exposed memory. The new semantic judgment has the form \( G \downarrow_{\text{node}} f(\overline{x}, M, \overline{y}) \). It exposes a memory tree \( M \) whose structure is isomorphic to the tree of instances in the datatransfer and translated imperative programs. Unlike in the Obc semantics, the domain \( V \) for memory is not instantiated by single constant values but rather by streams of constant values. For each variable \( x \) defined by a \( \text{fly} \) in the source program at any level of the call hierarchy, it specifies the sequence of values that should be taken over successive iterations by the corresponding \( \text{state}(x) \) in the translated code.

In defining the new model, we reuse the instantaneous semantic rules for expressions. The rules for nodes, ordinary equations, and node calls are redefined almost trivially. The new rule for \( \text{flys} \) is the only really interesting one:

\[
\begin{align*}
ms &= M.\text{values}(x) \\
ms(0) &= v_0 \\
H_0^{\text{fly}} e &::= \text{ck } \downarrow \text{ls} \\
\forall n, \quad \begin{cases} 
ms(n+1) = ms(n) \land \text{sx}(n) = \text{abs} & \text{if } ls(n) = \text{abs} \\
ms(n+1) = v \land \text{sx}(n) = \langle ms(n) \rangle & \text{if } ls(n) = \langle v \rangle 
\end{cases}
\end{align*}
\]

The stream \( ms \) corresponding to the memory for the variable \( x \) is retrieved from \( M.\text{values} \) at \( M.\text{mem} \). Initially, it takes the value \( v_0 \). At later instants, its value is maintained if the argument stream is absent; otherwise its next value is the value present on that stream. The stream \( \text{sx} \) associated with \( x \) is absent when the argument is and otherwise present with the current value of the memory. The behavior of this model is intentionally very close to that of the translated code.

If a node has a semantics, we show that it also has a semantics with exposed memories. In practice, this permits us to reason about memory values in the proofs of properties stated solely in terms of the standard dataflow semantics.

With a node memory \( M \) exposed, it becomes possible to relate it directly to a global memory \( \text{mem} \) at an instant \( n \). This relationship is expressed in the predicate \( \text{MemCorres}^f_{\text{node}}(M, \text{mem}) \) where \( G \) is implicit. The predicate requires that each equation for the node \( f \) in \( G \) satisfy \( \text{MemCorresEqn} \), Figure 7, which holds trivially for basic equations, and is otherwise defined directly for delays and recursively for node applications.

**Proof of translation.** The correctness of the \( n \)th execution of a translated step method is stated for a well-scheduled program \( G \) in terms of the semantics that exposes memories.

LEMMA 1. Given \( G \downarrow_{\text{node}} f(\overline{x}, M, \overline{y}) \) for a node \( f \) in \( G \), a global memory \( \text{mem} \) where \( \text{MemCorres}^f_{\text{node}}(M, \text{mem}) \) at an instant \( n \), then there is a global memory \( \text{mem}' \) such that

\[
\text{mem}|_{\text{step} r} = f_r.\text{step}(\overline{x}s(n)) \downarrow \text{mem}', \overline{y}s(n)
\]

and \( \text{MemCorres}^f_{\text{node}}(M, \text{mem}') \),

where we execute \( f \)’s step statement in translate \( G \) using the semantics of Obc method calls.

This lemma is shown by three nested inductions: over instants \( n \), node definitions in \( G \), and equations \( \text{eqns} \) within a node; and two case distinctions: on the three classes of equations, and whether the associated clock is true or not. Together with a lemma showing that \( f_r.\text{reset()} \) establishes \( \text{MemCorres}^f_{\text{node}}(M, \text{mem}) \), it implies the correctness result.

The formal proof is not trivial. The most subtle case is for node applications whose clock is false: since the associated imperative function is not executed we cannot appeal to the induction hypothesis on \( G \) and must reason that the instance memory does not change in this case.

