Co-Simulation of Hybrid Systems with SpaceEx and Uppaal
Bogomolov, Sergiy; Greitschus, Marius; Jensen, Peter Gjøl; Larsen, Kim Guldstrand; Mikuionis, Marius; Strump, Thomas; Tripakis, Stavros

Published in:
11th International Modelica Conference Proceedings

DOI (link to publication from Publisher):
10.3384/ecp15118159

Creative Commons License
CC BY-NC 4.0

Publication date:
2015

Document Version
Early version, also known as pre-print

Link to publication from Aalborg University

Citation for published version (APA):

General rights
Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

? Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
? You may not further distribute the material or use it for any profit-making activity or commercial gain
? You may freely distribute the URL identifying the publication in the public portal

Take down policy
If you believe that this document breaches copyright please contact us at vbn@aub.aau.dk providing details, and we will remove access to the work immediately and investigate your claim.

Downloaded from vbn.aau.dk on: oktober 25, 2018
Co-Simulation of Hybrid Systems with SpaceEx and Uppaal

Sergiy Bogomolov 1  Marius Greitschus 2  Peter G. Jensen 3  Kim G. Larsen 3  Marius Mikučionis 3  Thomas Strump 2  Stavros Tripakis 4

1 IST Austria, Austria  
2 University of Freiburg, Germany  
3 Aalborg University, Denmark  
4 Aalto University, Finland, and University of California, Berkeley, USA

Abstract
The Functional Mock-up Interface (FMI) is an industry standard which enables co-simulation of complex heterogeneous systems using multiple simulation engines. In this paper, we show how to use FMI in order to co-simulate hybrid systems modeled in the model checkers SpaceEx and Uppaal. We show how FMI components can be automatically generated from SpaceEx and Uppaal models. We also validate the co-simulation approach by comparing the simulations of a room heating benchmark in two cases: first, when a single model is simulated in SpaceEx; and second, when the model is split into two submodels, and co-simulated using SpaceEx and Uppaal. Finally, we perform a measurement experiment on a composite model to show a potential for statistical model checking using stochastic co-simulations.

Keywords: FMI, hybrid system, timed automaton

1 Introduction

Despite advances in model checking and other formal verification techniques, simulation remains the workhorse for system analysis. A plethora of simulation tools are available today, from academia as well as from industry. These tools support a large variety of modeling languages, targeted at different types of systems from various disciplines (e.g., mechanical, electrical, digital, continuous or discrete, or mixes thereof). Unfortunately, these tools can rarely interoperate. This is a problem because modern cyber-physical systems are highly complex and multidisciplinary, requiring specialized modeling languages and tools from several domains.

The Functional Mock-up Interface 1 (FMI) is a standard developed to address this problem. FMI defines an XML schema for describing simulation components and a C API that these components must implement. The components are called functional mock-up units, or FMUs. An FMU is typically generated automatically (exported) from some simulation tool, and corresponds to a (sub-)model designed in that tool. The submodels/FMUs are then imported into a host simulator. The host commands the simulation by calling the API methods of the FMUs, thus effectively achieving integration of the original simulation environments. FMI supports two integration modes: (a) model exchange, where the host simulator handles the numerical integration; and (b) co-simulation, where each FMU implements its own numerical integration mechanism (or any other internal mechanism to advance its state in time). Because each mode imposes its own requirements on FMUs (for instance, in model exchange, the FMUs must provide the host with information such as state derivatives, which are not necessary for co-simulation) the FMI APIs for the two modes are different.

In this paper, we use FMI in order to connect two state-of-the-art modeling and verification tools for cyber-physical systems: SpaceEx (Frehse et al., 2011) and Uppaal (Larsen et al., 1997). SpaceEx is a tool for modeling and verifying hybrid systems (Alur et al., 1995). Uppaal is primarily a model-checker for timed automata (Alur and Dill, 1994), however, it also supports statistical model-checking of hybrid systems (David et al., 2011).

Our goal is to integrate these two tools for co-simulation. That is, we want to be able to: (a) build a sub-model of the system (e.g., the model of the plant under control) in SpaceEx; (b) build another sub-model (e.g., the controller) in Uppaal; (c) automatically generate an FMU for each sub-model; (d) import the FMUs, connect and co-simulate them in a host environment.

The motivations for connecting SpaceEx and Uppaal in this manner are numerous. First, although both SpaceEx and Uppaal support simulation of hybrid systems, each tool offers its own modeling language, which is not compatible with that of the other tool. Translating from one language to the other is limited to common features supported by the tools. For example, even though the frameworks CIF (Agut et al.,
to show the potential for statistical model checking using FMI co-simulations.

The rest of the paper is organized as follows. In Sec. 2, we introduce the necessary background on FMI for this work. Afterwards, we present our translation of SPACE\textsc{ex} and UPPAAL models into FMUs in Sec. 3. This is followed by the case study in Sec. 4. We discuss related work in Sec. 5. Finally, we conclude the paper in Sec. 6.

