Abstract

Model-based design can shorten the development time of complex systems by the use of simulation techniques. However, it can be hard to simulate the system as a whole if it is developed in a concurrent fashion, by multiple and specialized teams. Co-simulation, with the support of the Functional Mock-up Interface (FMI) Standard, is proposed as a way to promote tool interoperability while protecting intellectual property of subsystems. The standard allows a uniform communication between subsystem simulators, but does not state how the inputs and outputs should be interpreted, nor how the subsystems should interact correctly. Semantic adaptations can be quickly made to correct the interactions with sub-system simulators that were produced with different assumptions, and avoid changing those sub-systems, their simulators, or the orchestration algorithm that computes the co-simulation. In this work, we explore how to describe common adaptations and what their meaning is in the context of FMI co-simulation. The result is a sound language that enables the implementation of adaptations with minimal effort. A distinct feature is that it describes adaptations for groups of interconnected subsystem simulators in the same way as for a single simulator, and the implementation is itself a simulator, thanks to a sound definition of hierarchical co-simulation. This work paves the way for research into the correct combination and interfacing of different adaptations.

Keywords
Co-simulation, Semantic adaptation, Functional Mockup Interface Standard, Hierarchical co-simulation

Introduction

The systems we engineer today are characterised by an increasing complexity. Model-based design can boost the development of such systems by enabling their analysis at higher levels of abstraction via simulation. However, it can be hard to simulate the system as a whole if it is developed in a distributed fashion, by multiple and specialized teams.

Two factors contribute to this difficulty: (i) specialized teams have their own tools; and (ii) some of the components of the system are provided by different suppliers, and have valuable Intellectual Property (IP).

Difficulty (i) is a natural consequence of using the most appropriate formalism for a specific domain. The same can be expected from external suppliers. During the development process, if a team wishes to understand how a component being developed behaves when interacting with the rest of the system, it is useful that the tool in use can import and simulate correctly the models created by the other teams (with different tools). As we show shortly, it can be hard to correctly simulate imported models, as these belong to potentially different domains, each with its own set of specialized simulators.

As for difficulty (ii), if the team is using externally supplied components and wishes to simulate them, it may not be able to import the components’ models because these contain protected IP. For the sufficiently complex components, a “lock-in” contract can be made to allow the team access to those models. However, the team will no longer be free to benchmark components from different competing suppliers.

Co-simulation, with the support of the Functional Mock-up Interface (FMI) Standard, is proposed as a way to promote tool interoperability while addressing the IP protection requirement. The models are exported as executable black boxes, that receive inputs and produce outputs, allowing for the
simulation of the component they stand for. In the FMI, each black box is called a Functional Mock-up Unit (FMU), the term adopted in this document.

The standard provides a common interface to allow a uniform communication with the black boxes, solving the combinatorial explosion of import/export formats. However, it does not ensure that the black boxes are interacted with in a semantically correct manner.

When a team is given an FMU that does not behave as it is expected to, we say that there is an interaction mismatch between the FMU and the rest of the system. Interaction mismatches can be roughly classified as:

- **Signal Data Mismatch** happens when the signals provided by the FMU are not compatible with the ones that are expected (e.g., different frame of reference or different physical units).

- **Model of Computation Mismatch** happens when the provided FMU assumes a different model of computation than the one actually used to compute the overall behavior of the system (e.g., FMUs exported by a timed automata modelling and simulation tool have to make assumptions about the other interacting FMUs).

- **Capability Mismatch** happens when a given FMU lacks some capabilities that affect the simulation performance (e.g., FMUs that lack higher order input extrapolation, an important capability that affects the accuracy and stability of the co-simulation).

Rather than asking the original producer of the FMU to correct an interaction mismatch, it can be useful that the team is able to correct it immediately. Note that any mismatch happens between a given FMU and a usage intent, and therefore it is not necessarily the case that the best correction of a mismatch is done by the producer (if the FMU is to be reused).

In fact, as show, some mismatches happen as a product of the (incorrect) handling of multiple interacting FMU’s, and the correction has to be done for that specific interaction.

The above arguments motivate the need for semantic adaptations, and lead to the following research question:

**RQ1** How can we describe the most common semantic adaptations on multiple types of black box FMUs in a productive manner, and realise them without violating modularity and transparency.

Informally, we call semantic adaptation of an FMU to the set of modifications made to the inputs/outputs and interaction with environment, of the FMU, with the purpose of correcting an interaction mismatch. This concept is formalized later in this work.

**Productivity** is related to the effort required to describe an adaptation. **Modularity** refers to the fact that any FMU should be adapted by changing how it is interacted with, and not how it is implemented. **Transparency** means that any tool that imports FMUs should not have to be changed in order to import, and interact with, an adapted FMU.

The descriptions should be made in an independently developed language because it is impractical that every tool capable of importing FMUs is able to implement the adaptations. Furthermore, one cannot expect that any user of an FMU has the ability to modify the importing tool to support these. Compared to implementing these adaptations manually, a language reduces the accidental complexity, prevents mistakes, and allows soundness analyses to be carried out.

In this paper, we build on prior work to define a language that allows for the descriptions of the most common semantic adaptations that can be used in FMI co-simulation, surveyed in. A distinct feature of the language proposed here is that it describes adaptations for groups of interconnected FMUs in the same way as for a single FMU, thanks to a sound definition of hierarchical co-simulation.

The definition of hierarchical co-simulation, and the semantics of the language, are presented in a bottom up approach, as illustrated in Figure 1. In the Background section, we introduce a co-simulation abstraction with simulation units and how these relate to FMUs. Section “Hierarchical Co-simulation for Semantic Adaptation” contains the formal foundations of a special kind of simulation unit that is the template to implement any semantic adaptation. In Section “Running Example”, a running example is described, and in Section “A DSL for Semantic Adaptation” the language and its semantics are described. Section “Evaluation” judges how well we have addressed the research question. Section “Discussion and Future Work” discusses the flaws of our approach and research opportunities. Finally, Sections “Related Work” and “Conclusion” present the related work and conclude, respectively.

**Background**

In this section, we introduce the concepts, terminology, and assumptions used throughout this document. We cover co-simulation, the Functional Mockup Interface (FMI) standard, semantic adaptation, and domain specific languages.

**Co-Simulation**

We briefly summarize the main concepts related to co-simulation and we refer the reader to for a more detailed introduction of each concept.

We call dynamical system to a model that has a notion of state and rules describing the evolution of

---

*See for an overview of capabilities of FMUs.
that state across time, starting from an initial state. Inputs and outputs can be defined, to describe the environment.

A simulator is an algorithm that takes a dynamical system and input signals, as input, and computes an approximated behavior trace of the dynamical system.

A simulation unit (also known as a simulation application) is the composition of a simulator together with a dynamical system, essentially representing a mockup of a real system. It accepts input trajectories and produces a behavior trace.

Simulation units (or just units) can be coupled through their inputs and outputs. A coupling restriction (or just couplings) is an output connected to an input. It means that the trajectory computed at that output — e.g., a function of the continuous time — must be equal to the one computed at the input, at all times.

The orchestrator (or master) is an algorithm that takes a set of simulation units and their coupling restrictions — that is, a co-simulation scenario — and computes the behavior trace of all units, trying to satisfy the coupling restrictions. In practice, these restrictions can only be satisfied at certain countable points in time, called communication points. These points are agreed upon by the simulation units and the orchestrator.

A basic orchestrator will, at each communication point, copy data points from outputs to inputs, and ask each unit to compute its own behavior trace until the next communication point. The collective behavior trace is called the co-simulation.

We capture the essence of a simulation unit with reference $i$, using the discrete time system notation, in one of the following four possible ways:

$$
\begin{align*}
\langle x_i(t + \hat{H}_i), \hat{H}_i \rangle &= F_i(t, H, x_i(t), u_i(t + H) \text{ or } u_i(t)) \\
y_i(t) &= G_i(t, x_i(t), u_i(t)) \text{ or } G_i(t, x_i(t)) \\
x_i(0) &= \text{Init}_i(u_i(0)) \text{ or } \text{Init}_i(\text{...})
\end{align*}
$$

where $t$ denotes the simulated time, $x_i$ denotes the state vector, $u_i$ the input vector, $\text{Init}_i$ computes the initial state, $H > 0$ denotes the requested communication step size, $0 < \hat{H}_i \leq H$ denotes the communication step size taken by the unit, $F_i$ is the state transition function, and $G_i$ the output function. Bold symbols will always refer to vectors in this paper.

The definition in Equation (1) covers the different kinds of simulation units considered (based on the orchestrators surveyed in [46]): Reactive Mealy, Reactive Moore, Delayed Mealy, and Delayed Moore. The difference is in where and when the unit expects inputs to be provided. For example, a delayed Moore unit can compute its output without requiring an input, and can compute its future state $(x_i(t + \hat{H}_i))$ with just the current input $u_i(t)$. A reactive Mealy unit, on the other hand: requires an initial input to compute the initial state; and needs to know the next input in order to compute the next state/output.

We use shortcuts such as $F_i(t, H, x_i(t), \ldots)$, $G_i(t, \ldots)$, and $\text{Init}_i(\ldots)$, to denote the appropriate function depending on the kind of unit $i$. Furthermore, we make the following remarks about each simulation unit $i$:

- $F_i$ and $G_i$ are mathematical functions (also denoted pure).
- The internal definition of $F_i$ and $G_i$ is unknown, but the kind of unit is known.
- If $\langle \cdot, \hat{H}_i \rangle = F_i(t, H, x_i(t), \ldots)$ and $\hat{H}_i < H$, then the unit rejects the step size $H$ requested. Furthermore, for any $\hat{H} \leq \hat{H}_i$, we assume that $\langle \cdot, \hat{H} \rangle = F_i(t, \hat{H}, x_i(t), \ldots)$.

Given a set of unique unit references $D = \{1, \ldots, n\}$, a co-simulation scenario is defined as the aggregation of each simulation unit definition, in Equation (1), plus a coupling function that defines the input of $i$ as a function of the outputs of units $\{j : j \in D \setminus \{i\}\}$. Formally, combining the notation used in [6,53], a scenario is given by:

$$
\begin{align*}
\langle x_i(t + \hat{H}_i), \hat{H}_i \rangle &= F_i(t, H, \ldots) \\
y_i(t) &= G_i(t, \ldots) \\
\underline{u}_i &= c_i(y_1, \ldots, y_{i-1}, y_{i+1}, \ldots, y_n) \\
x_i(0) &= \text{Init}_i(\ldots)
\end{align*}
$$

for each $i \in D$

where $c_i$ denotes the coupling function, and each $F_i, G_i$ follows one of the definitions in Equation (1). Commonly, $c_i$ is linear and maps at most one component of one of the inputs (the inputs/outputs are vector quantities), onto one component of the output. We assume that $c_i$ is linear.