Lemma 1 involves showing that an imperative step has a semantics. While Obc programs cannot diverge by construction, we must ensure that they do not ‘get stuck’. We rely on scheduling to guarantee that local variable reads succeed, on \( \text{MemCorres} \) to guarantee that state variable reads succeed, and on the existence of the dataflow semantics to show that operators are always defined. The last point is subtle but important: it entails an obligation to verify that source programs apply operators correctly.

### 3.3 Fusion Optimization

Translation generates code with too many guards, so an optimization pass is normally used to fuse adjacent conditionals. For example, the step code for tracker becomes:

\[
\begin{align*}
\text{s, p} &:= \text{d}_\text{integrator}.\text{step}(\text{acc}); \\
\text{x} &:= \text{rising}.\text{step}(s > \text{limit}); \\
\text{if x then (c := counter.\text{step}(0, 1, \text{false}); t := c) else t} &:= \text{state}(\text{pt}); \\
\text{state}(\text{pt}) &:= t
\end{align*}
\]

This optimization is effective as scheduling places similarly clocked equations together. We define it over two functions. The first simply divides sequential compositions in two:

\[
\begin{align*}
\text{fuse} (s_1; s_2) &\triangleq \text{zip } s_1 s_2 \\
\text{fuse } s &\triangleq s
\end{align*}
\]

The zip function is shown in Figure 8. It iteratively integrates statements from its second argument into the first and recursively performs the optimization.

\[\text{Note: Incorporating this optimization into the translation pass would complicate both compilation and proof. In particular, the induction along the list of equations would need to track and manipulate the ‘open’ if/else constructs.}\]

\[\text{Note: The Coq version is broken into three pieces to make termination manifest.}\]
zip (if e then s₁ else s₂) (if e then t₁ else t₂) ≜ if e then zip s₁ t₁ else zip s₂ t₂
zip (s₁; s₂) t ≜ s₁; (zip s₂ t)
zip s (t₁; t₂) ≜ zip (zip s t₁) t₂
zip s skip ≜ s
zip skip t ≜ t
zip s t ≜ s; t

Figure 8. Obc optimization: loop fusion.

While the first rule of zip does not preserve the semantics of s₁; s₂ in general,\(^5\) it does for the code produced by the translate function. To prove this, we must characterize an invariant that assures soundness, show that it holds of code produced by translate, and also that it is preserved by the successive transformations of fuse. We formalize a predicate Fusible over statements whose only non-trivial rule is

\[
\text{Fusible}(s₁) \iff \forall x \in \text{Free}(e), \neg \text{MayWrite}(x, s₁) \land \neg \text{MayWrite}(x, s₂)
\]

where MayWrite(e, s) is true iff s contains an assignment to e or state(e). The justification that this predicate holds requires a subtle technical argument about well-formed clocks in SN-Lustre programs.

We prove that fuse preserves Fusible by showing that it is a congruence for the relation \(s₁ \approx_{\text{fuse}} s₂\) that relates fusible statements that transform memory states identically.

As the treatment of fusion demonstrates, proving the correctness of optimizations may require reasoning across the whole framework, from the semantics of dataflow clocks, across the details of the translation function, and finally over equivalence relations on the imperative code.

4. Generation of Clight

The generation pass takes an Obc program, for instance, one produced by translation from SN-Lustre, and generates a Clight program for compilation by CompCert.

Generation treats the Obc classes in a program one by one. A record is declared for each with fields for every memory and instance. Another record is declared for each method with fields for the output variables. This is necessary since Obc allows multiple return values but Clight does not. We optimize the special cases of zero or one output. The records generated for the example tracker class are:

```c
struct tracker {  struct trackerstep {    int pt;    int p;    struct d_integrator s;    int t;    struct rising x;    struct counter c;  };  }
```

A separate function is generated for each class/method pair. Figure 9 shows the result for tracker/step. Two additional parameters precede the method inputs: self is a pointer to an instance memory and out is a pointer to memory for returning output values. Within the function, auxiliary variables are introduced for every invoked class/method to retrieve outputs, and for every local and output variable declared in the source method. Method calls are translated into calls to previously generated functions with two new argument values: an instance memory retrieved from self and, when necessary, an output record. A sequence of assignments is added after each call to copy from the auxiliary variable into local and output variables. These new assignments aside, the control structure of the source statement is reproduced exactly in the generated code. In expressions, variables are either translated into member accesses within out or self, or directly as registers, and constants and operators are propagated directly.