2 Background on FMI

Conceptually an FMU can be seen as a (timed) state machine. This machine has a set of input variables (or ports), a set of output variables, and a set of internal states. The machine interacts with its environment only by means of a clearly defined set of interface methods. These methods are specified in the FMI standard. For the purposes of this paper, and following the formalization presented by Broman et al. (2013), the key interface methods of FMI (for co-simulation) are:

- A method to initialize the state of the FMU. If \( S \) is the set of states of the FMU, then \( \text{init} \in S \).
- A method \( \text{set} \) to set a given input variable to a certain value. The signature of \( \text{set} \) is \( \text{set} : S \times U \times V \to S \), where \( U \) is the set of input variables of the FMU, and \( V \) is the set of all possible values (for simplicity we ignore typing and use a single universe \( V \) of values for all variables). Given state \( s \), input variable \( u \in U \), and value \( v \in V \), \( \text{set}(s,u,v) \) returns the new state obtained after setting \( u \) to \( v \).
- A method \( \text{get} \) which returns the value of a given output variable. Its signature is \( \text{get} : S \times Y \to V \), where \( Y \) is the set of output variables of the FMU. Given state \( s \) and output variable \( y \in Y \), \( \text{get}(s,y) \) returns the value of \( y \) in \( s \).
- A method \( \text{doStep} \) which advances the state of the machine in time. Its signature is \( \text{doStep} : S \times \mathbb{R}_{\geq 0} \to S \times \mathbb{R}_{\geq 0} \), where \( \mathbb{R}_{\geq 0} \) is the set of non-negative real numbers. The behavior of \( \text{doStep} \) is explained below.

As said above, an FMU is essentially a state machine: the \( \text{get} \) method corresponds to the output function of the machine, while the \( \text{doStep} \) method corresponds to the transition function. The difference is that \( \text{doStep} \) takes as input a time \( h \in \mathbb{R}_{\geq 0} \); in that sense, an FMU is a timed state machine.

The behavior of \( \text{doStep} \) is as follows. Given state \( s \in S \), and time step \( h \in \mathbb{R}_{\geq 0} \), a call to \( \text{doStep}(s,h) \) is interpreted as the co-simulation algorithm “asking” the FMU to perform a simulation step of length \( h \). For a
number of reasons, including numerical integration issues, the FMU may “accept” or “reject” this request. If it rejects, it means that it was not able to advance time by \( h \) (but may have been able to advance time by a smaller delay \( h' < h \)). Formally, doStep\((s,h)\) returns a pair \((s',h')\) where \( s' \in S \) is a state and \( h' \in \mathbb{R}_{\geq 0} \) is a time step, such that:

- either \( h' = h \), which is interpreted as \( F \) having accepted \( h \), and having moved to a new state \( s' \);
- or \( 0 \leq h' < h \), which is interpreted as \( F \) having rejected \( h \), but having made partial progress up to \( h' \), and having reached a new state \( s' \).

It is worth noting that FMUs are deterministic machines, in the sense that for a given sequence of inputs (i.e., a sequence of input values and time steps), the sequence of states and outputs that the machine produces is unique. This is because there is a unique initial state \( \text{init} \in S \), and \( \text{set}, \text{get}, \text{doStep} \) are all \textit{total functions}. Moreover, the fact that these functions are total implies that the machine is able to accept any input at any time, therefore, it is implicitly \textit{input-enabled}.

We also rely on zero-time steps in a sense of allowing doStep\((s,h)\) calls with \( h = 0 \) (despite that version 2.0 of the FMI standard forbids this), because they are essential for modeling discrete transitions like instantaneous mode switches in hybrid automata models.

In addition to the above, each FMU comes with a set of \textit{input-output dependencies}, \( D \subseteq U \times Y \). \( D \) specifies for each output variable which input variables it depends upon (if any): \( (u, y) \in D \) means that output variable \( y \) depends on input variable \( u \). This information is used to ensure that a network of FMUs has no cyclic dependencies, and also to determine the order in which all network values are computed during a simulation step (Broman et al., 2013).

FMI specifies the methods that every FMU must implement, but it does not specify the co-simulation algorithm (also called a master algorithm). In fact, devising such an algorithm with good properties is not a trivial problem, and has been the topic of previous work (Broman et al., 2013). In that work, two co-simulation algorithms were proposed and proved to have desirable properties, such as termination of a simulation step, and \textit{determinacy}. The determinacy property says that the results of a simulation do not depend on the order in which the algorithm chooses to call doStep over a set of FMUs. This ensures that the simulation results are well-defined and are not influenced by arbitrary factors such as FMU names, order of creation, geometrical position in the diagram of a graphical model, etc., as is often the case with simulation tools.

In a nutshell, the co-simulation method proposed by Broman et al. (2013) relies on the following principle. First, the co-simulation algorithm chooses a default time step, \( h_{\text{max}} \), called the \textit{maximum step size}. Second, the algorithm saves the state of each FMU in the model (FMI specifies methods for an FMU to export and import its state, although these are optional). Assuming there are \( n \) FMUs, \( F_1, \ldots, F_n \), the algorithm maintains \( n \) states, \( s_1, \ldots, s_n \). Third, the algorithm calls \( F_i, \text{doStep}(s_i, h_{\text{max}}) \) on each FMU \( F_i \), and collects the returned time steps \( h'_1, \ldots, h'_n \). There are two cases: either all FMUs accepted the proposed time step, i.e., \( h'_i = h'_2 = \ldots = h'_n = h_{\text{max}} \), in which case this simulation step is over, and the algorithm proceeds to the next one; or at least one FMU \( F_i \) rejected \( h_{\text{max}} \), i.e., \( h'_i < h_{\text{max}} \) for some \( i \). In the latter case, the algorithm computes the minimum of \( h'_1, \ldots, h'_n \), \( h_{\text{min}} = \min\{h'_1, \ldots, h'_n\} \), restores the saved state of each FMU, and tries again with new step size \( h_{\text{min}} \).

