Combination of Behavioral and Parametric Diagrams for Model-based Testing
Technical Report
December 2015

Christoph Hilken
University of Bremen
Bibliothekstr. 1
28359 Bremen
Germany
chilken@cs.uni-bremen.de

Felix Hübner
University of Bremen
Bibliothekstr. 1
28359 Bremen
Germany
felixh@cs.uni-bremen.de

Jan Peleska∗
University of Bremen
Bibliothekstr. 1
28359 Bremen
Germany
jp@cs.uni-bremen.de

ABSTRACT
Model-based testing is a promising approach to cope with the complexity of nowadays systems. It shifts the effort from manual test case specification to the modeling of the system’s behavior and enables automated test case generation. System behavior is often modeled by state machines. However, state machines are not appropriate to model all kinds of behavior. This work combines state machines and parametric diagrams of the Systems Modeling Language (SysML) to allow a higher degree of abstraction. This is shown by application to a real-word example: a safety function of the European Train Control System (ETCS).

Keywords
SysML, Model-based Testing, ETCS, Cyber-physical Systems, Hybrid Systems

1. INTRODUCTION

Model-based testing.

Model-based testing (MBT) has gained much attention during the last decade [17, 14, 3]. This is mainly due to the fact that MBT enables a high degree of automation, increasing the efficiency of test-related verification and validation activities in a considerable way. The main automation benefits are mechanized test case creation from the model, test data calculation by means of mathematical constraint solvers, test procedure generation using model-based code generation techniques, and compilation of traceability data relating testing artifacts to requirements by exploiting traceability mechanisms available in the modeling languages [13]. At the same time, MBT allows for the application of more complex test strategies. These provide higher test strength, but the test case generation algorithms involved can no longer be managed in a manual way; examples of these more complex strategies are given in [8, 15, 10].

For automated MBT, the modeling formalism applied needs to be associated with a formalized behavioral semantics describing how model states, inputs, and outputs evolve over time. For test models described in the SysML formalism [12], formalization options are described, for example, in [13, 9]. With these results at hand, model-based testing against concurrent real-time SysML models can be considered as a solved problem for continuous time/discrete control systems depending on notions of discrete or dense time, but producing discrete control outputs only. The system behavior over time is modeled, for example, by means of concurrent SysML state machines, whose trigger conditions depend on variable values and timer conditions. This is then formally specified by a transition relation describing how discrete control steps or time-delays are performed. Using, for example, an SMT solver that is also capable of floating point arithmetics, the possible transition steps can be calculated. Specific test objectives can be encoded as additional constraints used in conjunction with the transition relation, so that the solutions provided by the constraint solver describe at the same time valid state transitions of the model and suitable candidates for the test objective under consideration.

The challenge.

For real-time systems depending on mixed time-discrete and time-continuous evolutions of observables and/or control variables, no comprehensive MBT methodology exists yet. While the formal semantics of these so-called hybrid systems has been thoroughly investigated [7, 2], the automated calculation of suitable test data for practical MBT still remains a challenge. This is mainly due to the fact that best practices for specifying time-continuous evolutions in test models and creating associated concrete data by means of constraint solving are still subject to discussions.

Objectives and main contributions.

∗The author’s research has been partially funded by ITEA2 project openETCS under grant agreement 11025
In this paper, a novel approach to MBT in a hybrid system context is presented, based on the SysML modeling language. The utilization of blocks and associated diagrams for decomposing the functionality of the system under test (SUT) and the use of state machines is "imported" from proven MBT technology for time-discrete systems. These description means are, however, combined with an abstraction technique and extended by constraint blocks and parametric diagrams [12, Section 10] for modeling time-continuous dependencies between inputs, outputs, and model variables. From these descriptions means our proposed approach is able to generate concrete test cases in a fully automatic way.

Our approach is illustrated and a proof of concept is given by application to a complex real-world system. We create a test model of a control problem from the European Train Control System (ETCS), using the system requirement specification [16]. We describe how the expected EVC behavior can be modeled using the SysML subset indicated above, and the computational effort needed for automated test generation by means of an SMT solver is evaluated.