We do not detail Clight’s abstract syntax or the generation function. The former is described elsewhere [9, Figures 1, 2, and 3]. The latter is similar to the translate function of Section 3, although its definitions are more intricate since they must treat the low-level details inherent to Clight.

There are both big-step and small-step operational semantics for Clight. We reason about the correctness of generation in the big-step model since Obc methods always terminate and since their control structure is essentially preserved in the resulting code. We need not invoke CompCert’s framework for simulation proofs; we simply proceed by induction on arbitrary Obc statements.

The big-step judgment for a Clight statement is written:

\[
ge, e \vdash \text{stmt} \quad le, m, s \Downarrow le', m', oc
\]
A statement \( s \) is executed in a global environment \( ge \) that tracks global variable, function, and type declarations, and a local environment \( e \) that maps (local) variable names to their addresses and types. It may read and alter values in a temporary environment \( le \) and a memory \( m \) to produce updated versions \( le' \) and \( m' \), a trace of observable events \( t \), and an outcome \( oc \). Memories map addresses—that pair abstract block identifiers with integer offsets—to byte-level values and abstract permissions [41][1, §32]. In contrast to a local environment, a temporary environment maps variable names to values directly. The address-of operator \((\&)\) is only allowed for variables in \( e \) but reasoning about variables in \( le \) is easier because there is no indirection or aliasing.

We generate code that places the output blocks in \( e \), since we must pass their addresses in function calls, and the other variables in \( le \)—denoted by the \texttt{register} keyword in Figure 9. The semantic models can be instantiated to store function parameters in \( e \) or in \( le \); we choose the latter.

Reasoning about the generation pass presents two main technical challenges: integrating types and operators from Clight into Obc and SN-Lustre (Section 4.1), and reasoning in the memory model to relate the tree-like memories of Obc to nested \texttt{struct}s in the generated code and to handle multiple return values (Section 4.2). The solutions are applied to obtain the correctness result (Section 4.3).

4.1 Abstracting and Implementing Operators

The definitions and proofs about SN-Lustre, Obc, and the translation pass are defined relative to an \textit{operator interface}. Technically, we define the operator interface as a module type and the other components as functions over that type.

Figure 10 presents most of the interface. There are types for values (val), value types (type), constants (const), unary operators (unop), and binary operators (binop). We require that type contain an element \texttt{bool} and that \texttt{val} contain two elements \texttt{true} and \texttt{false}. There is a typing judgment \( \vdash_{wt} \) and (partial) functions that map constants and operators to their types and values.

The \texttt{bool} type and values are distinguished because they are required to define the semantics of sampling, merges,muxes, and clocks in SN-Lustre, and of conditionals in Obc.

Figure 10 omits the properties required of interface elements, namely \texttt{true} \( \neq \) \texttt{false}, \( \vdash_{wt} \texttt{true} : \texttt{bool}, \vdash_{wt} \texttt{false} : \texttt{bool}, \)

\( \vdash_{wt} \texttt{sem}_c c : \texttt{ty}_c c, \) and two type preservation properties. The type preservation property for unary operators \((\circ)\) is:

\[
(ty_{unop} \circ ty) = ty' \quad \vdash_{wt} v : ty \quad (sem_{unop} \circ v ty) = v'
\]

The property for binary operators \((\oplus)\) is similar.

Our compiler from SN-Lustre to Obc can thus be instantiated to any suitable language or for different variations of a given language. Here we instantiate it with CompCert’s values and semantics so as to compose both compilers.