Assuming that the FMUs satisfy the reasonable “monotonicity” property that if they were able to advance time by \( h'_i \) then they are also able to advance time by any smaller step, and by the fact that \( h_{\text{min}} \) is smaller than all \( h'_i \), the second attempt is guaranteed to succeed. That is, \( h_{\text{min}} \) will be accepted by all FMUs. As a result, at most after two attempts, a co-simulation step is successful, and the algorithm proceeds with the next step, repeating the same procedure as above.

The FMI standard sets out a framework where FMUs share the notion of time and exchange variable values via input-output ports: outputs from one FMU are mapped as inputs to other FMU(s) and so on. The output port values are said to be owned and controlled by the emitting FMU, whereas the inputs are computed and provided by another (outputting) FMU. The framework foresees that before producing an output an FMU may first need some input values and thus input-output dependency information is introduced. Overall the I/O port connectivity graph derived from the model of interconnected FMUs, together with the local I/O dependencies of each individual FMU, result in a global I/O dependency graph for the entire model (Broman et al., 2013).

Time and I/O values are synchronized by the co-simulation algorithm: the time is agreed by repeatedly consulting each FMU and the I/O values are propagated according to dependencies. The co-simulation algorithm assumes that each FMU provides a static dependency list of its ports before simulation starts, and that the resulting global I/O dependency graph is acyclic, and therefore there exists a schedule for computing the value of every input port before the value of a dependent output port is requested (Broman et al., 2013).

3 Translating Models into FMUs

The behavior of individual FMUs is provided by the model-checker’s simulation engines based on the guidelines described by Tripakis (2015). In particular, the report distinguishes continuous and discrete dynamics. The continuous behavior is modeled by differential equations over continuous variables whose values can be
shared among FMUs by the means of port connections. The output ports of an FMU are mapped to the owned/controlled variables which are read and written to, whereas input ports map to read-only variables within the FMU.

The discrete behavior is modeled by discrete transitions in the timed/hybrid automata control flow structure. The discrete transitions are designed to be executed with micro-steps of zero delay. Transitions can also be decorated with event labels and each tool supports its own kind(s) of synchronizing compositions internally and therefore the discrete transition synchronization is also handled individually within the tools. Tripakis (2015) provides the means of discrete transition synchronization by allocating two special port variables: one for incoming (input) synchronization and one for outgoing (output) synchronization. The domain of discrete input (output) ports coincides with the set of input (output) labels plus a special value absent which denotes no synchronization or an internal discrete transition.

3.1 UPPAAL

UPPAAL uses timed automata models (Alur and Dill, 1994), extended with discrete variables over structured types to describe behaviors of a timed system. In timed automata, the continuous dynamics is controlled by real-valued clock variables (with derivatives set to one) and discrete states complemented with integer variables – both of which are candidates for exchange via FMU input-output ports. Statistical model checking (SMC) extensions (David et al., 2011, 2015) allow a finer control of the clock derivatives by means of ordinary differential equations, moreover the discrete transitions are stochastic where the execution is determined by probability distributions over time and over branching edges. The stochastic semantics of a parallel composition is similar to the FMI co-simulation algorithm (Broman et al., 2013): the way the minimum delay is negotiated and thus the timed composition within the FMI framework is straightforward, and task is to find a systematic way of handling discrete synchronizations. UPPAAL also supports the maximal progress or ASAP semantics on edges labeled with urgent channels.

UPPAAL supports the notion of discrete I/O synchronization natively by means of input and output channel labels. Thus, its discrete input and output transitions can be mapped directly to the input/output port variables of an FMU that is dedicated to transfer the synchronization label name. Nonetheless, we distinguish the following kinds of transitions: internal (transitions without I/O channel synchronization or internally synchronized transitions for which channels are not marked as input or output), input transitions (labeled by an input synchronization where the channel name is marked as an FMU input), and output transitions (labeled by an output synchronization where channel is marked as an FMU output). The marked outputs are controlled by the UPPAAL simulation and are executed asynchronously irrespective of whether the receiving FMU is ready to synchronize. Meanwhile, the input transitions are executed only when there is a corresponding input label set on a discrete input port. At most one (internal, input or output) transition is allowed at a time, hence fine-grained simulation control can be achieved by the co-simulation algorithm.

UPPAAL FMUs do not introduce I/O dependencies between continuous variables because the models do not use algebraic expressions to compute variable values. Instead of algebraic expressions the automata use discrete transitions to update the variable values. However, only one discrete transition is allowed at a time, therefore all discrete outputs have dependencies on the inputs dedicated to synchronization labels which restrict the selection of a particular discrete transition and hence specific variable update.