2. BACKGROUND

In order to keep this paper self contained relevant parts of the European Train Control System and the Systems Modeling Language are introduced.

2.1 European Train Control System

The European Train Control System (ETCS) shall replace existing national train control systems by a modern unified system in the European Union. The system requirement specification is publicly available [16].

The onboard computer of ETCS trains (the so-called European Vital Computer (EVC)) performs — among other safety-relevant control tasks — target speed monitoring (TSM): the speed of the train, while approaching a target (for example, a train station or a level crossing), is monitored by the EVC, so that the ability to brake the train in time for the target is always ensured [16]. To this end, the speed-dependent braking curves of the train have to be taken into account, so that the braking distance required is correctly calculated.

In the remainder, the following variables are used: \( v_{\text{est}} \) is a system input describing the current train speed. The maximum allowed speed on the track section of the train is denoted by \( v_{\text{max}} \). The value of \( v_{\text{max}} \) should remain constant over time, while \( v_{\text{est}} \) should change according to the current acceleration \( a \). All track related locations, including the train position, are measured as track relative distances \( D \), starting from zero (start location) and ending at the target location \( D_{\text{Target}} \). The train position is given by a variable \( D_{\text{front}} \) and should always have values from the range \([0, D_{\text{Target}}]\). \( D_{\text{Target}} \) is constant during runtime, while \( D_{\text{front}} \) might change over time as the train moves forward towards the target location.

2.2 Systems modeling language

OMG’s System Modeling Language SysML [12] is a semiformal language to specify a model of the system’s structure and behavior. In addition, corresponding diagrams offer a graphical representation of parts of the model. Therefore, a system description is composed of different kinds of diagrams, for example block definition diagrams and internal block diagrams to describe the structure, complemented by activity and state machine diagrams to describe the behavior of the system. Our approach is working on state machines and constraint blocks. Therefore, to focus on the objectives of this paper, in the remainder only parametric diagrams and state machine diagrams are used.

### Constraint blocks.

Constraint blocks are used to express general dependencies between observables, such as physical laws. Parametric diagrams are used to bind the generic names of these observables to the concrete model variables that are restricted by these constraints. Using these syntactic entities, the boundary conditions restricting time-continuous inputs or outputs of the SUT can be suitably specified.

#### Example 1.

For the TSM function the braking distance has to be calculated. The motion of a train can be expressed by the equations for a translational acceleration, which are:

\[
\begin{align*}
\dot{v} &= a, \\
\dot{s} &= \frac{1}{2} a t^2 + v_0 t + s_0.
\end{align*}
\]

In Figure 1 both expressions are combined to calculate the braking distance \( s \). The constraints are bound to the system inputs \((v_{\text{est}}\) and \(a)\) in the parametric diagram.

SysML does not define a language to express constraints. In the remainder the natural mathematical notation is used, which is interpreted by our tools. In addition, it is possible to add conditions under which constraints are valid. Therefore, pseudocode is used. In particular, such conditioned constraints can be used to specify the so-called flow conditions of hybrid automata associated with control modes specified in state machines: following [7], these flow conditions describe the restrictions to be observed by time-continuous variables, while the system resides in the given control mode.

### State machines and testing strategies.

UML’s state machines have been adopted for SysML without changes. Hence, SysML state machines are a variant of Harel’s statechart formalism [6]. Therefore, many MBT-tools are using state machines as modeling formalism, e.g.[1, 4] and many testing strategies have been developed in respect to state machines. This includes conventional strategies based on coverage criteria, but also complete test strategies, that give a guarantee that certain types of faults are revealed, given an assumption that only pre-defined types of faults can occur. In addition, it was shown in [11] that test suites generated by a complete test strategy can have a very

![Figure 1: Parametric diagram for the calculation of the braking distance](image-url)
3. PROBLEM FORMULATION AND GENERAL IDEA

