Refactoring of Simulink Diagrams via Composition of Transformation Steps

Quang Minh Tran, Benjamin Wilmes
Berlin Institute of Technology
Daimler Center for Automotive IT Innovations (DCAITI)
Berlin, Germany
E-Mail: {quang.tranminh,benjamin.wilmes}@dcaiti.com

Christian Dziobek
Daimler AG
Mercedes-Benz Cars Development
Sindelfingen, Germany
E-Mail: christian.dziobek@daimler.com

Abstract—Model-based design has been increasingly adopted by the industry, especially the automotive industry, for the development of embedded software. Today, Matlab/Simulink by The MathWorks is widely employed as a modeling tool in which embedded software is modeled as data flow diagrams consisting of blocks and signals. While refactoring has become an established technique for improving the structure of code in textual programming languages, refactoring Simulink diagrams is relatively unexplored. This paper introduces a technique for specifying and implementing refactoring operations for Simulink diagrams by composing elementary and composite transformation steps. How the transformation steps can be leveraged to specify and implement complex refactoring operations is demonstrated based on the two refactoring examples Replace Goto/From With Explicit Signals and Merge Subsystems. Our prototypical implementation of a refactoring extension for Simulink is also briefly described.

Keywords—Simulink; Refactoring; Transformation

I. INTRODUCTION

The model-based design (MBD) paradigm has been widely adopted by the automotive industry to develop embedded software, with Matlab/Simulink [1] by The MathWorks being the defacto standard modeling tool. Using Simulink, software functionality is modeled as data flow diagrams by connecting functional blocks via data-carrying signals. Additional concepts of the Simulink modeling language address practical needs, like the readability of large models. For instance, model fragments can be hierarchically grouped into logical units called subsystems and related signals can be grouped into structured bus signals. The adoption of MBD using Simulink leads to models being central artifacts in development. Due to the continuously increasing software complexity and short development cycles, the creation and maintenance of models have become highly intensive and time-consuming activities.

Refactoring is an established restructuring technique which implies changing the structure of a development artifact without changing its observable behavior. Semiautomated or interactive refactoring operations have been integrated into textual programming environments like Eclipse or Visual Studio. However, at present, refactoring is practically non-existent in the Simulink Editor. The missing support for refactoring in Simulink has two potentially severe consequences. First, the model quality may be compromised if quality-improving model changes are not done due to tight development time, even if the modeler is aware of the structural deficits. Second, refactoring a huge Simulink model manually can be very labor-intensive and error-prone.

Thus, in this paper, we present a modular technique for refactoring Simulink diagrams based on the composition of predefined transformation steps. While the focus of this paper is on the underlying refactoring mechanism, we refer to a previous publication of ours for a wider spectrum of useful refactoring operations for Simulink diagrams [2].

The paper is structured as follows. In Section II, we present our meta-model for Simulink models, which serves as the basis for defining transformation steps. Our mechanism for composing transformation steps is described in Section III. How even complex refactoring operations can be specified and realized by utilizing primitive but powerful transformation steps is shown in Section IV. Insight into our prototypical implementation of the concept as an extension of the Simulink Editor is provided in Section V, followed by a summary of related work in Section VI, and our conclusion in Section VII.

II. SIMULINK META-MODEL FOR REFACTORING

The development of a refactoring technique for Simulink diagrams inevitably requires the existence of a meta-model. Unfortunately, to date, no official meta-model for Simulink diagrams has been published. Hence, we defined our own Simulink meta-model which, for the purpose of refactoring, implicitly meets the following criteria:

1) All necessary structural properties of diagrams that are required by refactorsions should be captured, including model hierarchy, signal properties and bus structure
2) Support for incomplete diagrams, such as those with unconnected signals, since some refactorings can be triggered at any time during the modeling
3) Layout information must be captured because the execution of a refactoring operation should preserve the layout as much as possible
4) Establish a degree of granularity that enables local structural changes during a refactoring operation without affecting irrelevant model parts
Figure 1 shows an excerpt of our meta-model as class diagram. In this meta-model, Block is the superclass for all other block types. A Block has a unique name in its hierarchical scope and an ordered list of Inport and Outport instances. Its position is stored in the field position. A Subsystem is a block that can contain child blocks including other subsystems. An entire model is also a Subsystem.