Let $i, j \in D$ be two different units, and $\bar{0}$ be the zero matrix of appropriate dimension. If $\frac{\partial c_i}{\partial y_j} \neq 0$, then $i$ gets part of its input from $j$. Informally, this means that at least one component of $u_i = c_i(\ldots)$ is determined by at least one component of $y_i$. We say

Figure 1. Overview of DSL semantics and document structure.
that a unit $i \in D$ depends algebraically on unit $j \in D$, with $i \neq j$, if $i$ gets part of its input from $j$ and $i$ is not a delayed Moore. So, e.g., if $i$ gets part of its input from $j$, but it is a delayed Moore, then $i$ does not depend algebraically on $j$. Using the algebraic dependency relationship, one can build a directed graph — called the dataflow graph — with one node $n_i$ per simulation unit $i \in D$, and an edge $(n_j, n_i)$ between two nodes $n_j, n_i$ whenever the unit $i$ depends algebraically on unit $j$. This procedure is based on the Causal Block Diagram Simulation algorithm\textsuperscript{43,66}. A topological order of the resulting graph gives an execution order that respects the units’ algebraic dependencies.

Depending on the coupling function and on the kind of simulation units being coupled, algebraic loops may occur. An algebraic loop includes any input/output/state that depends on itself, at the same time point\textsuperscript{35}. If an algebraic loop exists between the units, then it is not possible to compute a topological ordering of the dataflow graph. For now, we assume that such topological order can always be computed for a given co-simulation scenario. We denote that order via a mapping $\sigma : N \rightarrow D$, that returns the unit reference $\sigma(j)$ that is the $j$-th in the topological order. So $\sigma(1)$ gives a unit that is first in the topological order, i.e., has no algebraic dependencies.

With a well defined topological order, the orchestrator only has to provide inputs to, execute, and get outputs from, the units in that order. Algorithm 1 formalizes what is known in the state of the art as the Gauss-Seidel orchestrator. It computes the behavior trace of a given co-simulation scenario as described in Equation (2). To be concise, we abbreviate the output and state transition function calls, which depend on the kind of unit (lines 11, 19, and 21). Furthermore, the orchestrator provides the inputs ($uc_{\sigma(j)}$ or $up_{\sigma(j)}$ in line 19) that each unit expects, working for both reactive and delayed units alike. This is the main reason we single out this orchestrator in this work.

Without loss of generality, we assume the most basic step size control policy in Algorithm 1: the communication step size is never increased after being rejected by some unit\textsuperscript{*}. The orchestrator uses the most recent consistent state.

**Functional Mock-up Interface Standard (FMI)**

The FMI standard\textsuperscript{11} defines the interface and interaction pattern that allows simulation units to communicate. In the standard, a simulation unit is called a Functional Mockup Unit (FMU).

**FMUs and Simulation units.** This subsection establishes the equivalence between FMUs and simulation units (recall Figure 1), and the assumptions we make throughout this document.

---

### Algorithm 1: Gauss-seidel orchestrator for co-simulation scenarios can be topologically sorted.

**Data:** The stop time $T$, a starting communication step size $\bar{H}$, and a set of unit references $D = \{1, \ldots, n\}$.

1. $t := 0$; // Simulation time
2. $H := \bar{H}$; // Communication step size
3. Initialize variables
4. for $i = 1, \ldots, n$ do
5. \hspace{1em} $x_i := 0$; // State vector
6. \hspace{1em} $uc_i := y_i := 0$; // Current I/O variables
7. \hspace{1em} $up_i := 0$; // Previous input variables
8. end
9. // Compute initial states
10. for $j = 1, \ldots, n$ do
11. \hspace{1em} $uc_{\sigma(j)} := \sigma_j(y_1, \ldots, y_{\sigma(j)} - 1, y_{\sigma(j)} + 1, \ldots, y_n)$;
12. \hspace{1em} $x_{\sigma(j)} := Init_{\sigma(j)}(uc_{\sigma(j)})$ or $Init_{\sigma(j)}()$;
13. \hspace{1em} $y_{\sigma(j)} := G_{\sigma(j)}(t, x_{\sigma(j)}, uc_{\sigma(j)})$ or $G_{\sigma(j)}(t, x, uc_{\sigma(j)})$;
14. \hspace{1em} $up_{\sigma(j)} := uc_{\sigma(j)}$;
15. end
16. while $t < T$ do
17. \hspace{1em} accepted := false;
18. while not accepted do
19. \hspace{2em} for $j = 1, \ldots, n$ do
20. \hspace{3em} $uc_{\sigma(j)} := c_{\sigma(j)}(y_1, \ldots, y_{\sigma(j)} - 1, y_{\sigma(j)} + 1, \ldots, y_n)$;
21. \hspace{3em} $\sigma_j(x_{\sigma(j)}, \bar{H}_{\sigma(j)}) := F_{\sigma(j)}(t, H, x_{\sigma(j)}, uc_{\sigma(j)}, or up_{\sigma(j)});
22. \hspace{3em} $y_{\sigma(j)} := G_{\sigma(j)}(t + \bar{H}_{\sigma(j)}, x_{\sigma(j)}, uc_{\sigma(j)})$ or $G_{\sigma(j)}(t + \bar{H}_{\sigma(j)}, x, uc_{\sigma(j)})$;
23. \hspace{3em} end
24. \hspace{1em} $\bar{H} := \min_{i \in D} (\bar{H}_i)$;
25. \hspace{1em} if $\bar{H} < H$ then
26. \hspace{2em} $H := \bar{H}$;
27. \hspace{2em} else
28. \hspace{3em} accepted := true;
29. \hspace{2em} end
30. end
31. // Commit state and update I/O
32. for $j = 1, \ldots, n$ do
33. \hspace{1em} $x_i := \hat{x}_i$;
34. \hspace{1em} $up_i := uc_i$;
35. end
36. $t := t + H$; // Advance time

Given a simulation unit $i$ (described in Equation (1)) we define its equivalent FMU, and vice versa, as follows:

**FMU State** — The state of the FMU corresponds to the state of the unit $x_i$. The FMU does not make the state explicit, but instead implements functions that can be used to set and retrieve the state.

**Inputs** — FMUs have input ports, each accepting a scalar quantity. Each dimension in the input $u_i$ corresponds to one input port of the FMU.

*Algorithm 1 can be greatly optimized (e.g., rolling back as soon as a reject occurs).*
The FMU implements functions that allow the orchestrator to set those inputs (e.g., fmi2SetReal and fmi2SetInteger) and a single vector quantity $u_i$ can be set via multiple calls to those functions.

**Outputs** – The outputs of the FMU are analogous to the inputs. To obtain an output $y_i$, multiple calls are made to the dedicated functions (e.g., fmi2GetReal and fmi2GetInteger).

**Initial State** – The initial state computed by the $\text{Init}$ function corresponds to the computation performed by the FMU in the initialization mode. We assume that an initial state of a unit/FMU can always be found from the $\text{Init}(\ldots)$ function (and initial input, in case of a reactive unit). This is in accordance with the FMI Standard, but it restricts our scope to scenarios where the consistent initial state of one unit depends on factors (e.g., the initial state of another unit) other than its initial inputs.

**Co-simulation Step** – A state transition invocation $\left(\vec{x}_i, \vec{H}_i\right) := F_i(t, H, x_i, u_i)$ is mapped to (in order): an optional invocation to set the state of the FMU to $x_i$; multiple invocations to set the input $u_i$; an invocation to the fmi2DoStep function; a query to find out up to which time the FMU computed the step (to get $\vec{H}_i$); and an (optional) invocation to get the new state of the FMU $\vec{x}_i$. The manipulation of the state is optional for orchestration algorithms that do not perform rollback operations. However, in this document, we assume that the FMUs support rollback.

**Output Function** – If the unit is a Mealy unit, then the execution of the output function $y_i := G_i(t, x_i(t), u_i(t))$ corresponds to setting the inputs to the FMU, and then getting the outputs. If the unit is a Moore unit, then the outputs can be enquired without first setting the inputs.

It is the role of the orchestrator to set the appropriate inputs depending on whether the FMU is reactive or delayed, or mealy and Moore.

We define the type of the FMU by applying the following rules, in order:

1. If the unit does not disclose any input-to-output feedthrough, it is assumed to be Mealy.
2. If at least one output variable depends instantaneously on an input variable, we assume that the unit is Mealy.
3. If the previous two do not apply, the unit is assumed to be Moore.
4. If the capability flag canInterpolateInputs is set, then the unit is reactive.
5. Otherwise, the unit is delayed.

To establish the equivalence of the couplings restrictions of units and those of FMUs, we note that the definition of algebraic dependency remains the same between FMUs. Thus, the dataflow graph can be built as described in the previous subsection.

Having established the equivalence between simulation units and FMUs, we will henceforth use the terms interchangeably.

**Semantic Adaptation**

The interface of an FMU (or of a simulation unit) comprises not only the specification of the inputs and outputs, but also how it is to be interacted with. It may be the case that in different co-simulation scenarios, the same FMU has to be interacted with differently (e.g., for accuracy/performance concerns). While modifying the orchestrator to support a new interaction pattern will solve the problem, it is not ideal since: (i) the interaction pattern may be specific to a single FMU (therefore not reusable), and (ii) modifications to the orchestrator may require extensive testing to ensure that it retains its correctness properties (e.g., see).

Our work avoids changes to the underlying orchestration algorithm, and focuses those changes around the FMU itself in the form of semantic adaptations, using hierarchical co-simulation.

An adaptation targets an FMU, or group of FMUs, which we will call the internal FMU(s), and the end result of an adaptation is a new FMU, which we call external FMU. The external FMU interacts with the internal FMU(s), without requiring them to be modified (modularity). The adjectives internal and external reflect the hierarchical nature of our approach and are illustrated in Figure 2.
Adaptation, Time and Partial Derivative Adaptation, Accurate Threshold Crossing, and Re-Initialisation.
See and references thereof, for variants of these adaptations.

Conversion of Units and Reference Frame Translation. The conversion of units and reference frame adaptations, take an internal FMU and create an external FMU whose inputs/outputs are algebraic transformations of the input/outputs of the internal FMU.

Interpolation/Extrapolation of Inputs. An FMU that stands for a continuous system, such as a DC motor, approximates its behavior trace by discretizing the time continuum into a finite set of points and applying a numerical method at each of those points.

In co-simulation, when the orchestrator asks an FMU to compute the behavior trace over an interval of time, from $t$ to $t + H$, the FMU discretizes the interval and computes the internal solution at each of these points, called micro-steps. The most common FMUs assume that, in between $t$ and $t + H$ the inputs provided by the orchestrator are constant. Naturally, for large $H$, this assumption causes a significant error in the co-simulation.

Instead of reducing $H$, it is possible to adapt the FMU to better approximate its inputs. Essentially, the external FMU discretizes the interval $t \rightarrow t + H$ and runs the state transition function of the internal FMU multiple times, providing an approximated input at each of the time points. The internal FMU will still assume a constant input, but will do so in smaller intervals of time.

Fixed Point Iteration. If an algebraic loop exists, then the involved units will belong to the same cycle in the corresponding dependency graph.

As proposed in, given a co-simulation scenario (recall Equation (2)) that has one cycle involving at least two simulation units (non-trivial), it is possible to create an external FMU that replaces all the units in the cycle. All the couplings external to the cycle become couplings to the hierarchical simulation unit.

At each state transition of the external FMU, a fixed point iteration technique is applied to the inputs/outputs of the internal FMUs.

If a scenario has multiple non-trivial cycles, this adaptation can be applied to reduce the scenario to one where all the algebraic loops are solved. Algorithm 1 can then be applied to compute the co-simulation.

Multi-Rate Adaptation. For FMUs simulating first order Ordinary Differential Equations (ODE), the larger the interval between the points, the less accurate the computed behavior trace will be.

