Verification of embedded control programs

Thao Dang¹, Bertrand Jeannet², and Romain Testylier¹

Abstract—In this paper we are concerned with the problem of verifying embedded control programs. The approach we use combines the logico-numerical techniques developed for the verification of Lustre programs and the set-based image computation for continuous systems. The practical interest of this approach lies in the fact that there exists a tool for generating Lustre code for controllers described in Simulink. We also illustrate the approach with some experimental results obtained for a robotic controller for LEGO Mindstorm.

I. INTRODUCTION

Model-based design has emerged as an important paradigm in the development of cyber-physical systems. Its essence is the use of models in all the steps of the development process, from design to implementation. Using models is key to keeping the development costs manageable, since correcting bugs on models is much cheaper than on real prototypes. This paradigm can however be effective only when it is accomplished by powerful design tools for simulation, verification, synthesis, code generation, etc. In a cyber-physical system, the control and computation part is often realized as computer programs that will be executed to control the physical part. The design of the control and computation part can be assisted by a variety of tools, among which Simulink has become a de facto standard in industry and is extensively used in various application domains. Simulink supports the description and analysis of models in a graphical block-diagram language. The Simulink environment provides extensive capabilities for simulation and control design based on advanced algorithms. Embedded code generation from Simulink models is supported by tools, such as Real-Time Workshop Embedded Coder from Mathworks, TargetLink from dSpace. Although these code generators translate models into production-quality C code that can be deployed onto embedded systems, their major problem with respect to safety critical applications is the preservation of semantics. The question of fidelity between the generated code and the simulated models, that is what is simulated is what is executed, is not addressed by these tools. For this reason, an alternative approach is to translate Simulink models to an intermediate formal language, which permits formal reasoning on the relation between the models of each step in the design process. This work is developed within the framework of model-based development of embedded control design and we use Lustre as intermediate formal language and a tool for translating Simulink models to Lustre. The choice of Lustre in this approach is motivated by its formal semantics and a number of associated formal validation tools. Lustre is also supported by a code generator from SCADE² which is compliant the DO-178B Level A standard³.

This work contributes a method for the verification of Lustre programs generated from controller models specified in Simulink. Such Lustre programs contain both Boolean and numerical variables. Their verification often requires an exhaustive enumeration of modes corresponding to all possible values of the Boolean variables, which often leads to a state-space explosion. In this paper we employ a technique for verification of logico-numerical programs without resorting to the enumeration of the Boolean state space. On the other hand, to extend the technique to programs with non-linear computation, we combine it with a method for set-based image computation for polynomial functions.

II. FROM SIMULINK CONTROLLER MODELS TO LUSTRE PROGRAMS: A ROBOTIC EXAMPLE

A controller designed in Simulink can be translated to a Lustre program using the tool S2L [14]. The tool translates only the discrete-time blocks of Simulink, since implemented controllers are in discrete time. This means that to be translated, a controller that is designed in continuous time needs to be discretized using a sampling method. In addition, only a ”safe” subset of Simulink blocks can be translated. This restriction is needed to guarantee semantics preservation, that is the “informal” semantics of Simulink specified by the behavior of the simulator for a given set of simulation parameters and options.

Throughout the paper we use a case study of robotic planning to explain our method. This case study is simple (compared to advanced control and planification methods) but sufficiently complex to illustrate the interest of our method. The goal of this case study is to verify a controller for a LEGO-Mindstorm robot. The robot is in form of a two-wheeled vehicle and should follow a black line on the ground while avoiding obstacles. The trajectory of the robot is defined by 3 variables: its position \((x, y)\) on the plane and its orientation \(\theta\) (compared to the \(x\) axis). These state variables are controlled by two control inputs: the speeds \(v_l\) and \(v_r\) of the left and right wheels. The dynamics of the

¹See http://www-verimag.imag.fr/Tools,36.html for the available tools for Lustre
²SCADE is the graphical version of Lustre commercialized by Esterel Technologies, which is used in highest criticality applications, such as European avionic projects (Airbus A340-600, A380, Eurocopter).
³This industrial standard specifies objectives across the development cycle to achieve flight software certification.
robot is described by the following differential equations:

\[
\begin{align*}
\dot{x} &= (v_l + v_r) \cos(\theta) / 2 \\
\dot{y} &= (v_l + v_r) \sin(\theta) / 2 \\
\dot{\theta} &= (v_r - v_l) / l
\end{align*}
\]

where \( l \) is the distance between the two wheels.