3.2 SPACEEx

SPACEEx (Frehse et al., 2011) uses hybrid automata to describe system behavior where the continuous variable derivatives are constrained by differential equations. The continuous variables are candidates for input and output exchange via FMU ports. The discrete transitions of hybrid automata can be decorated with labels. Synchronization may involve multiple participating processes, but there is no notion of input and output – all processes are equal contributors, therefore the simulator needs to implement the input/output semantics required by FMI. We use a special label naming notation to mark input and output labels (see Fig. 6). The transitions with input labels are only executed when the discrete input variable of FMU is set to the corresponding label name. Meanwhile, the transitions with an output label are controlled by SPACEEx’ simulation, and are executed asynchronously by setting the discrete output variable with the label name irrespectively of whether the receiving FMU can synchronize with it. We ensure the SPACEEx FMU determinism by enforcing the must-semantics of discrete transitions in a hybrid automaton. In other words, a discrete transition is taken as soon as its guard is enabled. Finally, we resolve the non-determinism between input, output, and internal transitions in the following way: input transitions have priority over output transitions and output transitions are preferred over the internal ones.

Both UPPAAL and SPACEEx translations simulate the source models as they are without intermediate transformations, except of the following additions: 1) input enabledness is ensured by broadcast channels in UPPAAL modeling and asynchronous I/O is implemented for SPACEEx synchronization labels, 2) for determinization SPACEEx uses maximal progress whereas UPPAAL uses stochastic semantics with a possibility of urgent channels for maximal progress.
In this section, we discuss the co-simulation semantics and contrast it to those typically used by a model-checking tool. In particular, we demonstrate by example how the FMI co-simulation algorithm resolves input/output dependencies and contrast it with execution analysed in a model checker. Our goal is to offer insights in the differences of the two semantics.

Consider a system model shown in Fig. 1 which consists of four timed automata composed in parallel. Labels of the form \( a! \) denote sending output \( a \), whereas \( a? \) denotes receiving an input \( a \). The variable \( x \) is a clock measuring time starting from zero. The constraint \( x = 1 \) is a guard which allows the corresponding transition of the automaton to occur only if the guard is satisfied, i.e., in this case only when \( x \) equals 1. The automata synchronize in a chain: the first can output \( a \) to the second one, the second one can output \( b \) to the third one and so on.

In principle, the system can be loaded into an FMI model in any combination: individually (one automaton per FMU) or collectively (multiple automata per FMU), but before an FMU can be loaded into an FMI model, it must declare its input/output dependencies. According to Broman et al. (2013) each automaton should expose an input/output variable which will contain the synchronization label value. Automaton \( A_1 \) in the example above will have only an output variable, which may contain values \{\( a, \text{absent} \}\). Automaton \( A_2 \) will have an input variable ranging over \{\( a, \text{absent} \}\) and an output variable ranging over \{\( b, \text{absent} \}\), and so on. The special value \( \text{absent} \) denotes that currently there is no synchronization. Timed automata must declare a dependency between its input and output label variable in order to avoid simultaneous input and output synchronizations.

In addition, it is assumed that each FMU is input-enabled, meaning that it can handle (i.e., it is able to receive) any declared input at any time. If a component is not input-enabled and an input synchronization is triggered then simulation is aborted, to avoid such situation we allow only broadcast channels, which do not block the sender process and receiver may simply ignore the synchronization if has no receiving edge.

Suppose the automata from Fig. 1 are loaded within separate FMUs and connected according to synchronization labels. That is, the output of \( FMU(A_1) \) is connected to the input of \( FMU(A_2) \), the output of \( FMU(A_2) \) is connected to the input of \( FMU(A_3) \), and so on. The co-simulation algorithm would detect that it has to fulfill inputs values for the \( FMU(A_4), FMU(A_3), \) and \( FMU(A_2) \) in order to proceed, therefore the input/output value propagation will have to start with \( FMU(A_1) \) and then proceed to the \( FMU(A_2) \) etc.. Once the values of all input and output variables are propagated, the algorithm proceeds with advancing each FMU in time by calling \( doStep() \). It is this dynamic behavior in time which interests us in this example.

In particular, observe that \( A_{2,3,4} \) automata are non-deterministic in the sense that, according to UPPAAL semantics, at time \( x = 1 \) an automaton can either delay, or take an outputting transition, or synchronize on inputs. For instance, at time \( x = 1 \), \( A_2 \) can either emit \( b \), or receive \( a \) (which will be available in this case, because it is sent by \( A_1 \) at exactly that time), or let the time pass. In timed automata semantics, all these options are possible at the individual component level. Moreover, not only individual components can be non-deterministic, but their composition is non-deterministic as well, based on so-called interleaving semantics. This means that when multiple automata are enabled at a given time, the choice of which one to execute is arbitrary. Non-determinism is a useful abstraction and thus model reduction technique in verification and model-checking. The same is true when these tools are used for simulation, i.e. different simulations in UPPAAL may yield different results.

In FMI, the situation is very different, as all FMUs are treated as deterministic components, and their composition, ensured by the co-simulation algorithm, is guaranteed to yield deterministic results as well. Interestingly, in this example, if all automata decide to output at time \( x = 1 \), some of them will succeed outputting in parallel, while others will be preempted by incoming inputs. In particular, the master algorithm will request \( FMU(A_1) \) to produce its output, and thus \( FMU(A_2) \) will be busy handling an input and will not be producing output at that time. Since \( FMU(A_2) \) is not sending anything, then \( FMU(A_3) \) will be free to produce an output and hence preempt \( FMU(A_4) \).