Regardless of whether a signal is completely connected, i.e., constituting an uninterrupted path from one source block to one or more destination blocks, it is divided into one or more segments. A segment connects a source and a target port - which can be of the following types: A real port belongs to a block and is either an import (for an incoming signal) or export (for an outgoing signal). A virtual port is either a branching point of a signal or an end point of an incompletely connected signal. Both real ports and virtual ports have an \((x, y)\) position. In contrast to a virtual port, a real port has a port number. A segment is called unconnected if its source is a virtual port but not a branching point, or its target is a virtual port.

III. TRANSFORMATION STEPS AND THEIR COMPOSITION

Instead of formulating each refactoring operation individually, we have set the goal to define basic transformation and modification steps that can be aggregated for specifying and implementing complex refactoring operations. As a result, on top of the meta-model in Section II, we have identified a collection of transformation steps (see Table I). A transformation step modifies an instance of the meta-model, i.e., a Simulink model. While defining the steps, we had to address the following key questions:

**How powerful in terms of the effect should a transformation step be?**

The use of a powerful transformation step reduces the complexity of a refactoring specification but is more difficult to reuse. For instance, there are two possible ways to define the transformation step deleteBlock that deletes a block. One way is, if a block is removed, its incoming and outgoing segments remain and become unconnected segments. A more powerful version of deleteBlock would also remove the incoming and outgoing segments. The former is especially useful if after the deletion, the references to the now unconnected segments are still needed - if, for instance, the segments are rerouted to other blocks in a following transformation step. For the sake of reusability, we have decided to keep basic transformation steps as granular as possible. If necessary, more powerful versions are defined by composing more fine-grained steps, such as deleteBlockWithSignals, which is realized by using deleteBlock and then deleting the incoming and outgoing segments using deleteSegment.

**How can transformation steps be composed to define more complex transformation steps?**

We distinguish between elementary and composite transformation steps. An elementary step modifies an instance of the meta-model without using other transformation steps, while a composite step consists of an ordered list of (possibly elementary or composite) child steps. Performing an elementary step directly changes an instance of the meta-model. A composite step can be performed by executing each step in the list in the specified order. Back to the previous example, deleteBlock is an elementary step while deleteBlockWithSignals is a composite step.

**How should a step affect the layout?**

Layouting of Simulink models ultimately addresses the positioning of blocks and signals in the Simulink Editor. Since the layout plays a crucial role for the readability of a Simulink diagram and layouting thus needs to be considered by refactoring operations, transformation steps can receive
Precondition: \textit{sameSubsystem}(srcPort, targetPort)
\[
1: \text{function ADDSEGMENT}(\text{srcPort}, \text{targetPort}, \text{[name]})
2: \ p \leftarrow \text{srcPort}.\text{containingSystem}
3: \ \text{newSeg} \leftarrow \text{new Segment(name)}
4: \ \text{newSeg}.\text{source} \leftarrow \text{srcPort}
5: \ \text{newSeg}.\text{target} \leftarrow \text{targetPort}
6: \ \text{srcPort}.\text{outSegs} \leftarrow \text{srcPort}.\text{outSegs} \cup \{\text{newSeg}\}
7: \ \text{targetPort}.\text{inSeg} \leftarrow \text{newSeg}
8: \ \text{p}.\text{childSegments} \leftarrow \text{p}.\text{childSegments} \cup \{\text{newSeg}\}
9: \end function
\]

Figure 2. Algorithm for elementary transformation step \textit{addSegment}

layout information from parent composite transformation steps. If layout information is not provided, predefined layout heuristics or Simulink itself determine the layout. For an improved layout after refactoring, the automatic layouting algorithm for Simulink diagrams [3] can be used.

Based on these basic principles, we show by examples how elementary and composite steps modify an instance of the meta-model. Figure 2 depicts (informal) pseudo code for the elementary step \textit{addSegment} that adds a new segment from a source port \textit{srcPort} to a target port \textit{targetPort}. It presumes that \textit{srcPort} and \textit{targetPort} are in the same subsystem and \textit{targetPort} does not have an incoming segment. If these conditions are satisfied, a new segment \textit{newSeg} is created. A name is given depending on the type of refactoring in which this step is used. The segment’s start and end ports are set to \textit{srcPort} and \textit{targetPort}, respectively. Additionally, the new segment is added to the collection of outgoing segments of \textit{srcPort} and assigned to \textit{targetPort} as the incoming segment. Finally, \textit{newSeg} is added to the current subsystem.

