Distributed-Code Generation from Hybrid Systems Models for Time-delayed Multirate Systems

Madhukar Anand, Sebastian Fischmeister, Jesung Kim, and Insup Lee
Dept. of Computer and Information Science, University of Pennsylvania
Philadelphia, PA, USA
{anandm,sfischme,jesung,lee}@saul.cis.upenn.edu

ABSTRACT
Hybrid systems are an appropriate formalism to model embedded systems as they capture the theme of continuous dynamics with discrete control. A simple extension, a network of communicating hybrid automata, allows for modeling distributed embedded systems. Although it is possible to generate code from such models, it is difficult to provide formal guarantees in the code with respect to the model. One of the reasons for this is that, the model is set in continuous time and concurrent execution with instantaneous communication, whereas the generated code is set in discrete time with delayed communication. This can introduce semantic differences between the model and the code such as missed transitions, faulty transitions, and altered continuous behavior. The goal of faithful code generation is to minimize these differences.

In this paper, we propose a relaxed criteria of faithfulness, coined relative faithful implementation. Based on this criteria, we propose dynamically adjusting the guard at runtime using estimates of errors for preventing faulty transitions. We also identify a sufficient condition to ensure no missed transitions in the code.

Categories and Subject Descriptors: D.2.4 [Software/Program Verification]: Reliability.

General Terms: Reliability, Design, Verification.

Keywords: Hybrid Systems, Distributed Systems.

1. INTRODUCTION
Hybrid systems are an appropriate modeling paradigm for embedded control software, because it can be used to specify continuous change and discrete transition of system states [1, 17]. The benefits of precise modeling can be enhanced if code is generated automatically while maintaining a correspondence with the model.

Since computer systems are discrete, code generation from hybrid systems models requires a discretized hybrid systems model. Such a model uses a designer-specified rate by which the continuous state evolves. Since a single processor executes discretely, concurrency in the model is broken on several processors (either on a single system or multiple systems).

Discretization affects switching. In the discretized code, transitions can only occur at the defined rate, i.e., the sampling time. If a transition in the model is only enabled between the sampling times, then it will be missed, called missed transition, and the code will behave differently from the model. Example 1 below illustrates such a missed transition.

EXAMPLE 1. The system comprises two communicating hybrid automata \( A_0 \) and \( A_1 \). \( A_0 \) executes every \( 0.0002s \) and \( A_1 \) every \( 0.0003s \). The communication delay between them, denoted by \( \phi_{12} \), equals \( 0.0004s \) and \( P_A \) denotes the state of the automaton. The following execution trace shows that this discretization leads to a missed transition \( q_2 \) in \( A_1 \). In the trace, we use \( x(A_i) \) to denote the local estimate of \( x \) at \( A_i \).

\[ t \quad x(A_0) \quad P_{A_0} \quad y(A_1) \quad x(A_1) \quad P_{A_1} \]
\[ 0.0002 \quad 0.0002 \quad q_0 \quad 0.0008 \quad 0.0008 \quad q_1 \]
\[ 0.0003 \quad 0.0004 \quad q_0 \quad 0.0008 \quad 0.0004 \quad q_1 \]
\[ 0.0004 \quad 0.0004 \quad q_0 \quad 0.0003 \quad 0.0004 \quad q_1 \]
\[ 0.0006 \quad 0.0012 \quad q_2 \quad 0.0006 \quad 0.0008 \quad q_1 \]

This can be prevented, if the rate is high enough compared to the enabled interval of the guard [3]. On the other hand, the rate must be low enough to provide sufficient computing time on the target platform between each execution.

![Figure 1: An example for missed transition.](image)

A faulty transition is an invalid change of system state due to an incomplete snapshot of the system’s state. The transition would be disabled, if the snapshot were accurate. For instance, Example 1, if \( G_1 \) is changed to \( x \in [0.0005, 0.0008] \), then \( A_1 \) will transit to \( q_2 \) at time \( 0.0006 \). This transition is faulty, since at that instant \( x = 0.0012 \) and the guard should not be enabled.

**Problem Statement.** In this paper, we consider the problem of faithful implementation of hybrid-systems models with the following assumptions: (1) We consider a system of rectangular linear...
hybrid automata models in which the dependency of different automata is either through dynamics or through switching and the dependency graphs are acyclic. (2) We consider a distributed implementation of the different agents which may have different sampling rates and (3) the system is also assumed to be affected by communication delays that could be arbitrary but bounded. Given such a model, the focus of this paper is to address faulty transitions, missed transitions, and identify sufficient conditions for preventing them.