\textbf{Instantiating the interface.} The operator interface must be instantiated with parts of CompCert’s front-end [9, 10] for the implementation and verification of the generation pass. We instantiate \texttt{val} with the type of CompCert values, and type with a subset of Clight types comprising integer, boolean, and floating-point types (but not pointers, arrays or \texttt{struct}s). We reuse CompCert’s dynamic semantics and type system almost directly. We do, however, impose stricter typing rules than those of C concerning Boolean values (to ensure that the only values of type \texttt{bool} are the integers 0 and 1) and implicit casts in assignments (we require explicit casts when assigning an expression to a variable of a different type). These stricter typing rules, applied at the level of the Obc and SN-Lustre type systems, simplify the proof of semantic preservation from Obc to Clight. They ensure correspondence between updates in Obc that store values directly and assignments in Clight that involve implicit casts.

4.2 Relating Memories

The correctness of the generation pass hinges on a lemma that shows the invariance of the assertion

\[
m \models \texttt{match\_states} c f \texttt{mem} e \texttt{le} (b_{self}, sofs) b_{out}
\]

that, for a given class \( c \) and method \( f \), relates the Obc memories \texttt{mem} and \texttt{e} to the Clight environments \( e \) and \( le \), and the state record at address \((b_{self}, sofs)\) and output record at address \((b_{out}, 0)\) in the Clight memory \( m \).

In essence, this predicate describes —at a logical level—the layout of the Clight memory, following the idealized memory of the Obc program. By ensuring this invariant, our compiler is guaranteed to generate memory-safe code. We explain its most salient parts in the remainder of this section.

\textbf{Separation assertions.} Expressing and reasoning about the internal components of \texttt{match\_states} involves treating the contents of \( m \) with all the concomitant details of alignment, padding, and aliasing. For example, when an Obc method updates a state element, \texttt{state(pt)} := \texttt{t}, we must show for the generated code, \((*\texttt{self}).pt = (*\texttt{out}).t\), that the address \( \texttt{self} + \texttt{offsetof}\texttt{(struct tracker, pt)} \) is properly aligned for the field type, that there is enough space to write a value of that type, and that other fields, or indeed, all other memory locations, are unchanged.
\[ \text{staterep } [] f \ \text{mem} (b, ofs) \triangleq \text{sepfalse} \]

\[ \text{staterep} (\text{class } g \{ m^t \triangleq v, \ldots \} : p) f \ \text{mem} (b, ofs) \triangleq \text{staterep } p \ \text{mem} (b, ofs) \quad \text{if } g \neq f \]

\[ \text{staterep} (\text{class } f \{ m^t \triangleq v, \ldots \} : p) f \ \text{mem} (b, ofs) \triangleq \]

\[ \text{sepall } (\lambda x. \text{ty}(b, ofs + \text{field}_o(x, \{ m^t \triangleq v \})) \ [\text{mem.values}_x] m^t \]

\[ \ast \text{sepall } (\lambda x. \text{staterep } p \ c_x \ \text{meminstances}(x) (b, ofs + \text{field}_o(x, \{ m^t \triangleq v \})) \ [\text{mem.values}_x] m^t \]

\[ \text{Figure 11.} \quad \text{The staterep function yields a memory assertion, for a class named } f, \text{ comparing the Clight memory in block } b \text{ at offset } ofs \text{ against the contents of an Obc global memory } \text{mem}. \text{ Note that } [f]_x \triangleq (\lambda v. \ f \ x \text{ is undefined } \lor \ f \ x = v). \]

Separation logic [33, 59] is designed to solve such problems and we apply it in our proof. More precisely, we exploit the basic concepts and standard connectives to express and reason about assertions on memory, but we do not use it as a variant of Hoare logic. We considered using the Verified Software Toolchain (VST) [1] in our formalization, but it focuses on proofs of particular programs whereas we prove properties about a function that produces programs. The VST also includes many sophisticated features that we do not need. Instead, we extend a small library of separation assertions that was already developed within CompCert for reasoning about stack frames.

A memory assertion is a dependent record with two main elements: mfoot, a declared memory footprint (represented as a predicate over block/offset pairs), and mpred, a predicate over memories; and two technical obligations to ensure that mpred only refers to memory elements of mfoot. Separating conjunction (\( \ast \)) is defined, for instance, as:

\[ p \ast q \triangleq \{ \text{mfoot } = \lambda b. \text{of} s \ b \ \text{of} s + \text{mfoot } q \ b \ \text{of} s \}

\[ \text{mpred } = \lambda m. \text{mpred } p \ m \ \land \ \text{mpred } q \ m

\land \text{disjoint } (\text{mfoot } p) (\text{mfoot } q) \}

The notation \( m \models p \) is just another way to write mpred \( p \ m \).