As mentioned before, state machines have been shown suitable to model the behavior of reactive systems in the past and MBT approaches chose them as description means for their input models. With respect to upcoming cyberphysical-systems state machines are not convenient, especially, due to the integration of components, such as sensors and actuators in today’s embedded systems. Such hybrid systems can be described by a combination of state machines and conditioned constraint blocks. Constraint blocks and parametric diagrams are well suited to describe mathematical expressions, as for example physical laws, due to their declarative manner. State machines on the opposite are an imperative, state-based description means and, hence, are not convenient to specify physical properties of the system. There are many approaches in the field of control theory and simulation to handle physical properties, but in the field of MBT physical constraints are usually neglected or simplified and current test strategies do not take them into account. As a consequence, test cases generated using current tools and strategies may be unrealistic and violate physical constraints.

Example 2. Assume test cases shall be generated for the TSM function. Without considering physical laws, a concrete test case might be generated that contains a sequence of train speed values that have sudden jumps from very low to very high values. Such a behavior is unrealistic, since the acceleration of a train is bounded by the train engine and train brakes. Figure 2 describes the physical constraints, that model how the speed and position of a train should evolve over time. Considering this kind of physical constraints leads to test cases that are executable in a real environment.

To overcome this situation and to allow the utilization of parametric diagrams as well as constraint blocks, we propose the following approach:

In a first step using an abstracted version of the state machine, abstract test cases are generated. Such an abstracted version is a so-called simulation of the concrete system [5]; the former represents an over-approximation of the latter. The simulation can easily be obtained by using Boolean variables as guard condition for every transition in the state machine. These Boolean variables are not restricted on the state machine level but bound to concrete system variables in the parametric diagrams. Abstract test cases are constructed from this simulation model using an input equivalence class partitioning test strategy with proven error detection capabilities [10]. The inputs associated with an abstract test case are sequences of input equivalence classes. In the second step, the abstract input sequences are resolved to sequences of concrete model variable valuations, using a mathematical constraint solver. For this step, the bindings of abstract Boolean condition variables to concrete model variables defined by parametric diagrams are taken into account. Additional physical constraints describing the temporal evolution of system variables are considered in this step as well, which leads to realistic concrete test cases.

4. TWO-STEP APPROACH

Step 1. System abstraction.

In the first step, the system behavior is abstracted. The idea is to use an over-approximation of the specified system’s behavior: all traces possible in the concrete system based on the specification, considering all physical and system related constraints, are also observable in the abstract model. Whether a trace in the abstract model has a concrete solution in the real world, depends on the abstraction and constraints. This is a conservative approach, which guarantees that every system behavior is covered by the model, while unrealistic traces might remain. If in the second step, no concrete solution for an abstract test exists, it can be deduced, that this behavior is not possible (assuming that the constraints are correct), and therefore it is safe to neglect this test case.

Example 3. Consider the following complex boolean expression, taken from the ETCS specification [16]:

$$V_{est} > V_{max} + D V_{ck}(V_{max}) \lor D_{front} > D_{EBI}$$

This boolean expression describes a trigger condition for the onboard controller of the train to automatically activate the emergency brakes. Variable $V_{est}$ describes the current train speed. If this speed exceeds the maximum allowed speed $V_{max}$ by more than an offset $D V_{ck}$, the emergency brakes are activated. Alternatively, if the distance of the train $D_{front}$ is too close to the target location (this is expressed by $D_{front} > D_{EBI}$), the emergency brakes are activated as well.

Instead of explicitly describing this trigger condition by concrete model variables, an abstract input variable $t_{13}$ is introduced. The state machine transition that uses this abstract input variable looks like this: $[L1]_{t_{13}} \rightarrow [L2]$. In this case we introduced a new free input variable. It can be assumed, that the value of this variable can change arbitrarily over time. This is a safe overapproximation of the real system.
The replacement of a complex guard-condition by a new boolean input variable is always possible, but \( n \) replacements introduce \( n \) new inputs. In the worst case this results in \( 2^n \) new input equivalence classes of the SUT [10]: each feasible conjunction of the positive or negated Boolean expressions associated with each of the abstraction variables \( t \), specifies an input equivalence class. For realistic models, not all combinations of valuations of the complex conditions are possible. Using an SMT solver, impossible input combinations can be dropped a priori. These combinations are ignored in the test generation.