As witnessed from above, such FMI system selects a particular sequence of steps (which is expected) but is not able to simulate all possible execution orders as in original semantics even if we allow FMUs to determinize their actions by themselves, which means that FMI simulations are selecting a particular subset of all possible behaviors and some behaviors may not be reproducible in FMI. Also FMI simulations may contain parallel synchronizations (e.g. actions \( A_1 \xrightarrow{a} A_2 \) and \( A_3 \xrightarrow{c} A_4 \) at the same computation step) which are possible only in several steps in timed automata semantics (action \( a \) and only then action \( c \) within zero-time), hence the intermediate state between \( a \) and \( c \) actions might not be accessible in FMI without very fine grained control over individual \( doStep() \) calls in one zero-time computation step. However, the successor state of such parallel executions can be matched with a state after multiple transitions in the given automata semantics, hence the FMI simulation

3.3 Discussion on Co-Simulation Semantics

Consider a system model shown in Fig. 1 which consists of four timed automata composed in parallel. Labels of the form \( a! \) denote sending output \( a \), whereas \( a? \) denotes receiving an input \( a \). The variable \( x \) is a clock measuring time starting from zero. The constraint \( x = 1 \) is a guard which allows the corresponding transition of the automaton to occur only if the guard is satisfied, i.e., in this case only when \( x \) equals 1. The automata synchronize in a chain: the first can output \( a \) to the second one, the second one can output \( b \) to the third one and so on.

In principle, the system can be loaded into an FMI model in any combination: individually (one automaton per FMU) or collectively (multiple automata per FMU), but before an FMU can be loaded into an FMI model, it must declare its input/output dependencies. According to Broman et al. (2013) each automaton should expose an input/output variable which will contain the synchronization label value. Automaton \( A_1 \) in the example above will have only an output variable, which may contain values \{\( a, \text{absent} \}\). Automaton \( A_2 \) will have an input variable ranging over \{\( a, \text{absent} \}\) and an output variable ranging over \{\( b, \text{absent} \}\), and so on. The special value \( \text{absent} \) denotes that currently there is no synchronization. Timed automata must declare a dependency between its input and output label variable in order to avoid simultaneous input and output synchronizations.

In addition, it is assumed that each FMU is input-enabled, meaning that it can handle (i.e., it is able to receive) any declared input at any time. If a component is not input-enabled and an input synchronization is triggered then simulation is aborted, to avoid such situation we allow only broadcast channels, which do not block the sender process and receiver may simply ignore the synchronization if has no receiving edge.

Suppose the automata from Fig. 1 are loaded within separate FMUs and connected according to synchronization labels. That is, the output of \( FMU(A_1) \) is connected to the input of \( FMU(A_2) \), the output of \( FMU(A_2) \) is connected to the input of \( FMU(A_3) \), and so on. The co-simulation algorithm would detect that it has to fulfill inputs values for the \( FMU(A_4), FMU(A_3), \) and \( FMU(A_2) \) in order to proceed, therefore the input/output value propagation will have to start with \( FMU(A_1) \) and then proceed to the \( FMU(A_2) \) etc.. Once the values of all input and output variables are propagated, the algorithm proceeds with advancing each FMU in time by calling \( doStep() \). It is this dynamic behavior in time which interests us in this example.

In particular, observe that \( A_{2,3,4} \) automata are non-deterministic in the sense that, according to UPPAAL semantics, at time \( x = 1 \) an automaton can either delay, or take an outputting transition, or synchronize on inputs. For instance, at time \( x = 1 \), \( A_2 \) can either emit \( b \), or receive \( a \) (which will be available in this case, because it is sent by \( A_1 \) at exactly that time), or let the time pass. In timed automata semantics, all these options are possible at the individual component level. Moreover, not only individual components can be non-deterministic, but their composition is non-deterministic as well, based on so-called interleaving semantics. This means that when multiple automata are enabled at a given time, the choice of which one to execute is arbitrary. Non-determinism is a useful abstraction and thus model reduction technique in verification and model-checking. The same is true when these tools are used for simulation, i.e. different simulations in UPPAAL may yield different results.

In FMI, the situation is very different, as all FMUs are treated as deterministic components, and their composition, ensured by the co-simulation algorithm, is guaranteed to yield deterministic results as well. Interestingly, in this example, if all automata decide to output at time \( x = 1 \), some of them will succeed outputting in parallel, while others will be preempted by incoming inputs. In particular, the master algorithm will request \( FMU(A_1) \) to produce its output, and thus \( FMU(A_2) \) will be busy handling an input and will not be producing output at that time. Since \( FMU(A_2) \) is not sending anything, then \( FMU(A_3) \) will be free to produce an output and hence preempt \( FMU(A_4) \).

As witnessed from above, such FMI system selects a particular sequence of steps (which is expected) but is not able to simulate all possible execution orders as in original semantics even if we allow FMUs to determinize their actions by themselves, which means that FMI simulations are selecting a particular subset of all possible behaviors and some behaviors may not be reproducible in FMI. Also FMI simulations may contain parallel synchronizations (e.g. actions \( A_1 \xrightarrow{a} A_2 \) and \( A_3 \xrightarrow{c} A_4 \) at the same computation step) which are possible only in several steps in timed automata semantics (action \( a \) and only then action \( c \) within zero-time), hence the intermediate state between \( a \) and \( c \) actions might not be accessible in FMI without very fine grained control over individual \( doStep() \) calls in one zero-time computation step. However, the successor state of such parallel executions can be matched with a state after multiple transitions in the given automata semantics, hence the FMI simulation
states in between system computation steps are included in the original semantics, albeit definite proof requires more formal insight to examine all scenarios.