The multi-rate adaptation is used to increase the accuracy while not sacrificing the performance in a co-simulation. Applied to co-simulation, the technique, well known in the circuit simulation domain, consists of a groups of interconnected internal FMUs that communicate more frequently. This can serve two purposes: optimize the communication cost between the internal units, or optimize the accuracy of the co-simulation (especially when the internal units are physically tightly coupled).

Similarly to the input extrapolation adaptation, the state transition function of the external unit instructs the internal units to perform multiple steps and exchange values at each of those steps. The higher the rate of the adaptation, the higher the number of internal steps performed.

This adaptation can be combined with the approximation of inputs adaptation, to provide for approximated inputs at each of the internal state transition invocations.

Time and Partial Derivative Adaptation. Time and partial derivative information about each simulation unit’s outputs can be used to optimize the co-simulation process in many different ways (e.g., see).

In the FMI standard, since the FMUs can optionally provide time and partial derivative information, it is often the case that some units do not support it. To mitigate this, a derivative adaptation can be used to produce an external FMU that provides (numerically estimated) partial and time derivatives.

Accurate Threshold Crossing. A co-simulation trace is more accurate if all units exchange values at the time when a certain signal crosses a given threshold. The problem of accurately finding that time is well known in the hybrid system simulation domain and many techniques exist to address it. In FMI co-simulation, the most basic technique to accurately locate a crossing consists of rejecting a step size and proposing a new one, that possibly coincides with the threshold crossing moment.

The accurate zero crossing adaptation ensures that the external FMU rejects the proposed step size when one of the inputs of the internal FMU crosses a significant threshold.

Re-Initialisation. An internal FMU that is expecting a smooth input signal may yield unexpected behavior trace when given a discontinuous signal (we consider a discontinuous signal to be a sufficiently rapid changing one in between co-simulation communication points). For example, an FMU that is using a multi-step numerical solver which assumes the input to be continuous (see, e.g., for a possible solution to this problem).

A re-initialization adaptation ensures that the external unit: (1) locates accurately the time of the discontinuity (e.g., in the same manner as the accurate crossing adaptation), and (2) the external unit is properly reset before handling the new value...
of the input. In the FMI standard, item (2) requires three steps: save the unit state; reset and initialize the unit; and restore the state.

**Quantization.** Quantization is an adaptation commonly used to convert a continuous signal into a discrete event one. The (continuous) set of possible input values is discretized into regions and, during the co-simulation, whenever the continuous signal enters a new region, an event is produced\(^\text{13,53}\).

In co-simulation, this adaptation transforms an internal FMU that expects continuous inputs and produces continuous outputs, into an external FMU that deals with events (see, e.g., \(^\text{9,14,26,68}\)).

The realization of this adaptation is very similar to the zero crossing one, except that the thresholds to locate are induced by the input space discretization.

**Hold.** The hold family of adaptations can be seen as the dual of the multi-rate adaptations.

If an internal FMU should run slower than the rest of the simulation units, then it can be adapted with a hold adaptation. The external FMU will trick the orchestrator and obey to the proposed step sizes, but will avoid executing the internal FMU every time a step is requested. For example, if a zero order Hold adaptation is used, then the external unit will produce an output that is equal to the most recent output produced by the internal unit.

There are many variants of this adaptation, with varying degrees of accuracy, borrowed from well known approximation techniques\(^\text{10}\).

The two adaptations below are novel in FMI based co-simulation domain, but well known in the discrete event domain.

**Data Triggered Execution.** The data triggered execution is an adaptation most useful when the modeller knows that a particular internal FMU will only produce relevant behavior when certain conditions are true over its inputs. The adaptation executes the internal FMU only when these conditions are met.

**Timed Transitions.** The time transition adaptation can be used when the internal FMU is known to have internal state changes, triggered after a known amount of time. The adaptation will query the internal FMU to know when exactly should the next state transition function call take place, and will call it only when that time is arrived. It can be combined with the data triggered execution to achieve a lazy execution of units.

Each of the semantic adaptations described above has many variants that make its ad-hoc implementations not only error prone, but also tedious. Additionally, one can extract the shared commonalities in the implementation of all semantic adaptations. The interplay between many small variants and shared commonalities is one of the motivating factors to use a Domain Specific Language for the description of the adaptations.

**Domain-Specific Languages**

Domain-specific languages (DSLs) offer a way to deal with the essential complexity of a given domain, while avoiding its accidental complexity\(^\text{51}\).

We highlight two important advantages that come from the use of a DSL in the context of our contribution:

1. The most common tasks in the target domain are performed in a very simple, productive, and intuitive manner (for a trained domain expert) — the descriptions made in our DSL do not deal with the idiosyncrasies of an implementation of the FMI Standard, even though a FMI compliant external FMU can be generated.

2. By maximally constraining the user, a DSL allows the user to specify extra information that can be used to detect extra information that can be used to detect mistakes (a simple validation being the compatibility of units in inputs/outputs).

**Running Example**

To showcase the language, the case study we present is adapted from a power window system, described in\(^\text{33,67}\). This system is the familiar automated car window, which responds to the driver/passenger pressing up/down buttons to raise/lower it. If an obstruction is detected, the window retracts for a few moments to avoid injury. This example was chosen for its heterogeneity and need for semantic adaptations.

**The Example Scenario**

Figure 3 shows the co-simulation scenario of the power window, consisting of five FMUs, with the illustrated input and output ports. The figure is a block representation of a co-simulation scenario as described in Equation (2). The FMUs were produced by the authors using independent tools.

The environment FMU, coded manually, is an abstraction of the behavior of the driver and passenger. Whenever the driver/passenger pushes a button up/down, the respective output will pulse to signal the event. When the button is released, the stop output pulses.

The controller FMU, produced from the Yakindu Statecharts tool, represents the software subsystem that ensures the safe operation of the window. It gets boolean pulse inputs and decides whether the motor should go up or down, through its boolean pulse outputs. If an object is detected (that is, obj_detected pulses) and the passenger (or driver) has pushed the up button, then the controller should instruct the DC motor to go down for one second. This is done by...
pulsing the down output and, after 1 second, pulsing the stop output.

The power is an ODE based unit, exported with OpenModelica, representing the DC motor and the up/down switched circuit that drives the motor. Whenever the $u$ input is bigger than 0.5, the DC motor moves the window up. Analogously, whenever the $d$ input is bigger than 0.5, it moves the window down.

The window and obstacle are stateless units, coded manually, that map the inputs to the outputs using algebraic equations. The obstacle FMU outputs a force proportional to how compressed it is. Non-zero compression happens only when the input displacement exceeds a given threshold (e.g., 0.45m).

An object is detected when the armature_current spikes, caused by a sudden increase in the reaction_torque input of the DC motor, cause in turn by an increase in the reaction_force of the object being compressed.

As illustrated in the figure, all units in this example are reactive, so the controller, power, window, and obstacle form a single cycle. The power and controller are Moore and the remaining units are Mealy.

Figure 4 shows the behavior trace of the example produced via a monolithic model produced in OpenModelica. In the figure, the driver continuously pushes the up button, asking the controller to move the window up, but the controller detects an object at about 2.5 seconds (due to the armature current spike), which causes it to override the requests of the driver and retract the window for 1 second.

**Semantic Adaptations**

The scenario presented in Figure 3 cannot be used as is to compute a co-simulation as the one shown in Figure 4 because the FMUs are incompatible.

The adaptations that need to be made were introduced in the “Background” section, and are detailed in the list below and illustrated in Figure 5. These will be referred to throughout this document.

- **lazy_sa** – for controller:
  - execute only if the inputs change (data triggered execution).

- **controller_sa** – for lazy sa:
  - map the armature_current to a boolean signal object_detected that is true whenever there is a threshold crossing. The condition that defines the crossing is $|\text{armature_current}| > 5$ * and the lazy_sa unit should be invoked at the time of crossing.
  - convert output, taking into account the stop signal.

- **window_sa** – for window:
  - negate the reaction_torque value;

*The value 5 is used here for the purposes of illustration. In practice, it is obtained by calibration with the DC Motor.
• convert the units of height from centimetres to metres.

**power**-sa – for power:
• ignore the algebraic loop between controller and power, and between the power and window, by delaying the outputs of the power by one co-simulation step. This effectively makes the external FMU a delayed unit.

**loop**-sa – for window**-sa** and obstacle:
• solve the algebraic loop between obstacle and window*sa* by successive substitution providing an initial guess for height.

**rate**-sa – in order to prevent divergence in the fixed point iteration caused by the above adaptation, smaller communication step sizes should be taken between the obstacle and the window FMUs. To this end:
• use a multi-rate adaptation, where loop**-sa** is executed 10 times faster than the remaining scenario.
• interpolate the input signal motor_speed.

### Hierarchical Co-simulation for Semantic Adaptation

The most straightforward way of dealing with semantic adaptations is by creating a master algorithm that implements them. There are multiple problems with this approach: 1) it forces the master algorithm to be specific to the scenario, which hinders the potential for reuse; and 2) it violates the transparency principle by not allowing the FMU (plus adaptations) to be easily imported onto other tools that perform co-simulation, such as Simulink®, INTO-CPS [57], or DACCOSIM [40].

To avoid these problems, we implement the semantic adaptations as FMUs, in a hierarchical way. In fact, our language defines semantic adaptations (plus internal FMUs) as FMUs themselves, allowing for adaptations to be described "on top of" other adaptations. This way, the orchestrator and semantic adaptations can be clearly separated, as well as the semantic adaptations between themselves.

As part of our contribution, we extend the definitions provided in Section "Background" to explain what hierarchical co-simulation is, and we give an overview on how the main semantic adaptations are implemented.

### Hierarchical Co-simulation

Before giving the formal definition of hierarchical co-simulation, we start with an example of a "default" hierarchical co-simulation unit and does.

A default hierarchical simulation unit is one that wraps a set of connected internal units, along with their inter-dependencies, and behaves in a manner that is indistinguishable from any other simulation unit. The internal FMUs have internal inputs/outputs (in between the units) and external inputs/outputs. This is called the default hierarchical unit because it does not adapt the behavior of the internal units. It merely wraps them.

To give details about how the default hierarchical unit is constructed, we extend the definition of co-simulation scenario to make the distinction between internal and external inputs. Let \( u_{ext} \) denote the input vector that is external to the co-simulation scenario. A co-simulation scenario with \( D = \{1, \ldots, n\} \) units, and with external input \( u_{ext} \), is then described as:

\[
\begin{align*}
\left\{ \begin{array}{l}
\mathbf{x}(t + \bar{H}_i, \bar{H}) = F_i(t, H, \ldots) \\
y_i(t) = G_i(t, \ldots) \\
u_i(t) = c_i(u_{ext}(t), y_1(t), \ldots, y_{i-1}(t), y_{i+1}(t), \ldots, y_n(t)) \\
x_i(0) = \text{Init}(\ldots)
\end{array} \right.
\end{align*}
\]

for each \( i \in D \)

(3)

Given then a co-simulation scenario as defined in Equation (3), and assuming that the topological order \( \sigma : \mathbb{N} \to D \) is well defined, the default hierarchical reactive Mealy FMU is constructed by:

1. aggregating the state \( x_i \) and the previous input \( u_{ip} \) of each FMU \( i \), into a single entity \( x \) that becomes the state of the hierarchical unit;
2. implementing the state transition function as a single co-simulation step of Algorithm 1.

