REACHABILITY COMPUTATION FOR LINEAR HYBRID SYSTEMS

Gerardo Lafferriere [*] **George J. Pappas** [**]
Sergio Yovine [***]

[*] *Department of Mathematical Sciences*
Portland State University
Portland, OR 97207
gerardo@mth.pdx.edu
[**] *Department of Electrical Engineering and Computer Sciences*
University of California at Berkeley
Berkeley, CA 94720
gpappas@eecs.berkeley.edu
[***] *VERIMAG*
Centre Equation, 2 Avenue de Vignate
38610 Gieres, France
Sergio.Yovine@imag.fr

Abstract: Linear hybrid systems are finite state machines with linear vector fields of the form $\dot{x} = Ax$ in each discrete location. Very recently, the reachability problem for classes of linear hybrid systems was shown to be decidable. In this paper, the decidability result is extended to capture classes of linear hybrid systems where in each location the dynamics are of the form $\dot{x} = Ax + Bu$, for various types of inputs.

Keywords: Reachability, linear systems, finite state machines.

1. INTRODUCTION

Hybrid systems combine discrete event systems with differential equations. One of the most important problems for hybrid systems is the *reachability problem* which asks whether some unsafe region is reachable from an initial region. The reachability problem for hybrid systems is known to be difficult. In order to tackle the complexity of the reachability problem, we adopt a computational point of view.

The important issue for computational algorithms for systems with infinite state spaces is *decidability*. The known decidable classes include timed automata (Alur and Dill, 1994), multirate automata (Alur *et al.*, 1995), and rectangular hybrid are described by Henzinger *et al.* (1995) and the references therein. Recently, it was shown that *the reachability problem is decidable for a class of linear hybrid systems*, that is hybrid systems with linear vector fields of the form $\dot{x} = Ax$ at each discrete location (Lafferriere *et al.*, 1998*b*). This result is significant given the wide applicability of linear systems in control theory.

This decidability result is obtained using new mathematical and computational techniques. In particular, the notion of *o-minimality* from *model theory* is used to define a class of hybrid systems, called *o-minimal hybrid systems*. In (Lafferriere *et al.*, 1998*a*), it is shown that all o-minimal hybrid systems admit finite bisimulations, the first step in showing decidability. For decidability, one needs to make the bisimulation algorithm compu-

ematical logic as the main tool to symbolically represent sets with logic formulas. The main computational tool for set manipulation is *quantifier elimination*. Since quantifier elimination exists for the theory of reals with addition and multiplication (Tarski, 1951), one either finds or transforms into subclasses of o-minimal hybrid system which are definable in this decidable theory. This leads to a decidable class of linear hybrid systems.

In this paper, the above decidability result is extended to computationally solve the reachability problem for hybrid systems with continuous dynamics of the form $\dot{x} = Ax + Bu$ in each discrete location. This requires investigation of not only classes of linear vector fields, but also classes of inputs for which this problem is decidable. In addition, conditions on the interactions between the input and the natural dynamics of the system are also important for this extension.

2. HYBRID SYSTEMS

The class of hybrid systems to be considered is defined as follows.

Definition 1. A *hybrid system* is a tuple $H = (X, X_0, X_F, F, E, I, G, R)$ where

- $X = X_D \times X_C$ is the state space with $X_D = \{q_1, \ldots, q_n\}$ and X_C a manifold.
- $X_0 \subseteq X$ is the set of initial states
- $X_F \subseteq X$ is the set of final states
- $F : X \longrightarrow TX_C$ assigns to each discrete location $q \in X_D$ a vector field $F(q, \cdot)$
- $E \subseteq X_D \times X_D$ is the set of discrete transitions
- $I : X_D \longrightarrow 2^{X_C}$ assigns to each location a set $I(q) \subseteq X_C$ called the invariant.
- $G : E \longrightarrow X_D \times 2^{X_C}$ assigns to $e = (q_1, q_2) \in E$ a guard of the form $\{q_1\} \times U$, $U \subseteq I(q_1)$.
- $R : E \longrightarrow X_D \times 2^{X_C}$ assigns to $e = (q_1, q_2) \in E$ a reset of the form $\{q_2\} \times V$, $V \subseteq I(q_2)$.