Proposed Approach. In [14], we have proposed a scheme to prevent faulty transitions by shrinking the guard set to account for the errors due to mixed sampling times. However, introducing such a margin into the guard in turn increases the chances of a missed transition, since it reduces the time interval during which a transition must be taken. We mitigate this problem in this paper by proposing a runtime technique that extends on our previous approach. The idea is that some kinds of errors, e.g., communication delay between distributed components, are better analyzed at runtime as they depend on current dynamics. By instrumenting at runtime, we can avoid a pessimistic decision to avoid a faulty transition. Our scheme is based on the premise that the communication delay can be bounded. To show feasibility, we present a possible configuration of time-deterministic distributed system in Section 4.

Related work. Commercial modeling tools such as SIMULINK also support code generation and address the effect of errors in the code. However, their concerns are largely limited to numerical errors occurring each step during simulation, and the effect of such errors on to discrete behavior is not addressed rigorously. Synchronous languages for reactive systems, such as STATECHARTS [12], ESTEREL [6], and LUSTRE [11], also support code generation. However, they do not support hybrid systems modeling. SHIFT [7] is a language for hybrid automata that also supports code generation, but the focus is on dynamic networks. Model-based development of embedded systems is also promoted by other projects with orthogonal concerns: Ptolemy supports integration of heterogeneous models of computation [8] and GME supports metamodeling for development of domain-specific modeling languages [15].

Code generation from hybrid models was introduced with focus on single-thread execution in [3]. This was extended to multi-threaded models accounting for faulty transitions in [14] and unidirectional distributed systems with no delays in [4]. To guarantee value and time determinism in the discretized code, we use an E machine approach [13] with the concept of logical execution time in our on-going prototype. However, we plan to generate E code for the E machine directly from the CHARON model without using the Giotto specification language.

The remainder of the work is structured as follows: Section 2 introduces the system model. Section 3 presents our approach of how to prevent missed and faulty transitions. Section 4 provides an overview of the implementation and Section 5 describes the current and future work.

2. SYSTEM MODEL

In this framework, we assume that there is a network of hybrid automata (called agents) communicating via a set of shared variables. We denote a single agent by $A \equiv (A, SV)$ where $A$ is the model for the agent and $SV$ is the set of shared variables. A system of communicating hybrid agents is a tuple $C = \{A_0, \ldots, A_n\}$.

**Definition 1. (DCHA)** Given a system of communicating hybrid agents $C$, and a discrete time domain $LT = \{l_0, l_1, \ldots\}$ where $l_i \in \mathbb{Q}^+$ and $v_i, l_{i+1} > l_i$, the discretized system of communicating agents (DCHA) is given by $D = \{(A, SV, LT)_{l_0}, \ldots, (A, SV, LT)_{l_n}\}$. □

Note that the DCHA is the discretization of the continuous model that is implemented on an actual platform. The guarantees for faithful implementation are given with respect to this model. In [14], the authors provide a rigorous definition of a system of communicating agents and their semantics.

When the agents of the discretized model is mapped to real-time tasks for execution, each agent is assigned a period. This period is normally equal to the frequency of evaluation for an agent.

**Definition 2. (Code)** The code, implementing a DCHA $D$ denoted by $K$, is given by the tuple $K = \{(A, SV, LT, clk, h, \sigma)_{l_0}, \ldots, (A, SV, LT, clk, h, \sigma)_{l_n}\}$, where $clk$ represents the physical time, $\sigma$ represents the local copy of the shared variables, $h$ is the frequency of evaluation. □

The frequency $h$ reflects the logical time between evaluations, i.e., $(l_{i+1} - l_i)$. Typically, this is constant for any agent and is equal to the period of the real-time task on which the agent is mapped. In a multirate system, different agents have different rates. This version abstracts away several details, e.g., the set of synchronization mechanisms, the collection of subroutines implementing the DCHA, scheduling details, and exact mapping of code into memory. In [5], we provide a descriptive exposition of this.

Finally, we denote the logical communication time between agents $A_i$ and $A_j$ by $\phi(i, j)$ and a time-delayed code implementation of $D$ by $K, \phi$.

3. VALIDATION OF GENERATED CODE

To provide a faithful implementation, we have to guarantee the absence of faulty and missed transitions [4]. The implementation should also be independent of the scheduling. However, in our system, delays can vary for each agent and also between single executions of the same agent. Such behavior introduces errors resulting in incorrect behavior. For example, in one mode, an agent could depend on $x$ that arrives 0.0001 s late and in another mode, it could depend on $y$ that arrives 0.0002 s later. Therefore, we relax the criteria for a faithful implementation by requiring that code enters the state of the model no later than the maximum possible delay. Formally, we can define.