We define the standard assertions and connectives, including separating implication (\( \to \)) similarly. This differs from mechanizations that follow the standard approach [33, 59] of first constructing a separation algebra over ‘heaplets’ [1, 39]. It works well in practice for assertions that only use quantifiers in simple ways.

**State representation.** The trickiest part of match_states is an assertion that represents the layout of the nested state record: \( m \models \text{staterep } p \ f \ \text{mem} (b, ofs) \), relating a list of Obc classes \( p \), the name of a class \( f \), an Obc global memory \( \text{mem} \), and a piece of Clight memory in \( m \) starting at address \( (b, ofs) \).

The hierarchical Obc run-time state provides the blueprint for the nested Clight run-time state.

The assertion is defined as a Coq function by the three clauses shown in Figure 11. The function returns sepfalse for the empty program and skips over classes that are not named \( f \). The last clause is the interesting one. When the class named \( f \) is found, the memory is separated into two disjoint regions: one for the memory fields \( m^t \) and one for the instance fields \( v \). Each region is further divided by iterating a predicate over the fields using the sepall combinator:

\[ \text{sepall } p \ xs \triangleq \text{foldl } (\lambda ps \ x. \ p \ast x) \ xs \ \text{emp} \]

The predicate for memories uses contains \( ty (b, ofs + \text{field}_o(x, \{ m^t \triangleq v \})) \ [\text{mem.values}_x] m^t \ast \)

\[ \text{Clight’s field}_o \text{offset gives the offset of a field in the generated record. The memory range must contain the value mem.values}_x \text{if it is defined and is unconstrained otherwise. In either case, the predicate asserts that the memory exists, that it can be read and written, and that it is disjoint from that of other fields (and any other assertions at higher levels). The predicate for instances applies staterep recursively for each instance/class pair.} \]

The Clight memory model requires that the ownership of local memory, like that allocated for the output records, be explicitly surrendered when a function returns. Since the sepall assertions only retain access to the field memory and not to inter-field padding, match_states includes separating implications that allow field access to be exchanged to recover ownership of the original block of memory.

We obtain the correctness result by showing that each generated operation preserves match_states.

**4.3 Proof of Generation**

The core of the correctness proof shows invariance of match_states between an arbitrary Obc program and the Clight program generated from it. The proof proceeds by induction on the former; it is long and contains many technicalities. The semantics of operators is the same in both languages and implicit casts are justified using the type system sketched in Section 4.1. The separation assertions relate local and state variables to their implementations in Clight, we reason about them using standard techniques (associativity and commutativity of \( \ast \), and rules for loading and storing).

The instance predicates in match_states hold over the memory in block \( b \) in the range \([ofs, ofs + \text{sizeof}(ty)]) \ast \). Clight’s field offset gives the offset of a field in the generated record. The memory range must contain the value \( \text{mem.values}_x \) if it is defined and is unconstrained otherwise. In either case, the predicate asserts that the memory exists, that it can be read and written, and that it is disjoint from that of other fields (and any other assertions at higher levels). The predicate for instances applies staterep recursively for each instance/class pair.

The Clight memory model requires that the ownership of local memory, like that allocated for the output records, be explicitly surrendered when a function returns. Since the sepall assertions only retain access to the field memory and not to inter-field padding, match_states includes separating implications that allow field access to be exchanged to recover ownership of the original block of memory.

We obtain the correctness result by showing that each generated operation preserves match_states.