The robot is equipped with two light sensors and one ultrasonic sensor. The light sensors are used to measure the deviation of the robot from the black line, and the ultrasonic sensor to detect the presence of an obstacle in front of the robot. A controller was designed for this control objective as follows.

The controller consists of two modes: one is for following the line, and the other is for avoiding obstacles. The controller takes as input the data from the sensors and computes the desired speeds \( v_l \) and \( v_r \) of the left and right wheels which are realized by the robot motor. The control law for the line-following mode is:

\[
\begin{align*}
v_d &= 50c_d + 50c_g^2 \\
v_g &= 50c_g + 50c_g^2
\end{align*}
\]

where \( c_d \) and \( c_g \) are the values of the right and left light sensors (ranging in the interval \([0, 1]\)). The control law for the mode of obstacle avoidance mode is as follows. When an obstacle is detected, the robot should first stop, make 180-degree turn (for example towards the left), and then continue to follow the black line in the other direction. Since there is a delay in the motor’s actions, it is important to wait until the robot stops fully before turning around (otherwise, turning around while advancing may make the robot lose track of the black line). In addition, the 180-degree turn ends when a sequence of colors White-Black-White is detected on the left light sensor.

The above controller was implemented as a Simulink model. In order to analyze the behavior of the controller in the closed loop, the dynamics of the robot and the outputs \( c_d \) and \( c_g \) of the light sensors are also emulated by a Simulink model. Since the proposed verification method works for polynomial dynamics, the cosinus and sinus functions are approximated by piecewise linear functions.

The resulting model of the closed loop system contains many switch blocks and Boolean operators (such as to encode the piecewise linear approximation and detection of a White-Black-White sequence). The Simulink model can be translated to a Lustre program using the tool S2L.

Such a Lustre program corresponds to a discrete-time hybrid automaton, which can be defined by exhaustively enumerating all the possible Boolean values in the program to construct a control flow graph with only numerical variables. Although the above controller is simple in terms of control laws and planification strategies, by a simple exploration of the control flow graph of a medium-sized Lustre program, the discrete structure of the corresponding hybrid automata can already be very large and this discrete structure may contain a lot of modes sharing the same continuous dynamics.

To address this problem, we employ the logico-numerical techniques for Lustre programs.

### III. Techniques for Verification of Logico-Numerical Programs

A Lustre program defines a set of equations:

\[
\begin{align*}
s_1 &= f_1(s_1, \ldots, s_n, u_1, \ldots, u_m) \\
&\quad \ldots \\
s_n &= f_n(s_1, \ldots, s_n, u_1, \ldots, u_m)
\end{align*}
\]

where \( S_i \) are variables and \( u_i \) are input variables. The variables in Lustre are also called flows and they can be Boolean and numerical. A flow is a partial function \( y : N \to D_y \) where where \( N \) is the set of natural numbers modelling discrete time, and \( D_y \) is the domain of \( y \). Besides the usual arithmetic operations and control-flow operators (e.g., if then else), the functions \( f_i \) are made up of a number of temporal operators, in particular the \( \text{pre} \) operator which is used to specify a delay of one time unit, that is \( \text{pre}(y) \) defines a flow \( z \) such that the value of \( z \) at instant \( k \) is equal to the value of \( y \) at instant \((k-1)\).

The compilation of Lustre programs produces a symbolic transition systems with the following components: \( s = (b, x) \) is the vector of the state variables composed of Boolean variables \( b \) and numerical variables \( x \), \( u = (u_b, u_x) \) is the vector of the input variables composed of Boolean variables \( u_b \) and numerical variables \( u_x \). \( C(x, u_x) \) is the set of Boolean constraints over the numerical variables. The transitions are specified as:

\[
A(b, u_b, C) \rightarrow \begin{pmatrix} b' \\ x'
\end{pmatrix} = \begin{pmatrix} f^b(b, u_b, C) \\ f^x(b, u_b, C, x, u_x)
\end{pmatrix}
\]

where \( A(b, u_b, C) \) is an assertion constraining the input variables depending on the state variables. The numerical transition functions are written as a disjunction of guard actions: \( f^x(b, u_b, C, x, u_x) = \bigvee g_i(b, u_b, C) \rightarrow a_i(x, u_x) \).