The hybrid system H has two types of transitions $\stackrel{e}{\to}$ and $\stackrel{\tau}{\to}$ where $(q, x) \stackrel{e}{\to} (q', x')$ for $e \in E$ iff $(q, x) \in G(e)$ and $(q', x') \in R(e)$, and $(q_1, x_1) \stackrel{\tau}{\to} (q_2, x_2)$ iff $q_1 = q_2$ and there exists $\delta \geq 0$ and a curve $x : [0, \delta] \longrightarrow M$ with $x(0) = x_1$, $x(\delta) = x_2$ and for all $t \in [0, \delta]$ it satisfies $dx/dt = F(q_1, x(t))$ and $x(t) \in I(q_1)$. A trajectory of H is a finite concatenation of transitions. Having defined $\stackrel{\tau}{\to}$ and $\stackrel{e}{\to}$ allows us to formally define $Pre_\tau(P)$ and $Pre_e(P)$ for $e \in E$ and any region $P \subseteq X$ as

$$Pre_\sigma(P) = \{q \in Q \mid \exists p \in P \text{ and } q \stackrel{\sigma}{\to} p\} \quad (1)$$

where σ is either τ or any $e \in E$. The following bisimulation algorithm computes a finite bisimulation of the hybrid system (provided the algorithm terminates). Because of the initialized structure of the discrete transitions, the initial partition $\mathcal{S}_q$ is a partition which is compatible with all relevant sets of the hybrid system associated with each discrete location q.

Algorithm (Bisimulation Algorithm for Hybrid Systems)
Set $X/\sim\; = \bigcup_q \mathcal{S}_q$
for $q \in X_D$
  while $\exists\ P, P' \in \mathcal{S}_q$ such that $\emptyset \neq P \cap Pre_\tau(P') \neq P$
    Set $P_1 = P \cap Pre_\tau(P')$; $P_2 = P \setminus Pre_\tau(P')$
    refine $\mathcal{S}_q = (\mathcal{S}_q \setminus \{P\}) \cup \{P_1, P_2\}$
  end while
end for

There are two issues that need to be resolved. The first is to find classes of hybrid systems for which this algorithm terminates. The second is to find classes of systems for which the algorithm is computable. The termination issue is tackled by the notion of o-minimality from model theory.

3. MATHEMATICAL LOGIC AND MODEL THEORY

This section provides a brief introduction to mathematical logic and model theory. The reader is referred to Chang and Keisler (1977) for more details.

3.1 *Languages and formulas*

A *language* is a set of symbols separated into three groups: relations, functions and constants. The sets $\mathcal{L}_0 = \{<, +, -, 0, 1\}$, $\mathcal{L}_R = \{<, +, -, \cdot, 0, 1\}$, and $\mathcal{L}_{\exp} = \{<, +, -, \cdot, 0, 1, \exp\}$ are examples of languages where $<$ (less than) is the relation, $+$ (plus), $-$ (minus), $\cdot$ (product) and exp (exponentiation) are the functions, and 0 (zero) and 1 (one) are the constants.

Let $\mathcal{V} = \{x, y, z, x_0, x_1, \ldots\}$ be a countable set of *variables*. The set of *terms* of a language is inductively defined as follows. A term θ is a variable, a constant, or $F(\theta_1, \ldots, \theta_m)$, where F is a m-ary function and θ_i, $i = 1, \ldots, m$ are terms. For instance, $x - 2y + 3$ and $x + yz^2 - 1$ are terms of $\mathcal{L}_0$ and $\mathcal{L}_R$, respectively. In other words, terms of $\mathcal{L}_0$ are linear expressions and terms of $\mathcal{L}_R$ are polynomials with integer coefficients. Notice that integers are the only numbers allowed in expressions (they can be obtained by adding up the constant 1).

are terms and R is an n-ary relation. For example, $xy > 0$ and $x^2 + 1 = 0$ are terms of $\mathcal{L}_R$.

The set of *(first-order) formulas* is recursively defined as follows. A formula ϕ is an atomic formula, $\phi_1 \wedge \phi_2$, $\neg\phi_1$, $\forall x : \phi_1$ or $\exists x : \phi_1$, where ϕ_1 and ϕ_2 are formulas, x is a variable, $\wedge$ (conjunction) and $\neg$ (negation) are the boolean connectives, and $\forall$ (for all) and $\exists$ (there exists) are the quantifiers.