With this abstract description of the system behavior at hand, we are able to apply the test generation strategy presented in [10] and create a test suite, whose inputs are represented by abstract equivalence classes.

**Example 4.** For our running example, the TSM function, we introduced nine abstract input variables. From the \( 2^9 = 512 \) possible combinations, only eight combinations were valid. Two of the combinations can be considered as equivalent [10]. Thus, seven input equivalence classes were calculated from our abstract model.

The result of the first step in our test case generation is a set of abstract test cases. An abstract test case is a sequence of system states \( \langle s_0, \ldots, s_n \rangle \), where \( s_i : V_{\text{abstract}} \rightarrow D \) is the valuation function, mapping the (abstract) Boolean system variables \( t_k \) to their values \( \bar{s}_i(t_k) \) in the \( i \)-th state.

**Step 2. Concretization.**

While the first step defined abstract test cases that are related to computations in the abstract test model, the second step aims at constructing concrete test data.

A concrete test case is a sequence of concrete system states \( \langle \bar{s}_0, \ldots, \bar{s}_n \rangle \). Each \( \bar{s}_i \) is a valuation function \( \bar{s}_i : V \cup V_{\text{abstract}} \rightarrow D \), where the concrete system variables \( v \), including the inputs and outputs, as well as the abstraction variable introduced before, are mapped to concrete values \( \bar{s}_i(v) \in D \).

For the concretization of the test cases, parametric diagrams were added as an additional description means. The abstract variables introduced in Step 1 have to be related to concrete system variables by parametric diagrams. Two types of constraints can be defined: state invariants and temporal evolution.

State invariants constrain the concrete system variables in every state and, therefore, in every test step of the test case, these invariants have to be fulfilled. In a parametric diagram a state invariant can be described by a constraint property. A parametric diagram with \( m \) constraint properties that define state invariants (\( \text{INV}_1, \ldots, \text{INV}_m \)) yields the invariant:

\[
\text{INV} \equiv \bigwedge_{i=1}^m \text{INV}_i, \quad \text{INV}_i \in V_{\text{abstract}} \to \text{D}
\]

Invariants in the concrete system are also used to bind the values of abstract system variables to the concrete guard conditions they are abstracting.

**Example 5.** Considering trigger condition \( t_{13} \) from example 3 the corresponding parametric diagram is shown in Figure 4. In this diagram \( t_{13} \) gets bound to its concrete guard condition.

Given this invariant, and given an abstract test case \( \langle s_0, \ldots, s_n \rangle \), we want to calculate a concrete test case \( \langle \bar{s}_0, \ldots, \bar{s}_n \rangle \) that uses the same valuations of the abstract system variables – that is, the same equivalence classes – as prescribed by the abstract test case. This is encoded by the following formula, where \( \phi(t/x) \) denotes the formula derived from \( \phi \) by exchanging every free occurrence of \( x \) by the term \( t \).

\[
\bigwedge_{i=0}^n \bigwedge_{t \in V_{\text{abstract}}} s_i(t) = \bar{s}_i(t) \land \text{INV}[\bar{s}_i(v), v \in V \cup V_{\text{abstract}}]
\]

The above formula can be solved by an SMT solver. The result is a concrete mapping of variables from \( V \) to concrete values. Using these concrete values, a concrete test case \( \langle \bar{s}_0, \ldots, \bar{s}_n \rangle \) is calculated.

Apart from that, the temporal evolution of system variables is described by means of parametric diagrams. In contrast to an invariant, such constraints describe the change of values between two test steps. For example it may necessary to constrain the change of velocity and location due to the physical laws of acceleration. Figure 2 gives an example for this kind of constraints. This kind of constraints can best be described by differential equations.