An execution of such a program is a sequence

\[
s^0 \rightarrow s^1 \rightarrow s^2 \ldots
\]

such that

\[
\forall k \geq 0, A(b^k, u^k_b, C) \rightarrow s^{k+1} = f(s^k, u^k)
\]

where \( f = (f^b, f^x) \) is the vector of both Boolean and numerical transition functions. In this work, to define the control flow graph, we consider partitions defined by equivalence relations on the Boolean state variables. A fully partitioned control flow graph constructed by enumerating all Boolean states is characterized by the following equivalence relation:

\[
b_1 \sim b_2 \iff b_1 = b_2.
\]

To combine symbolically discrete and continuous behaviors, we adopt the abstract interpretation approach for approximations, with adaptive techniques to improve the precision when necessary. Concerning enumerating continuous modes, based on the observation that, in cyber-physical
systems, many discrete states are likely to share the same continuous dynamics, and that the number of different dynamic behaviours remains small in many cases, we can build an automaton with one location per continuous mode, and use techniques for discrete systems to handle with the complex transitions generated by the controller program.

On the other hand, separating discrete modes only when they do not share the same continuous mode may lead to a quite abstract view of the system w.r.t. the qualitative behaviors. We adapted the techniques developed for purely discrete systems [8] to refine the control structure w.r.t. precision requirements. In order to choose automatically a suitable partition for the verification of a safety property, that is precise enough to conclude if the property holds, and also the coarsest possible to limit the complexity of analysis, the adopted solution consists in starting from a very coarse partition (basically, distinguishing only between initial states, bad states and others), and in refining it dynamically according to heuristics. The above mentioned techniques for verification of logico-numerical programs were implemented in the tool NBAC [9], [8], which can also be used for hybrid systems with piecewise-constant dynamics. In this work, with view to verification of embedded control programs with nonlinear numerical transition functions, we combine the tool with a method for computing the image of polynomial transition functions.

Before continuing, we briefly discuss related work regarding symbolic treatment of discrete dynamics. The standard approach to combine continuous and (finite) discrete dynamics consists in encoding the values of the finite-state discrete variables in the control structure of the hybrid systems [7], [5]. However, this easily results in a combinatorial explosion of the number of control locations. More recently, [3] proposed a fully symbolic technique based on backward (greatest) fixpoint computation, in which sets of states are represented with a variant of Boolean circuits mixing Boolean variables and linear constraints. The technique relies on a semi-canonical representation for sets of states.

### IV. IMAGE COMPUTATION FOR PARAMETRIC POLYNOMIAL DYNAMICAL SYSTEMS

In this section, we present a technique for computing the image of a polynomial function, which is used to handle transition functions of Lustre programs. We also augment the framework with uncertain parameters, since they are very useful in practice to model imprecision in modelling. Compared to the technique presented in [11], this technique is more efficient in term of accuracy since it handles parameters in a more symbolic way.

We consider the following numerical transition function:

\[ x(k + 1) = \pi(x(k), p) \]

where \( x \in \mathbb{R}^n \) is the state variables, \( p \in P \subseteq \mathbb{R}^m \) is the vector of uncertain parameters. The set \( P \) is called the parameter set. We assume that \( P \) is a convex polyhedron. The initial state \( x(0) \) is inside some set \( X_0 \subseteq \mathbb{R}^n \). Note that the above function \( \pi \) plays the role of a guard action in the symbolic transition system associated with a Lustre program.