Formally, the unit is defined as:

\[
\begin{align*}
\left\{ \begin{array}{l}
x(t + \bar{H}, \bar{H}) = F(t, H, x(t), u_{ext}(t + H)) \\
y(t) = G(t, x(t), u_{ext}(t)) \\
x(0) = \text{Init}(u_{ext}(0))
\end{array} \right.
\end{align*}
\]

(4)

where: \( x = [u_{p1}, \ldots, u_{pn}, x_1, \ldots, x_n]^T \) is the total state vector and \([\cdot]^T\) is the matrix transpose operation; the initial state vector is calculated by the \( \text{Init} \) function, defined in Algorithm 2, which finds the initial inputs and states to each of the internal units depending on their types; \( u_{ext} \) is the external input vector; function \( G \) is described in Algorithm 3, which computes the outputs of all internal units from the given inputs; and function \( F \) is detailed in Algorithm 4, which executes a single co-simulation step of Algorithm 1 and returns the minimum step size selected.

The construction of the default hierarchical reactive Moore, delayed Mealy, or delayed Moore, is done similarly and we omit it. The next subsection presents similar constructions for all kinds of units, incorporating adaptations.

The default hierarchical unit gives the basic transformation that underlies the semantic adaptation of one, or a connected group of, internal FMUs. In the subsection below, we describe the generic mechanism that enables the creation of hierarchical units with semantic adaptations.
Generic Semantic Adaptation

Previous work\cite{10,11} supports the hypothesis that any semantic adaptation can be described by the following elements, that mediate the interactions of the external FMU with the internal units:

- **internal input rules**, detailing how the values stored internally are mapped into inputs of the internal FMUs;
- **control rules**, determining what happens when the state transition function of the external FMU is invoked;
- **internal output rules**, describing how the outputs of the internal FMUs are stored in the state of the external FMU;
- **external output rules**, detailing how the values stored in the state of the external FMU are mapped to output values of the external FMU;

This subsection formalizes how a generic external FMU incorporating the above rules is constructed.

To formalize the above rules, we define the state of the external FMU. The external FMU is constructed from a given co-simulation scenario, defined in Equation (3), with \( D = \{ 1, \ldots, n \} \) units and external input vector \( u_{ext} \). Its state is then defined as

\[
\mathbf{x} = [x_{in}, x_{ctrl}, x_{out}, x_1, \ldots, x_n]^T
\]
with \( x_{in}, x_{ctrl}, \) and \( x_{out}, \) denoting the input, output and control storage vectors, respectively, and \( x_1, \ldots, x_n \) being the internal units’ states. The vectors \( x_{in}, x_{ctrl}, \) and \( x_{out} \) form the semantic adaptation storage and depend on the adaptations implemented in the external FMU.

Depending on the kind of external FMU being constructed, its initial state is computed by

\[
\text{Init}(u_{ext}) = [x_{in}, x_{ctrl}, x_{out}, x_1, \ldots, x_n]^T
\]

or

\[
\text{Init}() = [x_{in}, x_{ctrl}, x_{out}, x_1, \ldots, x_n]^T
\]

where \( \text{Init}() \), to be detailed shortly, makes use of the initialization functions \( \text{Init}_i \) of the internal units to get their initial states.

We now introduce the formal representation of the semantic adaptation rules, introduced at the beginning of this subsection:

- The application of the external input rules to the provided input is

\[
\text{In}([x_{in}, x_{ctrl}, x_{out}]^T, u_{ext}) = \tilde{x}_{in}
\]

- The application of the internal input rules to create the internal input vector is denoted as

\[
\text{MapIn}([x_{in}, x_{ctrl}, x_{out}]^T, h, dt) = [\tilde{u}_1, \ldots, \tilde{u}_n]^T
\]

This function is used whenever the input to any of the internal units needs to be computed. It is used in the \( \text{Ctrl} \) rules (defined next) and in the output function of the external unit.

In most adaptations, this function is invoked immediately before a call to the state transition function \( F_i \) of any internal unit. In line with the FMU interface, \( h \) is the communication step size that will be passed to the state transition \( F_i \) invocation, \( dt \) is the displacement of the time in unit \( i \), relative to the external unit, and \( \tilde{u}_i \) denotes the vector that will be used as external input to unit \( i \), or ignored if the unit does not depend on the external input. Multiple calls to this function can be made: potentially one per internal state transition call.

- The application of the control rules, to compute the new state \( \tilde{x}_i \) of each internal unit \( i \), the step size advanced \( \tilde{H} \), and the new control/output storage state \( \tilde{x}_{ctrl}, \tilde{x}_{out} \) of the semantic adaptation, is

\[
\text{Ctrl}(t, H, [x_{in}, x_{ctrl}, x_{out}]^T, [x_1, \ldots, x_n]^T) =
\begin{bmatrix}
\tilde{x}_{ctrl}, \tilde{x}_{out}, \tilde{x}_1, \ldots, \tilde{x}_n^T \\
\tilde{H}
\end{bmatrix}
\]

This function invokes the \( \text{MapIn} / \text{MapOut} \) functions before/after a state transition of an internal unit is invoked.

- The application of the internal output rules

\[
\text{MapOut}([x_{in}, x_{ctrl}, x_{out}]^T, [y_1, \ldots, y_n]^T, h, dt) = \tilde{x}_{out}
\]

Analogously to the \( \text{MapIn} \), the invocation of this function is controlled by the \( \text{Ctrl} \) functions. Parameters \( h \) and \( dt \) denote the communication step size, and time displacement, passed as arguments to the most recently invoked state transition function \( F_i \).

- The application of the external output rules to compute the external outputs, from the semantic adaptation state

\[
\text{Out}([x_{in}, x_{ctrl}, x_{out}]^T) = y
\]

Intuitively, the internal input/output functions serve to decouple the rate of execution of the internal units, from the rate of execution of the external FMU.

A semantic adaptation is a concrete definition of:

- Storage structure — \( x_{in}, x_{ctrl}, \) and \( x_{out} \);
- Initialization — \( \text{Init}() \);
- External input rules — \( \text{In}() \);
- Internal input rules — \( \text{MapIn}() \);
- Control rules — \( \text{Ctrl}() \);
- Internal output rules — \( \text{MapOut}() \);
- External output rules — \( \text{Out}() \).

We now describe how these functions are used in the specification of an external FMU.

The generic external unit is defined exactly as a simulation unit (recall Equation (1)):

\[
\left\{ \begin{array}{l}
x(t + H) = F(t, H, x(t), u_{ext}(t + H)) \text{ or } u_{ext}(t)) \\
y(t) = G(t, x(t), u_{ext}(t)) \text{ or } G(t, x(t)) \\
x(0) = \text{Init}(u_{ext}) \text{ or } \text{Init}()
\end{array} \right.
\]

where \( x = [x_{in}, x_{ctrl}, x_{out}, x_1, \ldots, x_n]^T \) denotes the state of the external FMU. Both an external reactive or delayed unit has the same implementation of \( F \), described in Algorithm 5 (but note that the definition of \( \text{Ctrl} \) will likely differ). The definitions of \( G \) differ for a Mealy or Moore external unit, and are detailed in Algorithm 6.

In Algorithm 6, we stress the following:

- The definitions take into account that it may not be possible to sort the internal units topologically, so the semantic adaptations support dependency cycles.
- Multiple calls to \( G \) can be made without changing the state of the external unit.
- If a Moore external FMU has at least one internal unit which depends on external input, then this input must be stored in the input storage \( x_{in} \) of the semantic adaptation by the \( \text{In} \) function (Line 2 of Algorithm 5), and then retrieved by the \( \text{MapIn} \) function (Line 8 of Algorithm 6).

To make these definitions easier to understand, we provide two examples: the default reactive Mealy hierarchical unit presented in the sub-previous section, and the algebraic loop semantic adaptation that involves the obstacle and window_sau units of the power window example (loop_sau).
Algorithm 5: State transition function of the generic external FMU, defined in Equation (5).

1 Function

\[ F(t, H, [x_{in}, x_{ctrl}, x_{out}, x_1, \ldots, x_n]^T, u_{ext}) \]

2 \[
\hat{x}_{in} := In([x_{in}, x_{ctrl}, x_{out}]^T, u_{ext});
\]

3 \[
\langle \hat{x}_{ctrl}, \hat{x}_{out}, \hat{x}_1, \ldots, \hat{x}_n \rangle^T, H := Ctrl(t, H, [x_{in}, x_{ctrl}, x_{out}]^T, [x_1, \ldots, x_n]^T);\]

4 return \[ \langle [\hat{x}_{in}, \hat{x}_{ctrl}, \hat{x}_{out}, \hat{x}_1, \ldots, \hat{x}_n]^T, \hat{H} \rangle; \]

5 end

The default reactive Mealy hierarchical unit can be informally described as follows:

- The state \( x_{ctrl} \) of the semantic adaptation includes the previous inputs of the internal units;
- The \( \text{Init} \) function is analogue to the one described in Algorithm 2;
- The \( \text{In}, \text{MapIn}, \text{MapOut}, \) and \( \text{Out} \), are roughly identity functions;
- And the \( \text{Ctrl} \) function implements the body of \( F \), in Algorithm 4;

Formally, functions \( \text{Init} \) and \( \text{Ctrl} \) are defined in Algorithm 7, and:

\[
\begin{align*}
\text{Init}([x_{in}, x_{ctrl}, x_{out}]^T, u_{ext}) &= u_{ext} \\
\text{MapIn}([x_{in}, x_{ctrl}, x_{out}], h, dt) &= [\tilde{u}_1, \ldots, \tilde{u}_n]^T, \text{ with } \tilde{u}_i = x_{in} \\
\text{MapOut}([x_{in}, x_{ctrl}, x_{out}]^T, [y_1, \ldots, y_n]^T, h, dt) &= [y_1, \ldots, y_n]^T \\
\text{Out}([x_{in}, x_{ctrl}, x_{out}]^T) &= x_{out}
\end{align*}
\]

The second example refers to the adaptation loop \( \text{loop}_{sa} \), which essentially performs a fixed point iteration between the obstacle and window \( \text{sa} \) units, computing improved values for their input/outputs via successive substitution.

The external FMU, called \( \text{loop}_{sa} \) in Figure 5 is a reactive Moore unit, and has an input \( u_{ext} \in \mathbb{R}^2 \) with two dimensions — displacement and speed — and one output — tau. Whenever the state transition of the external FMU is called, a successive substitution is performed between the two internal units, using the most recently found value of disp as an initial guess. Formally, let the index 1 refer to the window \( \text{sa} \) unit, and 2 to obstacle, so that, e.g., \( u_{\text{w2}} \) refers to the input to the obstacle unit. For the sake of simplicity, we assume that the system starts with all inputs/outputs being zero. Then, the functions that characterize the adaptation are shown in Equation (7). Note that had we not assumed that the system starts with zero inputs/outputs, the \( \text{Init} \) would have to compute a fixed point iteration to find a consistent initial state. This is possible with our formalization.

Algorithm 6: Output functions of the generic external FMU, per kind of unit, defined in Equation (5).