3 Similarly to the standard \( e \to e' \) assertion [33, 59].
This configuration is quite pessimistic but suffices for the present analysis.

work [4] with the ‘trivial’ script and default parameters. The code can be executed in a ‘test mode’ that \(\text{scanf} -O1\) and GCC 4.4.8 and estimate the Worst-Case Execution Time (WCET) of both Heptagon and our prototype found the same optimal distribution [57]. The resulting test suite comprises 14 programs, adapted examples from the literature and the Lustre v4 distribution [57]. The first column shows the worst-case estimates in cycles for step functions produced by Vélus. These estimates compare favorably with those for generation with either Heptagon or Lustre v6 and then compilation with CompCert. Both Heptagon and Lustre (automatically) re-normalize the code to have one operator per equation, which can be costly for nested conditional statements, whereas our prototype simply maintains the (manually) normalized form. This re-normalization is unsurprising: both compilers must treat a richer input language, including arrays and automata, and both expect the generated code to be post-optimized by a C compiler. Compiling the generated code with GCC but still without any inlining greatly reduces the estimated WCETs, and the Heptagon code then outperforms the Vélus code. GCC applies ‘if-conversions’ to exploit predicated ARM instructions which avoids branching and thereby improves WCET estimates. The estimated WCETs for the Lustre v6 generated code only become competitive when inlining is enabled because Lustre v6 implements operators, like \(\text{pre}\) and \(\rightarrow\), using separate functions. CompCert can perform inlining, but the default heuristic has not yet been adapted for this particular case. We note also that we use the modular compilation scheme of Lustre v6, while the code generator also provides more aggressive schemes like clock enumeration and automaton minimization [29, 56].

Finally, we tested our prototype on a large industrial application (\(\approx 6,000\) nodes, \(\approx 162,000\) equations, \(\approx 12\) MB source file without comments). The source code was already normalized since it was generated with a graphical interface.

\begin{table}[h]
\centering
\begin{tabular}{|l|c|c|c|c|c|c|c|}
\hline
Function & \textbf{Vélus} & \textbf{Hept+CC} & \textbf{Hept+gcc} & \textbf{Hept+gcci} & \textbf{Lus6+CC} & \textbf{Lus6+gcc} & \textbf{Lus6+gcci} \\
\hline
avgvelocity & 315 & 385 (22\%) & 265 (-15\%) & 70 (-77\%) & 1150 (265\%) & 625 (98\%) & 350 (11\%) \\
count & 55 & 55 (0\%) & 25 (-54\%) & 25 (-54\%) & 300 (445\%) & 160 (190\%) & 50 (-9\%) \\
tracker & 680 & 790 (16\%) & 530 (-22\%) & 500 (-26\%) & 2610 (283\%) & 1515 (122\%) & 735 (8\%) \\
pip\_ex & 4415 & 4065 (-7\%) & 2565 (-41\%) & 2040 (-53\%) & 10845 (145\%) & 6245 (41\%) & 2905 (-34\%) \\
mp\_longitudinal & 5525 & 6465 (17\%) & 3465 (-37\%) & 2835 (-48\%) & 11675 (11\%) & 6785 (22\%) & 3135 (-43\%) \\
cruise & 1760 & 1875 (6\%) & 1230 (-30\%) & 1230 (-30\%) & 5855 (323\%) & 3595 (104\%) & 1965 (11\%) \\
risingedgeretrigger & 285 & 300 (5\%) & 190 (-33\%) & 190 (-33\%) & 1440 (405\%) & 820 (187\%) & 335 (17\%) \\
chrono & 410 & 425 (3\%) & 305 (-25\%) & 305 (-25\%) & 2490 (507\%) & 1500 (265\%) & 670 (63\%) \\
watchdog3 & 610 & 575 (-5\%) & 355 (-41\%) & 310 (-49\%) & 2015 (230\%) & 1135 (86\%) & 530 (-13\%) \\
functionalchain & 11550 & 13535 (17\%) & 8545 (-26\%) & 7525 (-34\%) & 23085 (99\%) & 14280 (23\%) & 8240 (-28\%) \\
landing\_gear & 9660 & 8475 (-12\%) & 5880 (-39\%) & 5810 (-39\%) & 25470 (163\%) & 15055 (55\%) & 8025 (-16\%) \\
minus & 890 & 900 (1\%) & 580 (-34\%) & 580 (-34\%) & 2825 (217\%) & 1620 (82\%) & 800 (-10\%) \\
prodcell & 1020 & 990 (-2\%) & 620 (-39\%) & 410 (-59\%) & 3615 (254\%) & 2050 (100\%) & 1070 (4\%) \\
ums\_verif & 2590 & 2285 (-11\%) & 1380 (-46\%) & 920 (-64\%) & 11725 (352\%) & 6730 (159\%) & 3420 (32\%) \\
\hline
\end{tabular}
\caption{WCET estimates in cycles [4] for step functions compiled for an armv7-a/vfpv3-d16 target with CompCert 2.6 (CC) and GCC 4.4.8 -O1 without inlining (gcc) and with inlining (gcci). Percentages indicate the difference relative to the first column.}
\end{table}