**Definition 3. (Relative Faithful Implementation)** Let $VC$ be the set of all variables and $\alpha_x$ be the maximum bound on the error of a variable $x$. Given a trace of states of the code $K$ for an agent $A_j$, $(q_0, q_1, \ldots)$, at physical time-stamps $(clk_0, clk_1, \ldots)$, if $\forall clk_k$

1. $\forall x \in VC$, $|x_D - x_K| < \alpha_x$, where $x_D$ and $x_K$ represent the value of variable in the code and the model respectively.

2. $\forall j \ni q_0, q_0 = q_0$, $(|t_D - t_K| < \phi_j + \varphi)$ where $q_c$ is the state of the code at logical time $t_K$, at physical time $clk$, $q_c$ is the projection of the state of the model onto the code for $A_j$ at logical time $t_D$, $\phi_j$ is the maximum skew due to different rates of updates.

then, code for $A_j$ is said to be a relative faithful implementation. If $\forall j$, the code for $A_j$ is a relative faithful implementation, then $K$ is a relative faithful implementation of $D$. □

Condition 1 states that the error in continuous variables be bounded and Condition 2 states that the delay in timing of transitions be bounded. We can now provide approaches that safeguard against faulty and missed transitions and thus allow for a relative faithful implementation of a hybrid-systems model.
3.1 Preventing Faulty Transitions

A faulty transition occurs either because the transition was taken on the basis of an older value of the variable or because of numerical errors in the variables.

Static instrumentation was introduced in [14] as an approach to prevent faulty transitions. The idea there was to instrument the guards and the invariants with maximum possible error in variables and switch conservatively.

**Definition 4. (Static Instrumentation)** Let \( p \) be the state of agent \( A \) with \( E_A(p) \) being the set of discrete transitions, \( I_A(p) \) the set of invariants in that state, and \( G_A(e), e \in E_A(p) \) the set of guards. Let \( \forall x, \delta_{p,x} \) give the upper bound on the numerical and timing errors. If \( l \) and \( u \) denote lower and upper bounds respectively, then,

1. \( \forall p, \forall I \in I_A(p), \)
   \[ l(I_x) \leftarrow l(I_x) + \delta_{p,x}, u(I_x) \leftarrow u(I_x) - \delta_{p,x}. \]
2. \( \forall p, \forall G \in G_A(e), \)
   \[ l(G_x) \leftarrow l(G_x) + \delta_{p,x}, u(G_x) \leftarrow u(G_x) - \delta_{p,x}. \]

where \( G_x, I_x \) denote the projection of the guard and invariant on the variable \( x \) in agent \( A \) in state \( p \).

Once the guards and the invariants have been instrumented, the code generated from \( D \) can be assured of no faulty transitions. This is true because the instrumented guard remains disabled for the duration of synchronization and possible numerical errors. For the proof of this claim, we refer the reader to [14]. Though guarantees can be given statically with this scheme, since the guard and invariant sets are shrunk, the probability of not taking a transition increases. Yet another disadvantage is that it is not always possible to determine this error bound beforehand as with most dynamic equations, it is only possible to get a local estimate of error which is only available at runtime.

The two factors contributing to the error leading to faulty transitions are (1) Numerical errors (e.g., round-off, truncation) and (2) Timing related errors. While the numerical errors are static, the timing errors change over time as they comprise of communication delay \( \phi \) which is bounded and the skew due to different frequencies of update that is dynamic. If \( h_i \) is the agent’s \( A_i \) execution frequency and depends on variable \( x \) updated by agent \( A_i \) with frequency \( h_i \) and \( \phi(i,j) \) is a fixed constant, then, the maximum skew

\[ \varphi_{max}^x \text{ is given by } \varphi(1,j)_{max} = \max_{i \in [1..N]} \left( n p_j - \frac{n p_j - \phi(i,j)}{p_i} p_i \right) \]

where \( N = \frac{LCM(p_j,j)}{p_j} \).

Now, if we instrument the guard taking this into account at runtime, we can improve over the static approach. Another advantage of a runtime approach is that we can have a value-specific simplification leading to less pessimistic bounds. For example, if \( \delta > 0 \), \( x \) is an increasing function of time and hence the update for \( x \) will be greater than the current estimate. Therefore, if we are evaluating a condition such as \( x > 5 \), we will not have to instrument it for timing errors.