1 Function \( G(t, [x_{in}, x_{ctrl}, x_{out}, x_1, \ldots, x_n]^T, u_{ext}) \)

2 \[ \hat{x}_{in} := In([x_{in}, x_{ctrl}, x_{out}]^T, u_{ext}); \]

3 if \( \sigma \) is defined then

4 for \( i = 1, \ldots, n \) do

5 \[ \tilde{u}_i := y_i := \tilde{y}_i := 0; \]

6 end

7 for \( j \in (1, \ldots, n) \) do

8 \[ [\tilde{u}_1, \ldots, \tilde{u}_n]^T := \text{MapIn}([x_{in}, x_{ctrl}, x_{out}]^T, 0, 0); \]

9 \[ u_{\sigma(j)} := c_{\sigma(j)}(\tilde{u}_{\sigma(j)}); \]

10 \[ y_{\sigma(j)} := G_{\sigma(j)}(t, x_{\sigma(j)}, u_{\sigma(j)}); \]

11 \[ \hat{x}_{out} := \text{MapOut}([x_{in}, x_{ctrl}, x_{out}]^T, [y_1, \ldots, y_n]^T, 0, 0); \]

12 end

13 else

14 \[ \hat{x}_{out} := x_{out}; \]

15 end

16 return \( y := \text{Out}([x_{in}, x_{ctrl}, \hat{x}_{out}]^T); \)

17 return \( y; \)

18 end

1 Function \( G(t, [x_{in}, x_{ctrl}, x_{out}, x_1, \ldots, x_n]^T, u_{ext}) \)

2 \[ \hat{x}_{in} := In([x_{in}, x_{ctrl}, x_{out}]^T, u_{ext}); \]

3 if \( \sigma \) is defined then

4 for \( i = 1, \ldots, n \) do

5 \[ u_{c_i} := y_i := \tilde{y}_i := 0; \]

6 end

7 for \( j \in (1, \ldots, n) \) do

8 \[ [\tilde{u}_1, \ldots, \tilde{u}_n]^T := \text{MapIn}([x_{in}, x_{ctrl}, x_{out}]^T, 0, 0); \]

9 \[ u_{\sigma(j)} := c_{\sigma(j)}(\tilde{u}_{\sigma(j)}); \]

10 \[ y_{\sigma(j)} := G_{\sigma(j)}(t, x_{\sigma(j)}, u_{\sigma(j)}); \]

11 \[ \hat{x}_{out} := \text{MapOut}([x_{in}, x_{ctrl}, x_{out}]^T, [y_1, \ldots, y_n]^T, 0, 0); \]

12 end

13 else

14 \[ \hat{x}_{out} := x_{out}; \]

15 end

16 return \( y := \text{Out}([x_{in}, x_{ctrl}, \hat{x}_{out}]^T); \)

17 return \( y; \)

18 end

The next section describes a DSL for the definition of such semantic adaptations. The examples provided in that section clarify the need for the semantic adaptation functions defined in the current section.
Algorithm 7: \textit{Init} and \textit{Ctrl} functions of the default reactive Mealy hierarchical unit.

1 \textbf{Function} \textit{Init}(u_{ext})
2 \hspace{1em} \textbf{for} \ i = 1, \ldots, n \ \textbf{do}
3 \hspace{2em} x_i := u_{pi} := y_i := 0 ;
4 \hspace{2em} \textbf{end}
5 \textbf{end}
6 \textbf{for} \ j \in \{1, \ldots, n\} \ \textbf{do}
7 \hspace{2em} \uparrow \sigma(j) := \langle \textit{Init}_1(\uparrow \sigma(j)) \rangle ;
8 \hspace{2em} \textit{In}(x_{\sigma(j)}) := \textit{Init}_1(\uparrow \sigma(j)) ;
9 \hspace{2em} \textbf{end}
10 \textbf{return} \begin{bmatrix} x_{in}, x_{ctrl}, x_{out}, x_1, \ldots, x_n \end{bmatrix}^T ;

Algorithm 8: \textit{Ctrl} function of external FMU loop\textsubscript{sa}, illustrated in Figure 5.

1 \textbf{Function} \textit{Ctrl}(t, H, [x_{in}, y_{p1}, x_{out}]^T, [x_1, x_2]^T)
2 \hspace{1em} u_{1} := 0 ;
3 \hspace{2em} y_{1} := y_{p1} ;
4 \hspace{2em} u_{2} := y_{p2} := y_{2} := 0 ;
5 \hspace{2em} \overline{u}_1, \overline{u}_2 := \textit{MapIn}([x_{in}, y_{p1}, x_{out}]^T, 0, 0) ;
6 \hspace{2em} \textbf{for} \ i \in \{1, \ldots, \text{MAX ITERATIONS}\} \ \textbf{do}
7 \hspace{3em} u_{2} := c_2(\overline{u}_2, y_{1}) ;
8 \hspace{3em} \langle \tilde{x}_2, \tilde{H}_2 \rangle := F_2(t, x_{in}, u_{2}) ;
9 \hspace{3em} y_{2} := G_2(t + \tilde{H}_2, \tilde{x}_2, u_{2}) ;
10 \hspace{3em} u_{4} := c_4(\overline{u}_1, y_{2}) ;
11 \hspace{3em} \langle \tilde{x}_1, \tilde{H}_1 \rangle := F_1(t, x_{in}, u_{4}) ;
12 \hspace{3em} y_{1} := G_1(t + \tilde{H}_1, \tilde{x}_1, u_{4}) ;
13 \hspace{3em} \textbf{if} \ |y_{1} - y_{p1}| \approx 0 \ \textbf{and} \ |y_{2} - y_{p2}| \approx 0 \ \textbf{then}
14 \hspace{4em} \tilde{x}_{out} := \textit{MapOut}([x_{in}, x_{ctrl}, x_{out}]^T, [y_{1}, y_{2}]^T, 0, 0) ;
15 \hspace{4em} \textbf{break} ;
16 \hspace{3em} \textbf{else}
17 \hspace{4em} y_{p1} := y_{1} ;
18 \hspace{4em} y_{p2} := y_{2} ;
19 \hspace{3em} \textbf{end}
20 \hspace{2em} \textbf{return} \begin{bmatrix} y_{1}, \tilde{x}_{out}, [\tilde{x}_1, \tilde{x}_2]^T, H \end{bmatrix} ;
21 \hspace{1em} \textbf{end}

A DSL for Semantic Adaptation

We introduce a DSL for the specification of the set of rules introduced in the previous section (which form a semantic adaptation). Since research in semantic adaptation is ongoing, the language should be expressive enough to cover future semantic adaptations. Additionally, the implementation should not violate the modularity and transparency principles.

To these ends, the DSL—named baseSA—mixes imperative concepts with convenient functions that perform common operations on simulation units. A description made in this DSL can be used to generate hierarchical units.

The language and the examples used in this paper are available for download*.

The baseSA allows the description of the internal FMUs and their couplings (that is, the internal scenario as described in Equation (3)), and how semantic adaptation rules \(\text{Init}, \text{In}, \text{MapIn}, \text{Ctrl}, \text{MapOut}, \text{and} \text{Out}\), are implemented.

The remainder of this section is organised as follows. First, the baseSA DSL is introduced by describing the semantic adaptations of the running example and what their intended meaning is. Then, a

\*https://msdl.untwerpen.be/git/claudio/

HybridCosimulation

\[ \text{Init}(u_{ext}) = [0, 0, 0, \text{Init}_1(0), \text{Init}_2(0)]^T \]
\[ \text{In}([x_{in}, x_{ctrl}, x_{out}]^T, u_{ext}) = u_{ext} \]
\[ \text{MapIn}([x_{in}, x_{ctrl}, x_{out}]^T, h, dt) = [x_{in}, 0]^T \]
\[ \text{MapOut}([x_{in}, x_{ctrl}, x_{out}]^T, [y_1, y_2]^T, h, dt) = y_1 \]
\[ \text{Out}([x_{in}, x_{ctrl}, x_{out}]^T) = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix} x_{out} \]

\(Ctrl\) is defined in Algorithm 8.
more detailed description of the language (syntax and semantics) is provided.

**The baseSA DSL**

```
1  semantic adaptation reactive really WindowSA windowSA
2  at "./path/to/WindowSA.fmu"

3  for inner fmu Window window
4  at "./path/to/Window.fmu"
5  with input ports displacement (rad), speed (rad/s), reaction_force (N)
6  with output ports height (m), reaction_torque (Nm)

7  in rules
8  {
9      true -> 0 -> {
10         windowSA.tau := - window.reaction_force;
11     }
12  }

13  output ports disp (m) <- window.height, tau

14  with output ports height (cm), reaction_torque (N.m)

15  with input ports displacement (rad), speed (rad/s), reaction_force (N)

16  for
17  at "./path/to/WindowSA.fmu"

18  at "./path/to/Window.fmu"
```

Listing 1: The simple data adaptation window.sa in baseSA.

Listing 1 shows the baseSA definition of the semantic adaptation that generates the window.sa in Figure 5. The first few lines (Line 1 and 2 in the example) declare the name of the semantic adaptation and where the resulting external FMU will be generated.

Following that, the internal scenario is declared. The example listing (Lines 4 – 7) declares a single internal FMU and its ports.

baseSA descriptions work by exclusion: the user only specifies what needs to be changed, and the rest is assumed from the information provided. Hence, Listing 1 only needs to declare the output ports of the external FMU (disp and tau), in Line 13, and how they get their values: disp gets its value implicitly from the height port, and tau gets its value explicitly (via the specification of output rules).

Lines 11–15 declare the output rules. These specify how the tau output port of the external FMU gets its value, and this is done by assigning it the value of the reaction_torque output port, of the window FMU. The examples declares a single output rule, but in general multiple output rules can be declared. In general, each output rule has three parts: a condition, a MapOutRule part (syntactically preceded by "\(\rightarrow\)"), and an OutRule part (syntactically preceded by "\(\Rightarrow\)"). The condition decides whether the rule should be applied, and the other two parts contribute to the definition of the corresponding functions MapOut and Out, respectively.

Following the exclusion principle, Listing 1 omits several bits of information about the external FMU, that are required for a full definition of a semantic adaptation: input ports; Init function; In function; MapIn function; and Ctrl function;

In general, this information is assumed by applying multiple conventions, detailed in Section “Semantics”. The intended behavior is to follow the default hierarchical unit definition wherever the information is omitted (recall Equation (6) and Algorithm 7). For the example in Listing 1, the following is applicable:

- The external FMU (windowSA) has an input port for every input port of any internal FMU that has no incoming connection. This means that windowSA has three input ports, each bound to the corresponding input port of the internal FMU window.
- Each of the input ports of the internal FMU that have no incoming connections, gets its value from the corresponding external input port declared by the previous convention. The implementation of bindings is made via a storage variable. In Listing 1, this means that an extra input rule is created to encode the transfer of values. The input storage variables are also created.
- Any output variable bindings are realized in a manner similar to the previous convention: add an output rule and declare the necessary output variables to perform the transfer of values.
- Any expression referring ot the output of any internal FMU, in the Out part of an output rule, is assumed to refer to the storage variable with the most recent value of that output (output variables are created for the outputs of each internal FMU). In Listing 1, this means that window.reaction_force, in Line 13, gets replaced by a reference to an output variable.
- After applying the previous two conventions, the implicit bindings are removed.

Prepared using sagej.cls
Listing 2: The adaptation window_sa in explicit form.