4 Case Study

We have implemented the FMI standard in the UPPAAL (Larsen et al., 1997) and SPACEEx (Frehse et al., 2011) model checkers by providing model export to FMU\(^3\). In this section, we present and evaluate the performance of the resulting FMI framework on a case study inspired by the well-known room heating benchmark originally proposed by Fehnker and Ivancic (2004). Our model consists of a room with a heater (Fig. 2) and a controller (Fig. 3) which regulates the heater behavior. We model the room and the controller as a SPACEEx and UPPAAL FMU, respectively (see Fig. 4). Our bang-bang controller turns the heater on and off as soon as some temperature thresholds \(T_{\text{low}}\) and \(T_{\text{high}}\) have been reached. The as-soon-as-possible behavior is enforced by using urgent channels which effectively make the controller deterministic. The room temperature \(T\) evolves according to the following differential equation:

\[
\dot{T} = k \cdot (T_{\text{env}} - T) + h_{\text{power}} \\
T_{\text{env}} = 0 \\
h_{\text{power}} = 0
\]

In other words, the room temperature depends linearly on the difference between the current room temperature \(T\) and outside temperature \(T_{\text{env}}\). We assume the outside temperature \(T_{\text{env}}\) and heater power \(h_{\text{power}}\) to be constant. The constant \(k\) defines the heat exchange rate between the room and outside environment. If the heater is off, the heater power is set to zero.

4.1 Evaluation

We evaluate our FMU framework by comparing simulation trajectories of the FMUs with ones produced by a SPACEEx model consisting of both the controller and room components. We consider three different simulation step values: 1 (see Fig. 5a), 0.1 (see Fig. 5b) and 0.01 (see Fig. 5c). Considering the simulations, we observe that the FMU trajectories overshoot the controller constraints in the sense that the controller exhibits a delayed reaction when the room temperature crosses the temperature thresholds. The behavior is justified by the fact that the method call doStep for every FMU relies only on the local information about the state evolution when making decisions, e.g., the controller FMU does not have any information about the room temperature evolution beyond the value which can be provided when

\[ T = k \cdot (T_{\text{env}} - T) \]

\[ T_{\text{env}} = 0 \]

\[ h_{\text{power}} = 0 \]

the method doStep is called. Therefore, the controller FMU detects that the guard is enabled only a simulation iteration later after this event has already happened. We observe that the impact of the overshooting can be made arbitrary small by choosing a small enough simulation step (see Fig. 5c vs. Fig. 5a and Fig. 5b).

We note that the overshooting problem is inherent to the considered master algorithm and can be circumvented by incorporating additional cross-component knowledge into the master algorithm. Overall, our experiments validate that on this case study our co-simulation framework based on SPACEEx and UPPAAL provides equivalent simulation results compared to the setting where all components are modelled in one tool.

4.2 Supervisory Control Example

In this section, we show how supervisory control systems similar to the benchmarks presented by Fehnker and Ivancic (2004) can be modeled using the FMI paradigm. Compared to Section 4.1, we consider a model of the building with two rooms sharing a common wall and a heater. In this setting, the room temperature is influ-

\[
T = k \cdot (T_{\text{env}} - T) + h_{\text{power}} \\
T_{\text{env}} = 0 \\
h_{\text{power}} = 0
\]

\[
T \leq T_{\text{low}} \\
\text{(on)} \\
h_{\text{on}}! \\
T \geq T_{\text{high}} \\
\text{(off)} \\
h_{\text{off}}!
\]

Figure 2. Room component modelled in SPACEEx. The component switches between “on” and “off” modes. The temperature variable \(T\) is exported as output and synchronizations labels \(h_{\text{on}}\) and \(h_{\text{off}}\) as inputs.

\[
T \geq T_{\text{high}} \\
\text{(off)} \\
h_{\text{off}}!
\]

Figure 3. Controller in UPPAAL uses urgent channels to ensure as-soon-as-possible transition trigger. Temperature \(T\) is an input and labels \(h_{\text{on}}\) and \(h_{\text{off}}\) are outputs.

\[
T \leq T_{\text{low}} \\
\text{(on)} \\
h_{\text{on}}!
\]

Figure 4. SPACEEx and UPPAAL FMUs connected using the room temperature \(T\) and heater mode \(h_{\text{mode}}\).
enced by both the outside temperature and heat transfer between the rooms. Figure 6 shows a hybrid automaton from SPACEEx modeling the room temperature dynamics. The difference from the previous example here is an extra term \((T_{other} - t) \times 0.2\) denoting a contribution from another room. Another room is modeled analogously except that it responds to \(heater2\_on\) and \(heater2\_off\) signals instead of \(heater1\_on\) and \(heater1\_off\).