**Definition 5. (Dynamic Instrumentation)** Let \( p \) be a state of agent \( A_i \) with \( E_{A_i}(p) \) being the set of discrete transitions, \( I_{A_i}(p) \) the set of invariants in that state, and \( G_{A_i}(e), e \in E_{A_i}(p) \) the set of guards. Let \( \forall x, \beta_{p,x}(t), \gamma_{p,x}(t), t \in T = [l_t, l_t + d] \) give the bound in \( T \) due to timing and numerical errors, respectively. If \( I \) is an indicator function, \( l \) and \( u \) denote the lower and upper bounds respectively, and the condition on derivatives holds in \( T \), then,

1. \( \forall p, \forall I \in I_{A_i}(p), \)
   \[ l(I_x) \leftarrow l(I_x) + I(\delta \geq 0, t \in T) \beta_{p,x}(t) + \gamma_{p,x}(t) \]
   \[ u(I_x) \leftarrow u(I_x) - I(\delta \leq 0, t \in T) \beta_{p,x}(t) - \gamma_{p,x}(t) \]

2. \( \forall p, \forall G \in G_{A_i}(e), \)
   \[ l(G_x) \leftarrow l(G_x) + I(\delta \geq 0, t \in T) \beta_{p,x}(t) + \gamma_{p,x}(t) \]
   \[ u(G_x) \leftarrow u(G_x) - I(\delta \leq 0, t \in T) \beta_{p,x}(t) - \gamma_{p,x}(t) \]

where \( G_x, I_x \) denote the projection of the guard and invariant on the variable \( x \) in agent \( A \) in state \( p \).

Note that with dynamic instrumentation, we have to calculate the error that would happen in the future, i.e., the error that would be caused when a particular value is read.

**Example 2.** Consider the case of Example 1 with \( G_1 : x \in [0.0001, 0.001] \). If we statically instrument, we have to change the guard \( G_1 \) to \( x \in [0.0005, 0.0006] \) because the maximum error \( \alpha_{e} = 2(\varphi(1,2)_{max} + \phi(1,2)) = 2(0.0001 + 0.0001) = 0.0004 \) and then we can have a run as,

<table>
<thead>
<tr>
<th>( t )</th>
<th>( x )</th>
<th>( P_{A_1} )</th>
<th>( y )</th>
<th>( z(A_1) )</th>
<th>( P_{A_2} )</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.0002</td>
<td>0.0004</td>
<td>0.0006</td>
<td>0.0006</td>
<td>0.0001</td>
<td>0.0004</td>
</tr>
<tr>
<td>0.0005</td>
<td>0.0006</td>
<td>0.0007</td>
<td>0.0007</td>
<td>0.0004</td>
<td>0.0001</td>
</tr>
<tr>
<td>0.0004</td>
<td>0.0008</td>
<td>0.0007</td>
<td>0.0007</td>
<td>0.0004</td>
<td>0.0001</td>
</tr>
<tr>
<td>0.0006</td>
<td>0.0012</td>
<td>0.0006</td>
<td>0.0006</td>
<td>0.0005</td>
<td>0.0003</td>
</tr>
</tbody>
</table>

The transition to \( q_2 \) is missed at time \( t = 0.0006 \). If we had used dynamic instrumentation, then, the guard would have been instrumented to \( x \in [0.0003, 0.0008] \), since the error \( \alpha_{e} = 2(\varphi(1,2) + \phi(1,2)) = 2(0+0.0001) = 0.0002 \) and the transition to \( q_2 \) would have been taken.

The theorem below formally states that dynamic instrumentation prevents faulty transitions.

**Theorem 1.** Let the code \( K \) of the model \( D \) be implemented on a distributed platform. For every agent \( A_i \), let \( p \) be the current state with \( I_{A_i}(p) \) the set of invariants in that state, and \( G_{A_i}(e) \) the set of guards. If \( \forall G \in G_{A_i}(e) \) that evaluate to true, \( G \) is dynamically instrumented with \( \beta_{p,x}(t) \) and \( \gamma_{p,x}(t) \) for all shared variables \( x \), then, there will be no faulty transitions.

The timing errors can be computed in practice, by delegating this responsibility to different entities such as the agent that writes the shared variable or the agent that reads with different advantages.

3.2 Preventing Missed Transitions

Missed transitions are transitions that are enabled in the model but not taken in the code. They occur either because the guard is not evaluated sufficiently or scheduling affected the order of evaluation. In general, a transition will not be missed, if it stays enabled long enough to be detected. The theorem below gives a sufficient condition to prevent missed transitions.