Listing 2 shows the same adaptation as Listing 1, after applying the conventions introduced above:

- All input ports and output ports of the external FMU are declared, with no implicit bindings defined.
- Input storage variables, and their initial values, are declared (stored_window_height, stored_window_reaction_torque, stored_window_displacement, stored_window_speed). These are part of the \( x_{in} \) state vector of the semantic adaptation.
- Output storage variables, and their initial values, are declared (windowSa.disp, windowSa.tau), comprising part of the \( x_{out} \) state vector.
- A parameter per storage variable is added to allow the configuration of the initial value of that variable (technical detail: the parameters are mapped to FMI parameters).
- Input rules, as the one in Lines 31–41, are in general comprised of two parts: the InRule part, which in the example assigns values to the input storage variables; and the MapInRule part, which assigns the stored values to the input ports of the internal FMUs in the example. These make up the respective functions \( In \) and \( MapIn \).
- The control rules make use of the special function \( H_{window} \), which automatically: uses the \( MapIn \) function to compute the inputs to the internal FMU \( window \), computes any extra internal input (this applies to internal interconnected units), invokes the state transition function of \( window \), with \( t \) and \( H \), and invokes the \( MapOut \) function to compute its outputs. \( do\_step \) also takes into account the type (Mealy/Moore and reactive/delayed) of the internal unit invoked. The returned value is the step size taken by the unit.
- Output rules defined the functions \( MapOut \) (which stores the outputs of \( window \) in the output storage variables), and \( Out \) (which sets the outputs of the external FMU from the output storage variables). Notice that the conversion of units between the height and disp ports is also done.

In any baseSA description, there is no need to define explicitly the initial state (computed by the \( Init \) function). It is inferred from the input, control and output storage variables, plus the information about the internal units (extracted from their xml description file).
the external output function is called, the values of the output ports returned (in the example, the value of tau) are the ones computed in the most recent state transition function.

The control block of Listing 3 implements Algorithm 8 with the following differences.

- As part of the semantics of the do_step function: the MapInRule and MapOutRule instructions (which are implicit in Listing 3 by convention) are executed automatically to set the inputs of the internal FMUs; and the inputs of each FMU, if unspecified by an assignment, are set according to the internal connectivity information declared in Lines 14–15.
- The convergence test (Line 33) is made only in the disp port (to simplify).
- The state manipulation of the internal FMUs is facilitated by the use of the save.state and rollback functions.

Listing 4: Adaptation that generates rate_sa.

```
semantic adaptation reactive more RateSA rate_sa
at "./path/to/RateSA.fmu"
for inner fmu LoopSA loop_sa
at "./path/to/LoopSA.fmu"
with input ports displacement (rad), speed (rad/s)
with output ports tau (N.m)
input ports speed
output ports tau <- loop_sa.tau

control var previous.speed := 0;
control var inner.time := t;
for (var i in 0 .. RATE) {
do_step(loop_sa, inner.time, micro_step);
inner.time := inner.time + micro_step;
}

previous.speed := current.speed;
return H;
}
}
in var current.speed := 0;
in rules {
true -> {
current.speed := speed;
} -> {
loop.sa.speed := previous.speed + (current.speed - previous.speed)*(dt + h);
};
}
```

Listing 5 implements adaptation lazy_sa. This adaptation assumes the default mappings for the inputs, but is declares them because they are referred to in the Ctrl block.

In general, every reference to an input port of the external FMU, made outside of the In block, is replaced with a reference to the variable that stores
the most recently given value of that. For example, the expression lazy.obj_detected, is replaced by the variable that stores that input.

The adaptation in Listing 5 performs two tasks: it keeps track of the previous value of each signal, and invokes the internal unit state transition function (i.e., do_step) whenever there is a change; and it keeps track of the next time to execute the internal unit (assuming that no inputs change) and invokes it when such time arrives, to cater for internal timed transitions. At the same time, the output signals are always available (held constant) because of the storage output variables.

Listing 6: Adaptation that generates controller_sa.

The adaptation controller_sa is shown in Listing 6. The control rules apply regula falsi to locate the crossing of the armature signal into the threshold T.

This example shows how the conditions in the output rules can be used to select which rules are applied. Informally, in general, at the end of each external state transition, when MapOut is invoked, all the conditions in the rules are evaluated. The ones that evaluate to true, are recorded as part of the x_out state. Afterwards, whenever Out is called, only the rules that evaluated to true contribute to the output of Out.

The power_sa adaptation was omitted due to its simplicity. It declares the external FMU as a delayed Moore and lists the output port bindings.

The above adaptations generate the FMUs for the co-simulation scenario illustrated in Figure 5. The orchestrator in Algorithm 1 then computes the results shown in Figure 6. Comparing these results with the ones in Figure 4, one sees that they are similar, except for the fact that the armature current has a higher peak in the co-simulation. This is because the threshold crossing adaptation was disabled, since the power FMU does not support rollback.

![Figure 6. Power window co-simulation results.](image-url)
In the following subsections, we describe the language (syntax and semantics) in more detail. The syntax is described using extended Backus–Naur form (EBNF) and the semantics are presented informally by describing a transformation of baseSA descriptions, to the Init, In, MapIn, Ctrl, MapOut, and Out functions, introduced in Line 13.

Syntax

The partial syntax of baseSA is detailed in Listing 7. We omit the definition of the most common symbols:
- **ID** is an identifier;
- **URL** is a URL;
- **PhysicalUnit** denotes any physical unit;
- **Expression** is an expression that defines a value, e.g., comparison, addition, constant, variable reference, etc.

Statement is a programming language statement. It includes if-statement, static for loop, local variable declarations, assignments, references to variables/parameters, built-in function calls, etc.

```plaintext
SemanticAdaptation = "semantic", "adaptation", KindInput, KindOutput, UnitInstance, UnitInterface,
  InnerUnits, 
  "input", "ports", Port, 
  "output", "ports", Port,
  [ParamDeclarations], [ControlRuleBlock], [InRulesBlock], [OutRulesBlock];
KindInput = "reactive" | "delayed";
KindOutput = "moore" | "mealy";

InRule = RuleCondition, "->", InRule, "-->", MapInRule, 
  InRulesBlock = ControlRule, VarDeclarations, 
  ControlRule = "control", "rules", "in", "rules", 
  ParamDeclarations = "param", SingleDeclaration, 
  SingleDeclaration = ID "=" Expression;
  Connection = ID "->" ID, ID "<-" ID;
  Port = ID, "[", PhysicalUnit "]", PortBinding;
  PortBinding = ("->", ID) | ("<-", ID);
  ControlRuleBlock = ControlRule, [InRulesBlock];
  InRulesBlock = "with", "input", "ports", Port,
  "in", "rules", 
  OutRulesBlock = RuleCondition "-->", MapOutRule "-->", OutRule "-->";
```

Listing 7: The (partial) EBNF grammar of baseSA.

Table 1 summarizes the special functions and variables.

The full grammar definition, and an editor of baseSA descriptions, developed with Xtext, is available for download *. Figure 7 shows the editor interface.

Semantics

In this subsection, we define the semantics by describing informally how each syntactic construction in baseSA is mapped to the definition of Init, In,
Table 1. List of built-in symbols and their meaning.

<table>
<thead>
<tr>
<th>Symbol</th>
<th>Availability</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>$t \in \mathbb{R}$</td>
<td>ControlRule ($\text{Ctrl}$)</td>
<td>Argument provided in the state transition function of the external FMU.</td>
</tr>
<tr>
<td>$H \in \mathbb{R}$</td>
<td>ControlRule ($\text{Ctrl}$)</td>
<td>Co-simulation step size passed as argument to the state transition function of the external FMU.</td>
</tr>
<tr>
<td>$do_step(f\text{mu}, t_i, h) \in \mathbb{R}$</td>
<td>ControlRule ($\text{Ctrl}$)</td>
<td>Asks an internal FMU to perform a co-simulation step and returns the size of the computed interval.</td>
</tr>
<tr>
<td>$h \in \mathbb{R}$</td>
<td>MapInRule, MapOutRule ($\text{MapIn}, \text{MapOut}$)</td>
<td>Co-simulation step size passed as argument to the state transition function of the internal FMU.</td>
</tr>
<tr>
<td>$dt \in \mathbb{R}$</td>
<td>MapInRule, MapOutRule ($\text{MapIn}, \text{MapOut}$)</td>
<td>Let $t_i$ denote the time given as argument to the state transition function of an internal FMU. Then $dt = t_i - t$.</td>
</tr>
<tr>
<td>$\text{save_state}(f\text{mu})$</td>
<td>ControlRule ($\text{Ctrl}$)</td>
<td>Stores the state of an internal FMU.</td>
</tr>
<tr>
<td>$\text{rollback}(f\text{mu})$</td>
<td>ControlRule ($\text{Ctrl}$)</td>
<td>Rolls back an internal FMU to the last saved state.</td>
</tr>
<tr>
<td>$\text{is_close}(x, y, \text{rtol}, \text{atol}) \in \mathbb{B}$</td>
<td>Everywhere</td>
<td>Approximate equality.</td>
</tr>
<tr>
<td>$\text{get_next_time_step}(f\text{mu}) \in \mathbb{R}$</td>
<td>Everywhere</td>
<td>Returns the maximum time step an internal FMU is willing to accept.</td>
</tr>
<tr>
<td>$\sin, \cos, \min, \ldots$</td>
<td>Everywhere</td>
<td>Implements the corresponding mathematical function.</td>
</tr>
</tbody>
</table>

... add an assignment $\text{stored}\_\text{sa}\_\text{ip} := \text{sa}\_\text{ip}$; in the $\text{MapInRule}$ part, for each input binding declared $\text{ip} \rightarrow f.\text{ip}$, create an assignment $f.\text{ip} := \text{ip}$. If units need to be converted, the right hand side of the assignment is replaced accordingly.

**RemoveInBindings** – For each input binding declared $\text{ip} \rightarrow f.\text{ip}$, replace it by just $\text{ip}$. Any physical unit declaration is also removed.

**AddOutPorts** – If no output ports are declared, create an output port declaration $\text{op} \leftarrow f.\text{op}$ per output port $\text{op}$ of each internal unit $f$.

**AddOutParams** – Analogous to **AddInParams**: for each output port declaration $\text{op} \leftarrow f.\text{op}$ of each internal FMU $f$, create a parameter declaration $\text{INIT}\_F.\text{OP} := \text{v}$ (if such parameter does not exist), with $\text{v}$ being the default value.

**AddOutVars** – Analogous to **AddInVars**: for each output port declaration $\text{op} \leftarrow f.\text{op}$ of each internal FMU $f$, create an output variable declaration $\text{stored}\_f.\text{op} := \text{INIT}\_F.\text{OP}$, if it does not exist already.

**AddOutRule** – Prepend a new output rule to the output rules block, with a true condition, and: in the $\text{MapOutRule}$ part, add an assignment $\text{stored}\_f.\text{op} := f.\text{op}$, per output port $\text{op}$ of each internal unit $f$; in the $\text{OutRule}$ part, for each output binding $\text{op} \leftarrow f.\text{op}$ declared, add an assignment $\text{sa}\_\text{op} := f.\text{op}$. If units need to be converted, the assignment is replaced accordingly.

**RemoveOutBindings** – Analogous to **RemoveInBins**: for each declared output binding $\text{op} \leftarrow f.\text{op}$, remove the binding (and any unit declaration), leaving just the output declaration $\text{op}$.