One way to approximate the image \( \pi(X) \) is to use special convex polyhedra with fixed geometric form, called template polyhedra [13], [2]. A template is a set of linear functions defined by an \( l \times n \) matrix \( H \) and a real-valued vector \( c \in \mathbb{R}^l \). Let \( H^i \) be the \( i^{th} \) row of the matrix \( H \) and \( H^i_k \) its \( k^{th} \) element, the corresponding template polyhedron is defined by

\[ \langle H, c \rangle = \{ x \mid \bigwedge_{i=1}^{l} H^i x \leq c_i \}. \]

The vector \( c \) is called a polyhedral coefficient vector. By choosing the templates it is possible to control both the precision and the geometric complexity of the approximations. However this problem needs a deep investigation and is not a topic of this paper.

The template matrix \( H \) is assumed to be given (which can be composed of the normal directions of regular polyhedra); to over-approximate the image \( \pi(X) \), the polyhedral coefficient vector \( c = (c_1, \ldots, c_l)^T \) can be determined as the solution the following optimization problems:

\[ \forall i \in \{1, \ldots, l\}, c_i = \max(\Sigma_{k=1}^{n} H^i_k \pi_k(x, p)) \]

\[ \text{subj. to } x \in X, p \in P. \]

This polynomial optimization problem is computationally difficult, we propose to use a linear relaxation in order to take advantage of well-developed linear programming techniques [12]. To this end, the Bernstein expansion can be used to compute affine bound functions of polynomials.

The essence of this method, proposed in [6], for computing an affine bound function is to find a hyperplane that is close to all the control points, using linear least squares approximation. For our problem, we need to extend this method to parametric polynomial functions.

A multi-index \( i = (i_1, i_2, \ldots, i_n) \) is a vector of \( n \) non-negative integers. Given two multi-indices \( i \) and \( d \), we write \( i \leq d \) if for all \( j \in \{1, \ldots, n\} \), \( i_j \leq d_j \). The parametric polynomial \( \pi : \mathbb{R}^n \times \mathbb{R}^m \to \mathbb{R}^n \) can be represented using the power base as

\[ \pi(x, p) = \sum_{i \in I_d} a_i(p) x^i \]

where \( a_i \) is a function \( \mathbb{R}^m \to \mathbb{R} \); \( I_d \) is the set of all multi-indices \( i \leq d \), that is

\[ I_d = \{ i \mid i \leq d \}. \]

In this work, we consider only a linear dependence of the coefficients \( a_i(p) \) on the parameters \( p \). The multi-index \( d \) is called the degree of \( \pi \). For \( x = (x_1, \ldots, x_n) \in \mathbb{R}^n \), the \( i^{th} \) Bernstein polynomial of degree \( d \) is defined as follows:

\[ B_{d,i}(x) = \beta_{d_i, i_1}(x_1) \ldots \beta_{d_i, i_n}(x_n) \]

where for a real number \( y \),

\[ \beta_{d_i, i_1}(y) = \left( \frac{d_j}{i_j} \right) y^{i_j}(1 - y)^{d_j - i_j}, \]

One way to approximate the image \( \pi(X) \) is to use special convex polyhedra with fixed geometric form, called template polyhedra [13], [2]. A template is a set of linear functions defined by an \( l \times n \) matrix \( H \) and a real-valued vector \( c \in \mathbb{R}^l \). Let \( H^i \) be the \( i^{th} \) row of the matrix \( H \) and \( H^i_k \) its \( k^{th} \) element, the corresponding template polyhedron is defined by

\[ \langle H, c \rangle = \{ x \mid \bigwedge_{i=1}^{l} H^i x \leq c_i \}. \]

The vector \( c \) is called a polyhedral coefficient vector. By choosing the templates it is possible to control both the precision and the geometric complexity of the approximations. However this problem needs a deep investigation and is not a topic of this paper.

The template matrix \( H \) is assumed to be given (which can be composed of the normal directions of regular polyhedra); to over-approximate the image \( \pi(X) \), the polyhedral coefficient vector \( c = (c_1, \ldots, c_l)^T \) can be determined as the solution the following optimization problems:

\[ \forall i \in \{1, \ldots, l\}, c_i = \max(\Sigma_{k=1}^{n} H^i_k \pi_k(x, p)) \]

\[ \text{subj. to } x \in X, p \in P. \]

This polynomial optimization problem is computationally difficult, we propose to use a linear relaxation in order to take advantage of well-developed linear programming techniques [12]. To this end, the Bernstein expansion can be used to compute affine bound functions of polynomials.