5. Experimental Results

Our prototype compiler, Vélus, generates code for the platforms supported by CompCert (PowerPC, ARM, and x86). The code can be executed in a ‘test mode’ that \(\text{scanf}\)s inputs and \(\text{printf}\)s outputs using an alternative (unverified) entry point. The verified integration of generated code into a complete system where it would be triggered by interrupts and interact with hardware is the subject of ongoing work.

As there is no standard benchmark suite for Lustre, we adapted examples from the literature and the Lustre v4 distribution [57]. The resulting test suite comprises 14 programs, totaling about 160 nodes and 960 equations. We compared the code generated by Vélus with that produced by the Heptagon 1.03 [23] and Lustre v6 [35, 57] academic compilers. For the example with the deepest nesting of clocks (3 levels), both Heptagon and our prototype found the same optimal schedule. Otherwise, we follow the approach of [23, §6.2] and estimate the Worst-Case Execution Time (WCET) of the generated code using the open-source OTAWA v5 framework [4] with the ‘trivial’ script and default parameters.\(^\text{10}\) For the targeted domain, an over-approximation to the WCET is usually more valuable than raw performance numbers. We compiled with CompCert 2.6 and GCC 4.8.4 (-O1) for the \texttt{arm-none-eabi} target (armv7-a) with a hardware floating-point unit (vfpv3-d16).

The results of our experiments are presented in Figure 12. The first column shows the worst-case estimates in cycles for the step functions produced by Vélus. These estimates compare favorably with those for generation with either Heptagon or Lustre v6 and then compilation with CompCert. Both Heptagon and Lustre (automatically) re-normalize the code to have one operator per equation, which can be costly for nested conditional statements, whereas our prototype simply maintains the (manually) normalized form. This re-normalization is unsurprising: both compilers must treat a richer input language, including arrays and automata, and both expect the generated code to be post-optimized by a C compiler. Compiling the generated code with GCC but still without any inlining greatly reduces the estimated WCETs, and the Heptagon code then outperforms the Vélus code. GCC applies ‘if-conversions’ to exploit predicated ARM instructions which avoids branching and thereby improves WCET estimates. The estimated WCETs for the Lustre v6 generated code only become competitive when inlining is enabled because Lustre v6 implements operators, like \(\text{pre}\) and \(\rightarrow\), using separate functions. CompCert can perform inlining, but the default heuristic has not yet been adapted for this particular case. We note also that we use the modular compilation scheme of Lustre v6, while the code generator also provides more aggressive schemes like clock enumeration and automaton minimization [29, 56].

Finally, we tested our prototype on a large industrial application (\(\approx 6,000\) nodes, \(\approx 162,000\) equations, \(\approx 12\) MB source file without comments). The source code was already normalized since it was generated with a graphical interface.

\(^{10}\)This configuration is quite pessimistic but suffices for the present analysis.
but some other modifications were required. Most notably, we had to remove constant lookup tables and replace calls to special operators implemented in C/Assembler with empty nodes. Modeling and verifying these features is ongoing work. Our prototype compiles the application to assembly code in ≈1 min 40 s demonstrating that the performance of the extracted compiler is adequate for real-world code.

These preliminary experiments show the practicability of our approach in terms of compilation time and the efficiency of the generated code.

6. Related Work

Related work falls into two interrelated categories: one focuses on modeling the semantics of languages inside an ITP and the other on the correctness of compilation. We limit our survey to work that focuses on the particularities of synchronous languages, as we do, but proving the correctness of general-purpose compilers is undeniably a related problem.