**CreateCtrlRules** – If there is no control rules block declared, create one, and: compute the topological order $\sigma$ of the internal scenario (if it cannot be computed, abort with an error); for each internal unit declaration $f$, in topological order, append $\text{var} H_f := \text{do\_step}(f, t, H)$; append (at the end of the block) either return $H_f$ if there is only one internal FMU, or return $\min(H_1, \ldots, H_n)$, where $H_i$ refers to each of the local variables declared in the previous assignments.

**ImplementInternalBinding** – For each connection in the internal scenario $f1.\text{op} \rightarrow f2.\text{ip}$, locate the $\text{do\_step}(f2, \ldots)$ instruction in the control rules block. Before this instruction, if there is no assignment of the form $f2.\text{ip} := \ldots$, insert $f2.\text{ip} := f1.\text{op}$ immediately before the instruction $\text{do\_step}(f2, \ldots)$.

**ReplacePortsRefsByVars** – For every input rule, go through the $\text{MapInRule}$ part and replace every reference to an external input port $\text{ip}$, by a reference to the $\text{stored}\_\text{sa}\_\text{ip}$ input variable. In the control rules block, replace every reference to an output port $\text{op}$ of an internal unit $f$ by a reference to the corresponding storage variable $\text{stored}\_f.\text{op}$. For each output rule, in the $\text{OutRule}$...
part, replace any reference to an output port op of an internal unit f by a reference to the corresponding storage variable stored_f_op.

Listing 2 is the result of applying the above rules to Listing 1.

**Mapping to Generic Semantic Adaptation.** Given a baseSA description in explicit form, we now explain how it is mapped to the formal definition of a generic external unit. In the generic external unit definition (recall Equation (5)), the elements that need to be defined are:

- The space of $x_{in}$, $x_{ctrl}$, and $x_{out}$;
- $Init(u_{ext})$ or $Init()$, depending on the kind of external unit;
- $In([x_{in}, x_{ctrl}, x_{out}]^T, u_{ext})$;
- $MapIn([x_{in}, x_{ctrl}, x_{out}]^T, h, dt)$;
- $Ctrl(t, H, [x_{in}, x_{ctrl}, x_{out}]^T, [x_1, \ldots, x_n]^T)$;
- $MapOut([x_{in}, x_{ctrl}, x_{out}]^T, [y_1, \ldots, y_n]^T, h, dt)$;
- $Out([x_{in}, x_{ctrl}, x_{out}]^T)$.

Each of the above elements is now defined.

Part of $x_{in}$ is determined by the input variables declared: $x_{in}$ has one dimension per declared input variable. The type of the dimension (real, boolean, etc.) corresponds to the type of the declared variable. In addition, $x_{in}$ has one boolean dimension per input rule. For example, is there are three numeric variables declared, and one input rule, then $x_{in} \in \mathbb{R}^3 \times \text{Bool}$.

Analogously to $x_{in}$, $x_{out}$ has one dimension per declared output variable, and an additional boolean dimension per declared output rule.

The control storage vector $x_{ctrl}$ has one dimension per declared control variable. Additionally, if the semantic adaptation is a reactive one and the initial baseSA description (not the explicit one) does not include any control rules, the $x_{ctrl}$ has one dimension per internal delayed unit.

The external input function

$$In([x_{in}, x_{ctrl}, x_{out}]^T, u_{ext}) = \bar{x}_{in} \tag{8}$$

is defined to perform the following steps in order:

1. Evaluate all conditions of the input rules in the order that they are declared, and for each condition, mark the corresponding location of $\bar{x}_{in}$ with the outcome (true or false).
2. For the input rules whose conditions evaluated to true in the previous step, execute the $InRule$ part, in the order that the rules are declared (this computes the remainder of $\bar{x}_{in}$).

Function

$$MapIn([x_{in}, x_{ctrl}, x_{out}]^T, h, dt) = [\bar{u}_1, \ldots, \bar{u}_n]^T \tag{8}$$

executes the $MapInRule$ part of the input rules whose condition evaluated to true (this information is stored in $x_{in}$) in order of their declaration. The executed input port assignments form $[\bar{u}_1, \ldots, \bar{u}_n]^T$.

Function

$$MapOut([x_{in}, x_{ctrl}, x_{out}]^T, [y_1, \ldots, y_n]^T, h, dt) = \bar{x}_{out}$$

is analogous to $In$. It evaluates all the conditions of the output rules in the order that they are declared, and for each of those conditions, marks the appropriate location of $\bar{x}_{out}$ with the outcome of the condition evaluation. Then it computes the remaining portion of $\bar{x}_{out}$ by executing the $MapOutRule$ part of each of the output rules whose conditions evaluated to true.

Function

$$Out([x_{in}, x_{ctrl}, x_{out}]^T) = y$$

is analogous to $MapIn$. It executes the $OutRule$ part of the output rules whose condition evaluated to true (in the order in which they are declared) to compute the output vector $y$.

The role of the initialization function (derived automatically from the baseSA description) is to find a consistent initial state, defining the initial values of the storage vectors $x_{in}$, $x_{out}$, and $x_{ctrl}$. If the semantic adaptation is declared as reactive, then $Init$ requires the initial input, according to Equation (1).

First, the parts of $x_{in}$, $x_{ctrl}$, and $x_{out}$ that correspond to the declared input/control/output variables are initialized according to the initial value that is declared for them.

If it exists, the part of $x_{ctrl}$ that corresponds to the previous inputs to the internal units is initialized by computing the initial input to all the internal units in the topological order (such order exists by assumption). This is similar to Algorithm 2, except that the functions $In$, $MapIn$, and $MapOut$, are invoked to adapt any external input to the internal units, and initialize the condition flags.

Function

$$Ctrl(t, H, [x_{in}, x_{ctrl}, x_{out}]^T, [x_1, \ldots, x_n]^T) = \langle \bar{x}_{ctrl}, \bar{x}_{out}, [\bar{x}_1, \ldots, \bar{x}_n]^T, \bar{H} \rangle$$

runs the instructions declared in the control rules block, in the order that they are declared. The assignments performed to control variables make up part of the output vector $\bar{x}_{ctrl}$. The executed assignments to the input ports of each internal FMU $i$, up to the instruction $do\ step(i, t_i, h_i)$, make up part of the unit input vector $u_i$.

Any variable reference in the control rules block refers to the most recently given value of that variable.

Each instruction $do\ step(i, t_i, h_i)$ maps to the following steps, performed in Ctrl:

- Invoke $MapIn$ function to compute the external input of unit $i$:

  $$[\ldots, \bar{u}_i, \ldots]^T := MapIn([x_{in}, \bar{x}_{ctrl}, x_{out}]^T, h_i, t_i - t) \tag{8}$$
Note that $\tilde{x}_{ctrl}$ represents the control state vector that was affected by the assignments made since the beginning of the execution of the $Ctrl$ function. $x_{in}$ and $x_{out}$ represent the (unchanged) vector provided as input to $Ctrl$.

- Merge the input vector $u_i$ computed by previous assignments with $\tilde{u}_i$ to form the unit input $uc_i$;
- Invoke the state transition function of the unit:

$$\begin{align*}
\langle \tilde{x}_i, \tilde{H}_i \rangle := F_i(t, H, x_i, uc_i \text{ or } u_p_i) \quad (9)
\end{align*}$$

- Get the output of the unit:

$$y_i := G_i(t + \tilde{H}_i, \tilde{x}_i, uc_i) \text{ or } G_{\sigma(j)}(t + \tilde{H}_i, \tilde{x}_i) \quad (10)$$

- Invoke the $MapOut$ function to compute an updated output storage vector:

$$\tilde{x}_{out} := MapOut([x_{in}, \tilde{x}_{ctrl}, x_{out}]^T, [y_1, \ldots, y_n]^T, \tilde{H}_i, t, t_i - t) \quad (11)$$

Finally, upon returning, if the external FMU is a reactive unit, and the initial baseSA description does not declare a control rules block, $Ctrl$ stores the most recent inputs provided to each delayed internal units in the $\tilde{x}_{ctrl}$ vector, to be used as delayed inputs in a subsequent external state transition call. This instruction is similar to Line 29 of Algorithm 7.

**Evaluation**

In this section, we judge how well our approach answers the research question posed in this work.

The requirements set by the research question are:

**Productivity** – Does the language have impact in the productivity of its users?

**Expressivity** – Is the language expressive enough to cover current and future needs?

**Modularity** – Does the internal FMUs need to be changed?

**Transparency** – Does the external FMU behave exactly as an FMU?

**Productivity**

In general, DSLs have the potential to boost its users’ productivity\(^{51,52}\). For baseSA, we describe an early experiment to assess the productivity.

**Goals.** Productivity is measured by comparing the time it takes for a trained user to: (1) create an external FMU using our DSL; and (2) code the same external FMU.

As a surrogate measure, we compare the approximate number of lines of code (LOC) required for a semantic adaptation coded by hand, with the LOC of the corresponding semantic adaptation expressed in baseSA.

**Experimental Setup.** As part of the development of the code generator, all semantic adaptations identified in Figure 5, except the rate\(_sa\), were coded by hand and the effort taken was recorded.

**Table 2. Effort in hand-coding hierarchical semantic adaptations.**

<table>
<thead>
<tr>
<th>Semantic Adaptation</th>
<th>LOC</th>
<th>Effort (man-hour)</th>
</tr>
</thead>
<tbody>
<tr>
<td>lazy(_sa)</td>
<td>700</td>
<td>9</td>
</tr>
<tr>
<td>controller(_sa)</td>
<td>750</td>
<td>24</td>
</tr>
<tr>
<td>power(_sa)</td>
<td>680</td>
<td>16</td>
</tr>
<tr>
<td>window(_sa)</td>
<td>690</td>
<td>8</td>
</tr>
<tr>
<td>loop(_sa)</td>
<td>690</td>
<td>16</td>
</tr>
<tr>
<td>Total</td>
<td>3510</td>
<td>73</td>
</tr>
</tbody>
</table>

As Table 2 shows, even though the semantic adaptations differ in complexity, they have a similar number of LOC. This is evidence that there is a large portion of code dedicated to common FMI-related management tasks. With baseSA, the user does not have to code:

- Memory management – The inputs, outputs, and local variables, of the external FMU are stored in dynamically allocated memory.
- Variable de-referencing – To set/get values to/from an internal FMU, a list of value references (integers which identify a variable) has to be provided. Any mistake here may cause the internal FMU to give wrong results, but not necessarily crash, which makes it hard to debug.
- State management – The external FMU has to support rollback, and for that, the state variables must be properly serialized and de-serialized. In the case study, each semantic adaptation requires approximately 140 LOC to implement the set/get state.
- Consistent inputs management – The external FMU which is reactive and has internal delayed units, has to keep track of the previous inputs to these.

**Threats to Validity.** LOC is only a surrogate measure for the productivity of a DSL, albeit a common one\(^{12}\), and depends on the programmer. However, the tasks described in the above list are handled automatically by the code generator of baseSA.

The values provided in Table 2 lack external validation. We intend to perform a second round of experiments, where we will ask a participant to code a semantic adaptation, then train him/her, and measure the effort it takes to code the same adaptation, in baseSA.
Expressivity

The baseSA DSL is imperative in the sense that it describes how the semantic adaptations are performed. However, it forces a structure in the definition of the semantic adaptations, aided by the distinction between data (input/output rules) and control adaptations. We argue that this structure does not restrict the expressiveness of the semantic adaptations.

To provide evidence for this, we describe how the adaptations used in the case study are representative of the semantic adaptations and coupling algorithms surveyed in [46].