The essence of this method, proposed in [6], for computing an affine bound function is to find a hyperplane that is close to all the control points, using linear least squares approximation. For our problem, we need to extend this method to parametric polynomial functions.

A multi-index \( i = (i_1, i_2, \ldots, i_n) \) is a vector of \( n \) non-negative integers. Given two multi-indices \( i \) and \( d \), we write \( i \leq d \) if for all \( j \in \{1, \ldots, n\} \), \( i_j \leq d_j \). The parametric polynomial \( \pi : \mathbb{R}^n \times \mathbb{R}^m \to \mathbb{R}^n \) can be represented using the power base as

\[ \pi(x, p) = \sum_{i \in I_d} a_i(p) x^i \]

where \( a_i \) is a function \( \mathbb{R}^m \to \mathbb{R} \); \( I_d \) is the set of all multi-indices \( i \leq d \), that is

\[ I_d = \{ i \mid i \leq d \}. \]

In this work, we consider only a linear dependence of the coefficients \( a_i(p) \) on the parameters \( p \). The multi-index \( d \) is called the degree of \( \pi \). For \( x = (x_1, \ldots, x_n) \in \mathbb{R}^n \), the \( i^{th} \) Bernstein polynomial of degree \( d \) is defined as follows:

\[ B_{d,i}(x) = \beta_{d_i, i_1}(x_1) \ldots \beta_{d_i, i_n}(x_n) \]

where for a real number \( y \),

\[ \beta_{d_i, i_1}(y) = \left( \frac{d_j}{i_j} \right) y^{i_j}(1 - y)^{d_j - i_j}, \]
and $\left( \frac{d}{v} \right)$ denotes $\left( \frac{d_1}{v_1}, \frac{d_2}{v_2}, \ldots, \frac{d_n}{v_n} \right)$. Then, for all $x$ inside the unit box $B = [0,1]^n$, the polynomial $\pi$ can be written using the Bernstein expansion as follows:

$$\pi(x) = \sum_{i \in I_d} g_i(p) B_{d,i}(x)$$

where for each $i \in I_d$ the Bernstein coefficient $g_i(p)$ is defined as:

$$g_i(p) = \sum_{j \leq i} \binom{i}{j} a_j(p). \quad (4)$$

Note if all the coefficients $a_j$ depend linearly on $p$, so do the Bernstein coefficients. The Bernstein coefficients allow to derive useful geometric properties, and we now show that they can be used to efficiently compute an affine approximation of the polynomial $\pi$. Since the Bernstein coefficients reflect the form of the image set, we will derive the lower affine function that fits "best" these Bernstein coefficients using linear least squares. Note that in the method proposed in [1] the direction of the bound function is determined for some midpoint in the parameter set and then is shifted down to remain a lower bound for all other parameters $p$. In this work, we improve this method by treating the parameters in a symbolic way.

Let the set of multi-indices be written as $I_d = \{ i^j \mid 1 \leq j \leq n_b \}$, and the set of all control points are denoted similarly. Let $A$ be a matrix of size $n_b \times (n + 1)$ ($n$ is the number of state variables of the dynamical systems in question) such that its elements are defined as follows. For all $1 \leq j \leq n_b$ and $1 \leq k \leq n$, $A^j_k = \frac{1}{d_k}$ and $A^j_{n+1} = 1$. Let $\zeta(p)$ be the solution of the following linear least squares approximation problem:

$$A^T A \zeta(p) = A^T g(p).$$

Then, we can write $\zeta(p) = (A^T A)^{-1} A^T g(p)$ where $(A^T A)^{-1}$ is the Moore Penrose pseudo inverse\textsuperscript{4} of $A$. From $\zeta(p)$, we can define an affine function

$$\tilde{l}(x) = \sum_{k=1}^{n} \zeta_k(p) x_k + \zeta_{n+1}(p)$$

which corresponds to the "median" axis of the convex hull of all the control points. Its distance to a control point $g^j(p)$ is

$$\tilde{l}(\frac{\tilde{y}}{d}) - g^j(p) = \sum_{k=1}^{n} \zeta_k(p) x_k + \zeta_{n+1}(p) - g^j(p).$$