Figure 3 shows the algorithm for \textit{addCrossHierarchicalSignal} which adds (possibly cross-hierarchical) signals from a source port to one or several target ports. Unlike \textit{addSegment}, \textit{addCrossHierarchicalSignal} is a composite transformation step because it makes use of other transformation steps such as \textit{addSegment}, \textit{addOutportBlock} and \textit{addInportBlock}. Note that the algorithm contains control structures, as well as other commands, and is not purely a list of transformation steps as indicated before. Due to space limitations, we abstract from implementation details here. As mentioned, the steps’ algorithms as described in this paper are executed on an instance of the meta-model. Each call of an elementary transformation step is registered in an ordered step list which is then executed step by step on the real Simulink model.

The precondition of \textit{addCrossHierarchicalSignal} states that the list of target ports must have at least one element and all target ports must be in the same subsystem. If a precondition is not satisfied, the entire refactoring in which this step is used will not be applied. If satisfied, it determines the subsystem where a forward constructed signal from source to target and a backwards constructed signal from target to source would meet (least common subsystem). Then, the signal is forwarded from the source

Precondition: \textit{sameSubsystem}(targetPorts) $\land$ targetPorts $\neq \emptyset$
\[
1: \text{function ADDCROSSHIERARCHICAL SIGNAL} (\text{srcPort}, \text{targetPorts})
2: \text{leastSub} \leftarrow \text{getLeastCommonSubsystem}(\text{srcPort}, \text{targetPorts})
3: \text{curPort} \leftarrow \text{srcPort}
4: \text{curSubSys} \leftarrow \text{curPort}.\text{containingSystem}
5: \text{while curSubSys} \neq \text{leastSub} \text{do}
6: \text{outBlock} = \text{addOutportBlock}(\text{curSubSys})
7: \text{ret} = \text{addSegment}(\text{curPort}.\text{outBlock}.\text{inport})
8: \text{curPort} \leftarrow \text{curSubSys}.\text{outportOf}(\text{outBlock})
9: \text{curSubSys} \leftarrow \text{curPort}.\text{containingSystem}
10: \text{end while}
11: \text{subsysPath} \leftarrow \text{getSubsystemPath}(\text{targetPorts}(1), \text{leastSub})
12: \text{for p} \in \text{sortByHierarchyTopDown}(\text{subsysPath}) \text{do}
13: \text{inBlock} = \text{addInportBlock}(p)
14: \text{addSegment}(\text{curPort}.p.\text{importOf}(\text{inBlock}))
15: \text{curPort} \leftarrow \text{inBlock}.\text{outport}
16: \text{end for}
17: \text{for tp} \in \text{targetPorts} \text{do}
18: \text{addSegment}(\text{curPort}.\text{tp})
19: \text{end for}
20: \end function
\]

Figure 3. Algorithm for composite transformation step \textit{addCrossHierarchicalSignal}

The transformation steps can be leveraged to formulate refactoring operations, as shown next using two examples:

(1) \textit{Replace Goto/From With Explicit Signals} creates explicit (possibly cross-hierarchical) signals from the source Goto block to all associated From blocks, (2) \textit{Merge Subsystems} merges two subsystems into a single subsystem. These two refactorings are part of our Simulink refactoring catalog [2].

A. Replace Goto/From With Explicit Signals

\textbf{Motivation:} An advantage of data flow diagrams such as Simulink is that the data flow between blocks is explicit thanks to visual signal connections. However, Simulink provides Goto/From blocks as a means to define implicit, non-visual signal connections between blocks that may reside on different model levels - usually to reduce the visual complexity. Similar to the Goto construct in imperative programming languages, the use of Goto/From blocks, especially of global scope, may dramatically reduce the understandability of the model because tracing the data flow becomes more difficult. Goto/From blocks can be replaced by explicit signal connections without changing the behavior of the model. This can be a tedious task when done manually.
Mechanics: Figure 6 shows (informal) pseudo code for Replace Goto/From with Explicit Signal. Take a look at Figure 4 for an example.

As first operation, the function buildSubsystemList is called to obtain a list of the corresponding From blocks' parent subsystems. This list is used for signal forwarding while avoiding redundant signal paths at the same time. Then, the transformation step replaceBlock is used to replace the Goto block with a Signal Conversion block. Signal Conversion blocks are used here solely for preserving signal names. More specifically, if the incoming signal of the Goto block has a different name than the signal names leaving the From blocks, the use of Signal Conversion blocks would allow these names to continue to exist after the refactoring.

Next in the algorithm, subsystemList is iterated. In each iteration, the From blocks within the current subsystem are replaced by Signal Conversion blocks. Finally, the composite transformation step addCrossHierarchicalSignal is used to forward signals to the Signal Conversion blocks.