Our controller consists of two parts: local bang-bang controller and a supervisor shown in Fig. 7. In order to model the transitions of the heaters between the rooms, we assume that the controllers can be turned on/off by the supervising controller. Therefore, the local controller has an extra mode besides \(On\) and \(Off\) which stands for the controller being currently deactivated. The supervising controller has two kinds of stochastic behavior: it can pick any pair of rooms (one recipient and another donor) to transfer the heat, and it can choose the timing of transfer. When a pair of rooms is selected (by choosing concrete room identifiers for \(rec\) and \(donor\) variables) the donor is disabled by moving from location \(move\) to location \(move\) and the recipient is enabled by going from \(move\) to \(idle\). The supervisor may stay in location \(idle\) arbitrary long, but the exact duration is decided by an exponential probability distribution of rate 1 which means the duration of 1/1 time units on average. Similarly the supervisor may stay in \(decide\) and \(move\) but the duration will be 1/10000 on average, i.e. denoting that the heater is moved rather quickly.

Figure 8 shows the overall component connectivity diagram where the supervisor is reading temperatures from each room and controls the local movable heater controllers. The movable heaters then may either turn on the heat in their room or let them cool off giving the heat to

**Figure 5.** Simulation trajectories: each red x is a data point reported by SPACEEx, and blue + reported by the co-simulation.

(a) Maximum step size 1.

(b) Maximum step size 0.1.

(c) Maximum step size 0.01.

**Figure 6.** Hybrid automaton for a heated room connected to another room. Inputs are temperatures \(T_{env}, T_{other}\) and labels \(IN\_heater1\_on\) and \(IN\_heater1\_off\), while output is temperature \(t\). We use the prefix \(IN\) to mark input labels.

(a) Local bang-bang controller which can be moved (disabled). The inscribed \(U\) means urgent location where time delay is not allowed. The inputs are temperature variable \(T[i]\) and labels \(enable[i]\) and \(disable[i]\) where \(i\) is the room index.

(b) Supervising controller moves the heaters between rooms by reading inputs on \(T[i]\) and sending outputs on labels \(enable[i]\) and \(disable[i]\) where \(i\) is the room index.

**Figure 7.** Two layers of UPPAAL controllers.
outside. The individual heated rooms are then connected to the outside temperature and to each other denoting the heat exchange. The splitter FMUs are repeaters needed to connect multiple components to the same signal.

In the following, we discuss the behavior of the resulting composed model. Figure 9 shows the temperature dynamics in each room. In particular, the plot shows that in the beginning the temperature drops until the supervisor detects a room temperature below $T_{get} = 17^\circ$, then around 6 time units a heater raises the temperature in room 1. The local controller keeps raising the temperature until it goes over 22$^\circ$ bound at around 7.5 time units. Notice that the temperature in room 2 also rises due to heat exchange between the rooms. Around 10 time units the supervisor decides to hand over the heater to room 2. At 14 time units the heater is switched back to room 1 and so on. We can conclude that even though the temperature drops well below 18$^\circ$ overall it seems that the controllers manage to sustain the temperature at the similar level without loosing control (without dropping to outside temperature level).

### 4.3 Stochastic Simulations and SMC

The following is a demonstration of statistical model checking (SMC) using the FMI framework. We show how the performance of two stochastic controllers simulated by UP2AAL can be compared using SMC approach together with the heated room simulation provided by SPACEEX. Figure 10 shows two controllers: (a) reacting within 1 time unit to 18.0$^\circ$ and 22.0$^\circ$ temperature bounds and (b) reacting within 2 time units to 19.0$^\circ$ and 21.0$^\circ$ temperature bounds. The channels used in these controllers are not urgent and therefore the delay between temperature detection and heater activation is decided stochastically based on uniform distribution over the allowed delay by invariants, i.e. the concrete delay will be chosen from [0, 1] for the first controller and from [0, 2] for the second one. The On and Off locations do not have any invariant and therefore in principle the process may stay there forever. In such cases UP2AAL uses an exponential (Poisson) probability distribution to decide a particular time delay and hence asks to provide a rate of the exponential. The higher the exponential rate, the shorter the delays, hence we can provide a high rate to ensure that the detecting transition is fired arbitrarily quickly.

In our setup, we would like to know which controller is better at keeping the room temperature within 18.0$^\circ$ and 22.0$^\circ$ bounds. In order to answer this question we set up two FMI models for each controller with an equal room, run 100 simulations with 100 time units in length and 0.05 granularity, compute the amount of time spent outside the temperature range for each simulation and then compute the confidence intervals for both models. Table 1 shows a summary of amounts of time during which the temperature was either below or above the range. The estimated time duration use confidence interval (CI) notation which means that if we repeat the measurement experiment then the real mean (which is unknown) will fall into the interval with a probability of 95%. The results show that the second controller was more successful at maintaining the lower bound of the temperature, but was more overshooting beyond the upper bound. In total, the first controller kept the temperature in good range longer by 8.57 time units on average, which is much larger than confidence interval, hence the first controller is better.

<table>
<thead>
<tr>
<th>Controller</th>
<th>Time below</th>
<th>Time above</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Wide and fast</td>
<td>7.56 ± 0.20</td>
<td>32.69 ± 3.36</td>
<td>40.26 ± 0.59</td>
</tr>
<tr>
<td>Narrow and slow</td>
<td>2.40 ± 0.19</td>
<td>46.43 ± 0.82</td>
<td>48.83 ± 0.79</td>
</tr>
</tbody>
</table>

### 5 Related Work

The FMI standard and corresponding documentation are constantly evolving, as new versions of the standard are developed. The web site\footnote{https://www.fmi-standard.org/} also contains a list of tools supporting FMI. Descriptions of FMI can also be found in the academic literature (Blochwitz et al., 2011).