In our approach we support the definition of linear differential equations. These equations can be discretized and translated to expressions relating pre and post states. These expressions contain unprimed variable symbols \( V \) describing the variable in the pre-state \( \bar{s}_i \), and primed variable symbols \( V' \) describing the post-state \( \bar{s}_{i+1} \). For example the differential equation \( \frac{dv}{dt} = a \) in Figure 2 can be expressed by the constraint: \( t' = t + \Delta t \land V_{\text{post}}' = V_{\text{post}} + a \cdot \Delta t \).

All constraint properties describing temporal evolution (\( \text{TEMP}_1, \ldots, \text{TEMP}_n \)) can be summarized:

\[
\text{TEMP} = \bigwedge_{k=1}^n \text{TEMP}_k
\]

We extend the SMT instance to generate concrete test cases respecting the time-continuous evolution defined by parametric diagrams as follows:

\[
\bigwedge_{i=0}^n \bigwedge_{t \in V_{\text{abstract}}} s_i(t) = \bar{s}_i(t) \land \text{INV}[\bar{s}_i(v), v \in V \cup V_{\text{abstract}}] \land \\
\bigwedge_{i=0}^{n-1} \text{TEMP}[s_i(v), v \in V \cup V_{\text{abstract}}], \quad s_{i+1}(v)' = v' \in V' \cup V_{\text{abstract}}
\]

**5. EVALUATION**

The approach described in the previous sections has been applied to the TSM function of the ETCS specification [16]. The specification describes the behavior in means of a transition table with the corresponding conditions for each transition. Such a specification can easily be modeled in terms of state machines and parametric diagrams: For every condition defined in the specification a Boolean variable \( t_{13}, \ldots, t_{13} \), \( r_0, \ldots, r_3 \) was introduced, named according to the definitions in [16]. The state machine can easily be deduced from the transition table using the introduced Boolean variables. The corresponding state machine of the TSM function is shown in Figure 3. The concrete conditions are described and the introduced Boolean variables are bound by means of parametric diagrams. An excerpt of the parametric diagram of the TSM function is given in Figure 4.

As a result, specifying the constraint blocks and creating the parametric diagram, as well as the state machine,
was straightforward, where only the mathematical equations given in [16] had to be translated to constraint blocks and some notations had to be slightly modified to be parsed by our tools. Thus, the modeling effort for this example was very low and we assume that the modelling effort should scale very well if other sub functions of this specification were modeled with the proposed approach.

**Model description.**

In the TSM function state changes occur as soon as trigger conditions are fulfilled. Under certain trigger conditions \((t_3, t_4, \ldots)\) the system changes its state machine state corresponding to different levels of intervention. \(t_3\) is a boolean expression that evaluates to true, if the current velocity \(v_{\text{rel}}\) and the current train front position \(D_{\text{front}}\) exceed the line indicated by \(D_f\) in Figure 5. Similarly, \(t_4, t_5, t_{10}\) and \(t_{13}\) return true if and only if the train exceeds \(D_p, D_n, D_{\text{crit}}\) or \(D_{\text{max}}\) respectively. When a new state machine state is entered, the \(entry\)-actions write to the output variables tco, sb and eb. These variables indicate the activation of the traction cut-off, the service brake and the emergency brake. The revocation conditions \(r_0, \ldots, r_3\) describe the conditions necessary for the intervention states to be revoked.

The **state invariants** of our test model are described by means of a parametric diagram and as an example in Figure 4 the definition of trigger condition \(t_{13}\) is shown. For convenience the definition of the other triggering conditions is omitted.