Several synchronous languages have been formalized in ITPs, including a subset of Lustre in Coq using conductive types [21], an Esterel-like language in HOL with a focus on program proof [62], a shallow embedding of Lucid Synchronे in Coq with a focus on its higher-order features and clock calculus [12], and a denotational semantics of Kahn networks in Coq [51]. Similarly, work on synchronous compilers in ITBs has remained close to the dataflow model: an unpublished report [24] for a Scade 3 compiler focused on semantic and clocking definitions; the Gene-Auto project showed the correctness of equation scheduling for a Simulink to C code generator [34]. None of this work treats the translation of synchronous dataflow programs to imperative code, nor the generation of executables.

Translation validation is an alternative to compiler verification. It was first applied to synchronous languages two decades ago [52], more recently to a subset of Simulink and its optimizing RTW compiler [60], and in ongoing work on an existing Signal compiler [48, 49]. It is attractive because it decouples compilation from proof—even if the former must usually be adapted to provide hints for the latter—which can be an important practical advantage. While translation validation gives strong formal guarantees if the validator is verified [40], this is not the case for the work cited here.

One motivation for verifying a Lustre compiler is to ensure that properties verified on models also hold on generated code. An alternative is to also compile the properties and to reverify them at the code level [22]. This is an interesting approach, but it has two disadvantages: the compilation of properties and the re-verification must be trusted; verification may succeed on the model but fail on the code.

7. Concluding Remarks

We present a formalization and proof of correctness in Coq for a compiler that builds on CompCert to transform (normalized) Lustre into assembly code. The proof establishes that the dataflow semantics of source programs is correctly implemented by the generated code.

The definitions of the dataflow and imperative languages, and the translation function are adapted from previous work [2, 3, 8]. The correctness proof of the SN-Lustre to Obc translation is new. Several details of our formalization contribute to this success, but the introduction of the intermediate semantic model is central. It divides the proof into two more manageable parts and permits the statement and proof of the correctness lemma. It is used only in the proof and has no impact on compilation. The intermediate semantics encompasses both the dataflow and imperative ones: erasing memories gives the usual dataflow semantics; taking an instantaneous snapshot gives the usual imperative one.

The inclusion of the fuse optimization is all but obligatory in the clock-directed approach to compiling Lustre. Its verification within Obc demonstrates two advantages of the presented framework: properties of the source language and translation function can be exploited to justify optimizations on the imperative code; and the simple imperative language is a useful setting both for implementation and verification of optimizations. Expressing fuse, Fusible, and the proof of fuse s ≈ s in Clight is surely possible but also surely more demanding. Even as we progressively enrich our language with types, operators, and external functions from Clight, we do not expect the trade-off to change fundamentally.

In our implementation and verification of the translation of SN-Lustre to Obc and of the optimization of Obc, we abstract over the exact values, types, and operators used in programs. This separation of concerns simplifies formalization and proof, and, in principle, permits instantiations with different host languages. We integrate elements of Clight into our dataflow and intermediate languages and exploit CompCert to generate assembly code and formally relate the behavior of this code to the source dataflow model.

We describe a generation pass that transforms generic Obc programs into Clight programs. The use of a small library of separation assertions is decisive in expressing and reasoning about the invariants needed to show correctness: namely that the tree structures in the Obc semantics are correctly compiled into nested records in Clight. It quarantines the problem of reasoning about alignment and aliasing, and greatly simplifies the correctness proof of the generation pass, which is already challenging enough.

Acknowledgments

This work was partially funded by the ITEA 3 project 14014 ASSUME (Affordable Safe & Secure Mobility Evolution) and the Emergence(s) program of Paris.

We are indebted to Cédric Auger, Jean-Louis Colaço, and Grégoire Hamon for their earlier work. We thank Guillaume Baudart for his suggestions, Adrien Guattari, Soukayna M’Sirdi, and Jean Souyris for their help with the industrial application.
References


N. Halbwachs, P. Raymond, and C. Ratel. 