B. Merge Subsystems

Motivation: During creation and maintenance of a Simulink model, reorganizing activities are frequent. In particular, it is often necessary to combine functionalities residing in separate subsystems into a single subsystem. With the current modeling support of the Simulink Editor, the modeler would have to cut and paste the content of one subsystem into the other subsystem. Then, the signals must be reconnected to re-establish the initial signal relationships. If lots of signals must be connected manually, this activity becomes both labor-intensive and error-prone.

Mechanics: While Figure 5 provides an example of this refactoring, Figure 7 shows (informal) pseudo code for merging two subsystems A and B. The precondition specifies that A and B must be non-atomic (virtual) subsystems. This restriction exists since merging atomic subsystems may change the behavior of the model.

In essence, the algorithm uses suitable transformation steps to move the content of B to A (line 18), adjust the signal connections (line 23, 26, 32, and 35) and finally delete B (line 37). Before the actual transformation, some book keeping needs to be done. In particular, inSegsOfB contains all incoming segments of B. Hash tables inMap and outMap are used to keep track of the references between import and output blocks of B to the source ports of the signals reaching them for reconnecting signals.

For inMap, if an import block inpBlock of B has an incoming segment entering B at the import corresponding to inpBlock, we store the mapping between inpBlock and the source port of that incoming segment returned by getSrcPort. In this context, the source port is returned by getSrcPort, which checks if the root source of the segment is an output of A. If yes, we go into A and retrieve the source of the signal within A. Otherwise, the source port is the root source of the segment and is located on the common parent subsystem of A and B. For outMap, if an output block outBlock of B has an outgoing segment leaving B at the output corresponding to outBlock, we store the mapping between outBlock and the segment.

For transformation, moveBlocks(B.content,A) moves B’s content to A. inMap is used to reestablish incoming signal connections to the blocks that used to be in B. In particular, for each import block inpBlock stored in inMap, if the source port of inpBlock in A, inpBlock is replaced by a Signal Conversion block before connecting the source port to the import of that Signal Conversion block. Otherwise, the source port is connected to the import of A corresponding to inpBlock. Then, using outMap, the initial outgoing segments of B are rerouted to the outputs of A corresponding with the output blocks being moved from B. Finally, the initial incoming segments of B and B itself are removed.

It should be noted that the decision of which of the two subsystems to be merged serve the role of A and B in the...
Precondition: \neg \text{A.isAtomic} \land \neg \text{B.isAtomic}
1: function \text{MERGE}()\text{SUBSYSTEMS(A, B)}
2: inpSegsOfB \leftarrow B.inSegments
3: for inpBlock \in inpBlocksInB do
4:  \text{inp} \leftarrow B.inportOf(inpBlock)
5:  \text{inpMap} \leftarrow getSrcPort(inpBlock, A)
6:  \text{inpMap.endB} \leftarrow getSrcPort(inpBlock, B)
7:  \text{inpMap.value(inpBlock)} \leftarrow getSrcPort(inpBlock, A)
8: end if
9: end for
10: outBlocksInB \leftarrow B.outportBlocks
11: for outBlock \in outBlocksInB do
12:  outSeg \leftarrow outport.outSegment
13:  outSeg \leftarrow outport.outSegment
14:  outMap.value(outBlock) \leftarrow outSeg
15: end if
16: end function
17: moveBlocks(B.content, A)
18: for inpBlock \in inpMap.keys do
19:  inpBlock \leftarrow inpMap.keys
20:  srcPort \leftarrow inpMap.inMap
21:  if srcPort.containingSubsystem = A then
22:     converter \leftarrow \text{replaceBlock(inpBlock,'Signal Conversion')}
23:     addSegment(srcPort, converter.inport)
24:     addSegment(srcPort, converter.targetPort)
25:     end if
26: end for
27: end function
28: for outBlock \in outMap.keys do
29:  outBlock \leftarrow outMap.keys
30:  outPort \leftarrow A.outportOf(outBlock)
31:  outSeg \leftarrow outPort.outSegment
32:  outPort \leftarrow outPort.outSegment
33: end for
34: for seg \in inSegsOfB do
35:  deleteSegment(seg)
36: end for
37: deleteBlocks(B)
38: end function

Figure 5. Example for refactoring Merge Subsystems

Figure 7. Algorithm for refactoring Merge Subsystems

algorithm affects the port order of the imports and exports within the resulting merged subsystem. This is due to the way Simulink automatically assigns port numbers when a port is added or deleted. However, since the port order does not affect the behavior, the refactoring does not change the model behavior. We have also defined a refactoring operation called Reorder Ports that can be used to rearrange the port order of imports or exports of subsystems. If required, this refactoring can be used to achieve the desired port order.