**Theorem 2.** Let the code \( K \) of the model \( D \) be implemented on a distributed platform, \( h_i \) and \( h_j \) be the frequency of sampling in \( A_i \) and \( A_j \) respectively, and the instrumented guard set \( G(p) \leq I(p) \) be such that an active transition is enabled in state \( p \). If \( G(p) \) and \( I(p) \) overlap by \( \Omega = 2h_j + \sum_{k \in A_k} \beta_{k} \delta_{k} + \varphi_{k} \delta_{k} \), where the summation is over all frequencies along the longest path in the switching dependency graph \( E \) of \( G \), and \( \varphi_{k} \delta_{k} \) represents the maximum skew at agent \( A_k \) due to \( A_j \), then, the transition will be detected and will not be missed if they are taken as soon as enabled.

**Proof.** (sketch) First consider the case where the transitions are dependent on independent updates (such as sensor readings) but not updates from other agents. In the code, evaluation of the guard condition might be scheduled at some time \( j h_i, j \in Z^* \) but...
a guard may be enabled in the model immediately after that i.e., at time \( jh + \varepsilon, \varepsilon > 0 \). This will be detected in the code during the next evaluation which may be scheduled as late as \( (j + 2)h - \eta \). Since we assume that the transition will be taken as soon as detected, this transition will be taken at \( (j + 2)h - \eta \). Letting \( \eta, \varepsilon \rightarrow 0 \), in the worst case, the instrumented guard and invariant should overlap by at least \( 2h \) to avoid missing the transition.

Now consider the general case where the transition may depend on updates from other agents. Here, an update may be delayed by the communication \( (\phi) \) and the skew due to different rates \( (\psi) \) in addition to \( 2h \) at each agent due to sampling. By observing that this delay could be as large as sum of delays along the longest path in the switching dependency graph, we get the desired result.

The condition for overlap can be checked statically starting from the node with least dependency and iterating until the most dependent. Theorems 1 and 2 provide a sufficient condition to ensure a relative faithful implementation that we record in the following corollary.

**Corollary 1.** Let the code \( K \) of the model \( D \) be implemented on a distributed platform. If, (1) in the code for every agent \( A_i \), every dynamically instrumented \( G \in G_i \), and corresponding invariant \( I \) satisfy the condition of overlap in Theorem 2, and (2) all variables in \( K \) have bounded error, then, \( K \) is a relative faithful implementation of \( D \).

4. IMPLEMENTATION

We use CHARON [2] to implement ideas introduced in this paper. CHARON allows for modular specification of interacting hybrid systems and supports automatic code generation [3]. Once the model is specified in CHARON, the code generator for an agent takes a sampling period as an input and produces code that approximates the continuous behavior of the model. During this process, the guards and invariants are statically instrumented for numerical errors and the condition of overlap in Theorem 2 is used to check whether the code can be guaranteed against missed transitions. The responsibility of dynamic instrumentation to ensure no faulty transitions is delegated to the runtime environment. Specifically, the runtime environment, (1) ensures that the updates are delivered time deterministically, and, (2) provides the values of delay \( (\phi) \) and skew \( (\psi) \) for dynamic instrumentation.

As an underlying infrastructure of the distributed real-time system in our ongoing implementation, we use an existing system [16, 10, 9], which supports the concept of logical execution time (LET) [13]. This system provides time and value determinism across distributed nodes and is therefore apt to implement distributed CHARON models with constant and guaranteed delays at the cost of runtime performance. The worst-case execution time of an agent and the worst-case communication time together define the LET of the agent and thus its highest execution rate. This allows us to run each agent on a different node in a distributed system while it is oblivious of the distribution. If required, all the modified values can be communicated after the agent finishes execution and before the end of the LET. The limiting factor here is the shared communication channel between nodes. Given the agents’ frequencies and their node assignment, we can calculate the required communication, build a communication schedule, and validate that there are no collisions on the network. In [16], the authors describe how this can be done with tasks and the algorithms and tools can be reused for this implementation.

5. CONCLUSIONS

Hybrid systems based code generation is a promising yet challenging approach for producing reliable embedded software. Providing formal guarantees is difficult due to the semantic differences between the model and the code arising as a result of discretization and communication delays.

In this paper, we have presented an approach to guarantee faithful switching semantics that involves preventing missed and faulty transitions. In contrast to related and prior work, we have defined a notion of relative faithful implementation for systems with multiple rates of evolution. Based on this notion, we have proposed a runtime instrumentation technique to prevent faulty transitions and identified sufficient conditions to prevent missed transitions in the generated code. We have also sketched an implementation to show how timing delays can be bounded and accounted for, in the model.

In the future, we will consider fully integrating the infrastructure into the CHARON development environment and provide comprehensive techniques to detect missed transitions in the code at runtime.

6. REFERENCES