Discussions about the limitations of FMI can be found in the works by Broman et al. (2013, 2015). Broman et al. (2013) also formalize the main methods of FMI (get, set, doStep) by establishing a contract (pre/post-conditions) for each method and propose a master algorithm (i.e., a co-simulation algorithm). Furthermore, the authors prove its termination, determinacy, and other properties. However, the paper does not discuss how FMUs can be created. A different, master-slave based, co-simulation approach is proposed by Bastian et al. (2011), but formal properties such as determinacy are not discussed in this work.

Broman et al. (2015) defines a suite of test models that should be supported by a hybrid co-simulation environment, giving a mathematical model of an ideal behavior, plus a discussion of practical implementation considerations. Furthermore, the paper describes a set of basic modeling components in the spirit of Ptolemy actors (constant, gain, adder, integrator, etc.). Finally, the authors provide a kind of denotational description for each component (input and output signals), but no encoding into FMUs is discussed.

The FMU generation problem for various formalisms is discussed by Tripakis (2015). This work only refers to a generic model of timed machines which does not include the particularities of UP2AAL’s timed automata. In addition, hybrid automata are not considered in this work.

Recently, the co-simulation algorithm presented by Broman et al. (2013) has been implemented in the
open-source Ptolemy tool. As mentioned above, we use this framework in order to import FMUs into Ptolemy and co-simulate them. However, this framework does not address the FMU generation problem.

Feldman et al. (2014) present a plugin for Rhapsody for generating FMUs from Statechart SysML blocks. They provide high level guidelines for how to generate Statechart FMUs, but do not provide a formalization. Pohlmann et al. (2012) also discuss how to encode statechart models, described in MechatronicUML.

Co-simulation is one, but not the only approach to solve the tool interoperability problem. A further attempt to solve this problem is the Hybrid Systems Interchange Format (HSIF), designed with the goal of being “a sort of ‘maximum common denominator’ among all hybrid system modeling environments” (Pinto et al., 2005). HSIF aims therefore at defining a “maximal syntax” where all the syntax of different languages could be translated into. It could be seen as a type of XML schema for hybrid systems. HSIF is primarily aimed at enabling model translation between different hybrid system tools. Bak et al. (2015) present the tool Hyst which provides an automatic source-to-source model translation between a number of up-to-date hybrid model checkers. In their approach, Bak et al. (2015) do not use any intermediate format like HSIF. Both model translation-based approaches outlined above provide support only of the common subset of the tool features. Our co-simulation framework does not limit the model designer to use the “maximal syntax” among all the tools because every FMU takes care of its features independently.

6 Conclusions

We have shown how two state-of-the-art modeling and verification tools for hybrid systems, SPACEX and UPPAAL, can be integrated using the FMI co-simulation standard. The result is a powerful framework which allows users to build submodels separately in each of the two tools (as well as in other tools potentially), then generate individual FMUs for each submodel, and then combine all the FMUs into a single model, which can be co-simulated within the Ptolemy FMI implementation. We demonstrated the feasibility of our framework by comparing the co-simulation results of a simple two-FMU model to the simulation results that are obtained when we model and simulate the entire system in a single tool. By empirical evaluation on case studies we found that, provided time steps are small enough, individual components can ensure timely reactions to continuous signals, and the simulations can be made arbitrary close to self-contained model simulation. In addition to individual tool export to FMUs, we showed

Figure 8. Ptolemy diagram for supervisory control of two heated rooms.

Figure 9. Temperature trajectories for each of the rooms composed with stochastic supervising controller.

Figure 10. Stochastic controllers use regular (non-urgent) channels, therefore timings are stochastic: delays are distributed uniformly (when clock invariant is used) and exponentially (in locations On and Off).
how the non-deterministic models can be determined using stochastic semantics and included into FMI co-simulation. We also provided an example how statistical model checking can be performed using numerous FMI simulations which is an essential feature evaluating stochastic behavior. The integration of model-checkers into co-simulation frameworks provides further possibilities of analyzing early design models like conformance monitoring by checking that a simulation trace of a refined (e.g., hybrid) model is included in a more a abstract (e.g., timed automata) specification. We envision our work being a further step towards integrating tools developed in the formal methods community into the industrial system design and modeling workflow of cyber-physical systems.

7 Acknowledgments

We are grateful to Christopher Brooks, Fabio Cremona, and Edward Lee from UC Berkeley, for their work on the Ptolemy framework. This work was partly supported by the European Research Council (ERC) under grant 267989 (QUAREM), by the Austrian Science Fund (FWF) under grants S11402-N23 (RISE) and Z211-N23 (Wittgenstein Award), by the German Research Foundation (DFG) as part of the Transregional Collaborative Research Center “Automatic Verification and Analysis of Complex Systems” (SFB/TR 14 AVACS), by EU FET project SENSATION, the Sino-Danish Center IDEA4CPS and the Center DiCyPS of the Danish Innovation Foundation, by Academy of Finland, by the National Science Foundation (awards #1329759 and #1139138), and by the Industrial Cyber-Physical Systems Research Center (iCyPhy) supported by IBM and United Technologies Corporations.

References


Jens Bastian, Christoph Clauß, Susann Wolf, and Peter Schneider. Master for Co-Simulation Using FMI. In 8th International Modelica Conference, 2011.