Examples of $\mathcal{L}_R$-formulas are $\forall x\ \forall y : xy > 0$, $\exists x : x^2 - 2 = 0$, and $\exists w : xw^2 + yw + z = 0$. The occurrence of a variable in a formula is *free* if it is not inside the scope of a quantifier; otherwise, it is *bound*. For example, x, y, and z are free and w is bound in the last example. We often write $\phi(x_1, \ldots, x_n)$ to indicate that $x_1, \ldots, x_n$ are the free variables of the formula ϕ. A *sentence* of $\mathcal{L}_R$ is a formula with no free variables. The first two examples are sentences.

3.2 *Models*

A *model* of a language consists of a non-empty set S and an interpretation of the relations, functions and constants. For example, $(\mathbb{R}, <, +, -, \cdot, 0, 1)$ and $(\mathbb{Q}, <, +, -, \cdot, 0, 1)$, are *models* of $\mathcal{L}_R$ with the usual meaning of the symbols.

A set $Y \subseteq S^n$ is *definable* in a language if there exists a formula $\phi(x_1, \ldots, x_n)$ such that $Y = \{(a_1, \ldots, a_n) \in S^n \mid \phi(a_1, \ldots, a_n)\}$. For example, over $\mathbb{R}$, the formula $x^2 - 2 = 0$ defines the set $\{\sqrt{2}, -\sqrt{2}\}$. Two formulas $\phi(x_1, \ldots, x_n)$ and $\psi(x_1, \ldots, x_n)$ are *equivalent* in a model, denoted by $\phi \equiv \psi$, if for every assignment $(a_1, \ldots, a_n)$ of $(x_1, \ldots, x_n)$, $\phi(a_1, \ldots, a_n)$ is true if and only if $\psi(a_1, \ldots, a_n)$ is true. Equivalent formulas define the same set.

3.3 *Theories*

A *theory* is a subset of sentences. Any model of a language defines a theory: *the set of all sentences which hold in the model.* By abuse of notation the theory $(\mathbb{R}, <, +, -, \cdot, 0, 1)$ will refer to the collection of formulas of $\mathcal{L}_R$ which are true in the model (and similarly for other languages). Only models over $\mathbb{R}$ will be considered.

Definition 2. The theory of $\mathcal{L}$ is *o-minimal* ("order minimal") if every definable subset of $\mathbb{R}$ is a finite union of points and intervals (possibly unbounded).

Examples of o-minimal theories include

(2) $\mathcal{R}_{\sin} = (\mathbb{R}, <, +, -, \cdot, \sin|_{[-\pi,\pi]}, 0, 1)$ (van den Dries, 1986),
(3) $\mathcal{R}_{\exp} = (\mathbb{R}, <, +, -, \cdot, \exp, 0, 1)$ (Wilkie, 1996), and
(4) $\mathcal{R}_{\sin,\exp} = (\mathbb{R}, <, +, -, \cdot, \sin|_{[-\pi,\pi]}, \exp, 0, 1)$ (van den Dries and Miller, 1994).

Based on this notion a new class of hybrid systems is defined.

Definition 3. A hybrid system $H = (X, X_0, X_F, F, E, I, G, R)$ is said to be *o-minimal* if

- $X_C = \mathbb{R}^n$
- for each $q \in X_D$ the flow of F_q is complete (flow is definable for all time)
- for each $q \in X_D$ all the relevant sets (guards, invariants, initial conditions etc) and the flow of F_q are definable in an o-minimal extension of $(\mathbb{R}, <, +, -, \cdot, 0, 1)$.

Termination of the bisimulation algorithm is guaranteed by the following theorem.

Theorem 4. Every o-minimal hybrid system admits a finite bisimulation. In particular, the bisimulation algorithm, terminates for o-minimal hybrid systems.

PROOF. See (Lafferriere *et al.*, 1998*a*).