All these triggering conditions depend on the train velocity and track relative distance (position). The conditions can be calculated from the EBD curve, compare Figure 5. The EBD curve (emergency brake deceleration) maps the track relative distance to a velocity, assuming a negative acceleration \(A_{\text{safe}}\) such that zero speed is reached at the target position \(D_{\text{target}}\). The curve can be calculated by calculation of the braking distance \(s\) as shown in Figure 1. The value of \(A_{\text{safe}}\) is a conservative approximation that describes the least deceleration that the emergency brakes of the train achieve when fully activated. Because the full activation of the emergency brakes needs some time, the curve \(D_{\text{lim}}\) is shifted by a fixed time delay. If the current distance-velocity-pair describing the train speed and position is right or above this line, the emergency brakes are triggered. In a similar way the other triggering conditions are defined by shifting the EBD curve by fixed time delays.

The **temporal evolution** of our test model is described by the parametric diagram shown in Figure 2. The constraints in this diagram restrict the time-continuous evolution of the trains front position and the velocity. Therefore the acceleration \(a\) of the train is restricted to values in the range \([-10, 2]\). If the emergency brakes are activated (the train is in the state machine state EB), the deceleration must be higher than prescribed by \(A_{\text{safe}}\). As described in section 4, the discretized versions of the constraints were used.

In total, our test model used for this evaluation contains 52 constraint properties, including those shown in Figure 1, 2 and 4.

**Test case generation.**

Table 1 gives a summary of our test case generation. The test case generation was performed on a system with 24 CPU cores running at 2.8 GHz with 16 GB RAM. In the first step of our test case generation 50 abstract test cases were calculated, which took less than a second. Then it took three and a half hours to generate the concrete test cases from the 50 abstract test cases. The main reason, that this computationally expensive task, is still applicable in acceptable time, is that SMT instances for every test case can be solved in parallel. The test case generation was run in parallel and on average 16.7 of the 24 CPUs were utilized.

Given that the results from this TSM model can be generalized to hybrid systems of comparable complexity, the runtime is acceptable. Yet, the results indicate, that the utilization of a different computational backend, e.g. analytical or numerical solvers, should be considered.

For every abstract test case a concrete solution was calculated. Thus, our model abstraction seems appropriate. This might not always be the case. If the test case generation of abstract test cases yields a low percentage of solvable concrete test cases, this might be an indicator for inappropriate abstraction or too restrictive constraints.

<table>
<thead>
<tr>
<th>number of test cases generated</th>
<th>50</th>
</tr>
</thead>
<tbody>
<tr>
<td>avg. test case length (test steps)</td>
<td>2.84</td>
</tr>
<tr>
<td>time for abstract test case generation</td>
<td>0.8s</td>
</tr>
<tr>
<td>time for calculation of concrete test cases</td>
<td>3.5h</td>
</tr>
<tr>
<td>memory for test case generation</td>
<td>8.4 GB</td>
</tr>
</tbody>
</table>

**Table 1:** Results of the Test Case Generation for the TSM Model

6. **CONCLUSION**

We have presented a novel approach to automated model-based testing of mixed time-discrete and time-continuous systems. It has been shown that SysML is a suitable formalism for creating test models of this kind: time-discrete control aspects are reflected by state machines, and time-continuous constraints are represented by constraint blocks and parametric diagrams. Concrete test cases were generated in a two-step approach. At first, an abstraction of the concrete system was constructed, and an input equivalence class partition strategy with proven fault detection capabilities was applied to generate a test suite of abstract test cases. The inputs associated with each of these test cases
are sequences of input equivalence classes. For the second step, concrete solutions for each of these abstract test cases were generated. For this purpose, each abstract test case is transformed into constraints over concrete model variables, and the additional constraints coming from time-continuous conditions specified in constraint blocks are taken into account as well.

Using an SMT solver that is capable of processing floating point arithmetics, it was shown that the approach is suitable for practical application. To this end, tests for a safety-relevant control function of the ETCS onboard controller have been generated in an automated way. The results show that MBT is feasible for hybrid systems that are comparable to the complexity of this real-world example. The model-based test generation and execution process has been implemented in an experimental version of an industrial-strength testing tool. As mentioned, we are currently exploring numerical solvers or computer algebra systems to get a better performance and to support more complex mathematical expressions.

7. REFERENCES