It now remains to shift it downward by the amount:

$$\delta = \max \big\{ \sum_{k=1}^{n} \zeta_k(p) \frac{\tilde{y}}{d} + \zeta_{n+1}(p) - g^j(p) \mid 0 \leq j \leq n_b \land p \in P \big\}$$

where $\frac{\tilde{y}}{d} = (i^j_1/d_1, \ldots, i^j_n/d_n)$. Due to linear dependence of the Bernstein coefficients on $p$, the above optimization problem is a set of linear programs and can thus be solved efficiently. This results in a parametric lower bound function

$$l(x, p) = \tilde{l}(x, p) - \delta,$$

for all $x \in B$.

As mentioned earlier, the methods to compute affine bound functions for polynomials can be applied only when the set $P$ is inside the unit box $B$ anchored at the origin. To extend it to polyhedral domains, we transform the polyhedron to the unit box by two methods: (1) via an (oriented) box approximation, and (2) by rewriting the polynomials using a change of variables. In this paper, we describe only the second method.

The initial polyhedron $X$ can be mapped to the unit box $B$ by a change of variables. We assume that the polyhedron $X$ is bounded and let $V = \{v_1, \ldots, v_l\}$ be the set of its vertices. We first express the coordinates of a point $x \in X$ as a linear combination of the vertices of $X$, that is $x = \sum_{j=1}^{l} \alpha_j v_j = \nu(\alpha_1, \ldots, \alpha_l)$ such that

$$\forall j \in \{1, \ldots, l\} \alpha_j \geq 0 \quad (5)$$

$$\sum_{j=1}^{l} \alpha_j = 1. \quad (6)$$

We then substitute $x$ in $\pi$ with $\nu(\alpha_1, \ldots, \alpha_l)$ to yield a new polynomial in $\alpha_1, \ldots, \alpha_l$. It is not hard to see that a bound function computed for the polynomial $\xi(\tilde{a})$ on this unit box is also a bound function for the original polynomial $\pi$ on the polyhedron $P$.

The above image computation is similar to a number of existing numerical techniques in the use of linear approximation. Its novelty resides in the efficient way of computing linear approximations. Indeed, a common method to approximate a non-linear function by a piecewise linear one, as in the hybridization approach [1], [4] for hybrid systems, requires non-linear optimization. In the hybrid systems verification, polynomial optimization is used to compute barrier certificates [10].

V. EXPERIMENTATION

We have integrated the above image computation method with the tool NBAC and tested the resulting prototype on the controller for LEGO robot trajectory planning.

To prove that the controller assures that the robot follows the line, we verify for a family of black lines that the distance from the centroid of the robot to the line never exceeds some threshold (which is 0.1 for this experimentation). The Lustre program, shown in Appendix, contains three nodes corresponding to Simulink subsystems: controller (that computes the wheel speeds), position (that emulates the evolution of the position of the robot), track (that emulates the output of the light sensors). The assertions in the program specify the constraints of the initial conditions. The execution of the prototype tool allowed to detect a violation of the desired property, as shown by a divergence of the distance from the centroid of the robot to the black line in Figure 1.
VI. Conclusion

In this paper we described an approach to verification of embedded control programs with nonlinear transition functions. An important advantage of the approach is that it can handle practical control systems described in Simulink. We intend to explore more case studies in order to evaluate the prototype tool. As future work, we plan to investigate specific treatments for common operations in a program (such as resets and equalities) that may make the resulting optimization problems become flat.

REFERENCES


APPENDIX

node system(orientation0:real)
returns (ok:bool;
posx,posy,orientation, cd,cg,dist,d,
ud,ug:real)
let
assert(orientation0>-5.0*pi/180.0
and orientation0<5.0*pi/180.0);
assert(y0>=-0.1 and y0<=-0.1);
(ud,ug) = controller(cd,cg,dist);
(posx,posy,orientation) = position(0.0, 0.0, orientation0, 0.0 -> (pre ud), (0.0 -> (pre ug)));
(cd,cg,dist,d) = track(0.0 -> pre posx, 0.0 -> pre posy, 0.0 -> pre orientation);
ok = true -> (pre ok) and (d)>=-0.004 and (d)<=0.004;
tel