V. IMPLEMENTATION

We have implemented a prototype in Matlab’s m language that integrates refactoring support directly into Simulink Editor. Specifically, a refactoring operation can be directly triggered in Simulink Editor via a menu item or shortcut. Based on the Template design pattern [4], the prototype implements a generic workflow of refactorings as a graphical wizard. The behavior of a specific refactoring operation such as the required graphical dialogs for user input and the (interactive) specification of transformation steps can be easily defined and integrated into the prototype.

The meta-model in Section II is implemented as Matlab classes. The elementary and composite transformation steps in Section III are provided in the form of Matlab functions. In addition, the prototype also contains a collection of Matlab functions for model analysis that are useful for refactoring purposes. For instance, the functions getLeastCommonSubsystem used from Figure 3 and buildSubsystemList from Figure 6 are stored in a special collection since they are needed by multiple refactorings. The functions for transformation steps and model analysis serve as a high-level and compact API for formulating refactoring operations.

The prototype also features a graphical preview that shows the list of transformation steps to be executed in a tree. Moreover, it shows the Simulink diagram before and after a refactoring operation.

We have tested our prototype on several industrial Simulink models from the automotive domain at Daimler. The biggest time factor turned out to be the time required to convert a Simulink model into an instance of the meta-model. In an extreme case, for a model of about 20,000 blocks and a refactoring operation that affects almost the entire model, the parsing time took roughly 10 minutes. For most models and operations, however, the parsing time was just a matter of seconds. The transformation itself usually took only seconds, or at most, a few minutes.

VI. RELATED WORK

In textual programming, refactoring has become a standard technique for restructuring code without changing its observable behavior [5], such as for object-oriented languages [5] and functional languages [6]. Modern Integrated Development Environments (IDEs) like Eclipse, NetBeans and Visual Studio offer built-in support for refactoring.

In model-based development, UML models have been targeted for refactoring support [7]. Refactoring of data flow
diagrams such as Simulink, however, is only scantily researched. Sui et al. [8] propose an implementation approach for an automated refactoring tool aimed at visual dataflow programming languages. However, the focus of their paper is rather on the tool architecture aspect than on specifying refactoring operations modularly.

The current version of Simulink Editor does not provide refactoring support. Tools such as Model Advisor [9] or Model Examiner [10] can automatically detect violations of modeling guidelines and do provide, to a limited extent, so-called repair scripts for repairing guideline violations. Nonetheless, the focus of these tools is on automated detection of guideline violations and not on providing complex refactoring operations with possible user input or interaction.

The approach which is most related to our work has been developed in the MATE project [11]. It is an approach to visual specification and transformation for Simulink and Stateflow models based on graph transformation techniques. Specifically, modeling guideline violations and possible repair scripts are formulated in the graphical specification language called Story Driven Modeling (SDM). It turned out, however, that a purely visual specification language, such as SDM, is not powerful enough for complex real specification scenarios such as those including regular expressions, complex mathematical calculations and complex navigation through a network of linked objects.

VII. Conclusion and Future Work

In this paper, we have introduced our technique for specifying and implementing complex refactoring operations for Simulink diagrams based on the composition of transformation steps. The concept has been successfully implemented as a prototype that integrates refactoring support into Simulink Editor. Using the infrastructure provided by the prototype, we were able to implement many refactoring operations from our catalog [2] with little effort.

As the next step, we plan to extend our catalog and tool with further useful refactoring operations. Our future work will also address the automated identification of model constructs for which the application of certain refactorings is recommendable - so-called model smells, in analogy to code smells known from code refactoring [5]. There exist several techniques for Clone Detection in a Simulink diagram, as explored by Deissenboeck et al. [12] and Petersen [13], which could be used to identify similar or identical fragments in a Simulink diagram and suggest applicable refactoring operations for eliminating them. Moreover, we plan to evaluate the developed techniques and tool in real development environments at Daimler.

In addition, having automated transformation and refactoring techniques for Simulink models on hand, advanced applications are rendered possible. For instance, Simulink models could be automatically optimized by search-based algorithms using our transformation steps, as suggested for code [14], with respect to measurable model quality criteria, which already exist for Simulink diagrams [15].

REFERENCES


[4] E. Gamma, R. Helm, R. Johnson, and J. Vlissides, Design Patterns: Elements of Reusable Object-Oriented Software. Addison-Wesley, 1995.