Extrapolation/interpolation schemes These techniques, used in [17,20–22,35,76], are similar to the rate_sa.

Jacobi-based orchestration This orchestration algorithm, used in [10,23,36,38,40,48,54,83], is similar to the Gauss-seidel coupling except that it assumes that all units are delayed. A way to implement it as a semantic adaptation is to define a control rule that sets explicitly the inputs to the internal FMUs, and then invokes the do_step function on them.

Algebraic constraint couplings This coupling technique, reported in [5,49,74,75], can be implemented by a fixed point iteration (recall adaptation loop_sa) and extra algebraic computations on the units inputs and outputs.

Semi-implicit coupling These techniques, presented in [70–74], are similar to the ones above, except they perform two iterations only.

Error control Richard extrapolation, [6,40] can be implemented by creating a semantic adaptation which runs a whole scenario at twice the rate of the original one; Multi-order input extrapolation, [23,24] amounts to implementing two approximation schemes (see item above) and run in parallel; Embedded method, [50] requires that a semantic adaptation is implemented to perform a discretized numerical integration of some of the signals in the internal scenario; Energy based, [69] techniques can be implementing by coding semantic adaptations which monitor for energy dissipativity in some of the signals in the internal FMUs.

We do restrict the expressiveness of the language, with the intent of guaranteeing that it terminates:

- No function definitions are allowed;
- No recursive definitions of semantic adaptations are allowed;
- For loops must have a static range.

These restrictions make expressing some of the above techniques more cumbersome, but not impossible.

Modularity

The simulation unit specification, introduced in Equation (1) was shown to be a valid abstraction of an implementation of an FMU in Section “Background”. Furthermore, it is clear that changing the implementation of any of the functions Init, G, F implies a change in the FMU implementation. In Section “Semantics”, these functions are invoked as part of the implementation of each semantic adaptation, but never changed, thus showing that the corresponding FMU implementations are not affected by the implementation of the language.

Transparency

Section “Hierarchical Co-simulation for Semantic Adaptation” describes how a generic semantic adaptation forms a simulation unit that obeys the definition in Equation (1) (see Equation (5)). Furthermore, Section “Semantics” describes how a baseSA is implemented by “filling in” the semantic adaptation functions, that are used in Section “Hierarchical Co-simulation for Semantic Adaptation”. The semantics does not require the hierarchical unit definition, in Equation (5), to be changed. Therefore, our approach does not violate transparency.

Discussion and Future Work

This section discusses some of the characteristics and limitations of our contribution, and research opportunities for the future.

Automatic Semantic Adaptation Identification. Throughout this work, we assumed that the user knows that an adaptation is required in order to make the co-simulation possible. An interesting research direction is to explore how means can be employed in trying to identify the need for specific semantic adaptations.

Runtime Performance. Despite not being our primary goal, the performance of the generated FMU should be similar to a custom coded one. To this end, the code generator under development performs most tasks at compile time. However, we have not carried out any experiments to measure the performance of the generated code.

A research direction is to explore how to merge multiple adaptations, to avoid generating the intermediate hierarchical FMUs. For example, in Figure 5, adaptations loop_sa and rate_sa could be merged into one single adaptation, provided that the user has no intention of using loop_sa for other purposes. However, while it is clear what the result should be for this example, in general this is a non-trivial task: When can two arbitrary adaptations be merged?
Solving this problem brings a performance benefit, but also provides new insights into the nature of adaptations. In addition, one can ask: *If two semantic adaptations can be combined, are they commutative?* This question is important because it allows us to optimize: if there is a semantic adaptations which will cause rollbacks, we want it to be the first to execute, to avoid wasting computation. An example of this is the controller_sa, which performs the crossing location before the lazy_sa gets the opportunity to run.

Trying to answer the above questions will inevitably lead to another question: *what is the right level of abstraction to analyze the combination of semantic adaptations?* This question is related to the next discussion topic.

**Usability and Productivity.** As part of trying to find out what the right level of abstraction to analyze semantic adaptations is, we are developing a new language that allows for a more declarative description of semantic adaptations. This language, as opposed to baseSA, allows for a much more concise description of the most common semantic adaptations by just enumerating what semantic adaptations should be used to form the external FMU.

The descriptions made in this language compile to baseSA, whose role is to provide a solid foundation. The main benefits of using this language are:

1. The user does not to know how semantic adaptations are implemented.
2. It is minimal, meaning that it enables the user to specify common semantic adaptations (e.g., multi-rate, successive substitution) as concisely as describing them in natural language;
3. It further restricts the user into using well known semantic adaptations, which prevents mistakes.
4. It may provide insight into the research questions identified in the previous subsection.

Each adaptation has some degree of configuration. For example, the multi-rate is configurable with an input approximation adaptation. This highlights another interesting research direction, related to the combination of semantic adaptations: *how and when can semantic adaptations interface with each other?* In this example, it is clear that any input approximation adaptations can complement a multi-rate adaptation, but what are the essential characteristics of input approximation and multi-rate adaptations, that make them so compatible? The same question applies to output approximation adaptations (the family of Hold adaptations) and the lazy related ones. A possible direction to explore is to look at the object oriented world, and study how can semantic adaptations define interfaces and specialization, so that their interaction is well defined.

**Discrete Event FMU Implementation.** The current version of the FMI standard (version 2.0) lacks essential features to enable accurate hybrid co-simulation (see, e.g., [1][18][19][26][29][37]).

Until new extensions are made, there are many different ways in which a cyber system (e.g., a state chart) can be simulated in an FMU [29][30][37][64][65][78]. At least one of the implementations the authors used before (the Stategraph[44]), already includes semantic adaptations, to facilitate its integration with the FMI.

Our work shows that, when implementing an FMU that simulates a cyber system, it is best to leave as many semantic adaptations as possible out. The more adaptations an FMU already contains, the harder it is to adapt it to other contexts.

**Related Work**

Outside the context of FMI, the problem of composing and adapting operational semantics of multiple languages is discussed in [15][16][31][32][56][60][62][82] and references therein.

Within the context of FMI, we can divide the related works in two categories: (A) those whose prime purpose is to describe co-simulation scenarios; and (B) those that target the description of orchestration algorithms. Both these categories do not target primarily the description of semantic adaptations, but can potentially be extended to include simple descriptions. Due to our pure hierarchical co-simulation approach, our contribution complements any of these works.

Under Category (A), we highlight [29][57][40], and [25]. These works introduce a language for the description of a co-simulation scenario, with the purpose of running a co-simulation. The work in [40][57] assumes that a generic orchestration algorithm is used, whereas [25][79] aim at generating an orchestrator that is specific to the scenario described. Our DSL allows for the description of a co-simulation scenario, and a
specific master algorithm can be generated from that description.

DACCOSIM\textsuperscript{40} follows a related approach with respect to hierarchical co-simulation, allowing the scenario to grouped by computational nodes. In contrast to our work, this hierarchy is computational and not functional. Moreover, it is not transparent, as the distinction is made between local (internal to computational nodes) and global orchestrators. Nevertheless, each FMU is wrapped with code that performs error control, highlighting the need for semantic adaptation.

In Category (B), we highlight\textsuperscript{25, 41, and 8}. The work in\textsuperscript{25} allows the description of master algorithms using the Business Process Modelling Notation. We argue that the visual notation for the description of an orchestration algorithm works well for simple cases, with two units. However, when multiple semantic adaptations become necessary, or the number of simulation units increases, the visual notation rapidly becomes cluttered. The work does not describe any intention of using the notation to describe semantic adaptations, but the notation has an extension mechanism that can in principle be used to describe simple semantic adaptations.

The most related to our own is\textsuperscript{8}. It introduces an object oriented framework for co-simulation that allows for both the development of FMUs, as well as for orchestration algorithms, in C++. Class specialization is used extensively to maximize reuse, sharing some of the benefits with our contribution. The main difference to our work is the level of abstraction and the intention to use semantic adaptations. While their work is capable of expressing semantic adaptations, our work is targeted towards that purpose. One can position their work as helping develop FMUs for simulators that need to support the FMI Standard, and our work can be used to adapt already existing FMUs. Furthermore, the description of a complex adaptations such as rate\textsubscript{sa} is more compact in our DSL.

**Conclusion**

This paper addressed the problem of describing the most common semantic adaptations on multiple types of black box simulation units in a productive manner while avoiding the modification of the units (modularity) and tools for co-simulation (transparency).

To make this possible, we propose a DSL, available for download\textsuperscript{*}, that is both expressive (due to its imperative nature) but also productive (due to its conventions and high level constructs). Each description refers to a group of interconnected FMUs and dictates how those FMUs interact with the environment.

The essential mechanism that enables the semantic adaptations is the concept of hierarchical co-simulation, formalized in this work. The meaning of each adaptation is given by mapping it onto hierarchical co-simulation units, which in turn is mapped to units and FMUs, as illustrated in Figure 1.

The main distinguishing factor from the related work, is our focus in semantic adaptations for FMI based co-simulation, which imposes the modularity and transparency requirements.

This work opens up new opportunities for research into semantic adaptations, for example, how to find higher levels of abstraction to describe semantic adaptations, and explore how different semantic adaptations can interface and complement each other. We intend to explore these in the future.

**Funding**

This research was partially supported by Flanders Make vzw, the strategic research centre for the manufacturing industry, a PhD fellowship grant from the Agency for Innovation by Science and Technology in Flanders (IWT, dossier number 151067), and by the INTO-CPS project, funded by the European Commission's Horizon 2020 programme (grant number 664047).

**Acknowledgements**

We are grateful to Prof. Peter Gorm Larsen (Aarhus University) for the in-depth discussions regarding the potential applications of our work. Furthermore, we are in debt to the reviewers for the in-depth improvements made on this manuscript.

**Author Biographies**

**Cláudio Gomes** is a PhD student in the University of Antwerp and member of the Modelling, Simulation and Design Lab (MSDL), interested in co-simulation of hybrid systems.

**Bart Meyers** is a research engineer at Flanders Make vzw, and is associated with the University of Antwerp, where he obtained his PhD. His main research topic is modelling of software-intensive systems.

**Joachim Denil** is an Assistant Professor in the University of Antwerp and member of the Cosys-lab. His main research interest is multi-paradigm modelling of software-intensive systems.

**Casper Thule** is a PhD student at Aarhus University and member of the Software Engineering Research Group. His area of research is hybrid co-simulation and cyber-physical systems.

**Kenneth Lausdahl** is a PostDoc at Aarhus University and member of the Software Engineering group, interested in tool automation and co-simulation of hybrid systems.

\*https://msdl.uantwerpen.be/git/ciaulo/HybridCosimulation
**Hans Vangheluwe** is a Professor at the University of Antwerp (Belgium), an Adjunct Professor at McGill University (Canada) and an Adjunct Professor at the National University of Defense Technology (NUDT) in Changsha, China. He heads the MSDL research lab. His research interest is the multi-paradigm modelling of complex, software-intensive, cyber-physical systems.

**Paul De Meulenaere** is a Professor at the University of Antwerp and member of the research group CoSys-Lab. His main research field is model-based development methods for embedded real-time systems.

### References


Wirth N. Extended backus-naur form (EBNF), 1996.


Broman D, Greenberg L, Lee EA et al. Requirements for Hybrid Cosimulation Standards. In 18th International Conference on Hybrid Systems: Computation and