4. DECIDABLE LINEAR HYBRID SYSTEMS

Based on the previous result, if one can find or transform into classes of o-minimal hybrid systems for which each step of the bisimulation algorithm is decidable then one obtains a class of hybrid systems for which the following problem is decidable.

Problem 5. (Reachability Problem) Is there a trajectory of H which starts in X_0 and ends in X_F?

This is achieved for a class of linear hybrid systems. For this class, each step of the algorithm can be posed as a quantifier elimination problem in the theory $(\mathbb{R}, +, -, \cdot, <, 0, 1)$ which is known to be decidable (Tarski, 1951).

Theorem 6. Consider an o-minimal hybrid system H where $X_C = \mathbb{R}^n$, all relevant sets are semialgebraic and for all $q \in X_D$, $F(q, x) = Ax$ where A is a matrix with rational entries and A belongs in one of the following classes

- A is nilpotent

- A has purely imaginary eigenvalues of the form $i\omega$, $\omega \in \mathbb{Q}$ and its real Jordan form is block diagonal with 2×2 blocks.

Then the reachability problem for the linear hybrid system H is decidable.

PROOF. See (Lafferriere *et al.*, 1998*b*).

5. EXTENSIONS

The goal of this paper is to extend Theorem 6 to include linear hybrid systems where in each discrete locations the dynamics are of the form

$$\dot{x} = Ax + Bu$$

for various types of inputs u. This requires an extension of the definition of hybrid systems to allow time-varying dynamics.

Definition 7. A hybrid system $H = (X, X_0, X_F, F, E, I, G, R)$ with X, X_0, X_F, E, I, G, and R as in Definition 1 is said to have *time-varying dynamics* if

$$F : X \times \mathbb{R} \to TX_C,$$

that is, for each $q \in X_D$, $F(q, \cdot, \cdot)$ defines a time-varying vector field on X_C.

The main result is the following.

Theorem 8. Consider a hybrid system H with time-varying dynamics where all relevant sets are semialgebraic and for all $q \in X_D$ the dynamics are of the form $\dot{x} = Ax + Bu$ where A,B are matrices with rational entries and in addition one of the following holds:

- A is nilpotent and each entry of u is a polynomial in t
- A is diagonalizable, has real rational eigenvalues, and each entry of u is of the form $e^{\mu t}$ with $\mu \in \mathbb{Q}$ not an eigenvalue of A.
- A has purely imaginary eigenvalues of the form $i\omega$, $\omega \in \mathbb{Q}$, and the entries in the input u are of the form $\sin(\alpha t)$ or $\cos(\alpha t)$ with $\alpha \in \mathbb{Q}$ and $\alpha \neq \pm\omega$.

Then the reachability problem for the hybrid system H is decidable.

PROOF. First an associated (time-invariant) hybrid system as in Definition 1 is constructed, which has an equivalent reachability problem. This is done by introducing time as a new variable x_{n+1}. More precisely, we define the new hybrid system $\widetilde{H} = (\widetilde{X}, \widetilde{X}_0, \widetilde{X}_F, \widetilde{F}, \widetilde{E}, \widetilde{I}, \widetilde{G}, \widetilde{R})$ is defined

$n+1$

$\widetilde{G} : \widetilde{E} \to X_D \times 2^{\mathbb{R}^{n+1}}$, $\widetilde{R} : \widetilde{E} \to X_D \times 2^{\mathbb{R}^{n+1}}$, and, for $\tilde{x} = (x, x_{n+1})^T \in \mathbb{R}^{n+1}$,

$$\widetilde{F}(q, \tilde{x}) = \begin{pmatrix} A_q x + B_q u_q(x_{n+1}) \\ 1 \end{pmatrix}.$$

Let

$$\phi_q(s, t, x_0) = e^{A_q(t-s)} x_0 + \int_s^t e^{A_q(t-\tau)} B_q u_q(\tau)\, d\tau.$$

Then the flow of the vector field $\widetilde{F}_q = \widetilde{F}(q, \cdot)$ is given by

$$\phi_{\widetilde{F}_q}(t, (x_0, s_0)) = (\phi_q(s_0, t, x_0), t + s_0)^T.$$

This shows that if ϕ_q is definable in an o-minimal theory then so is the flow of the corresponding vector field $\widetilde{F}_q$. Moreover, the reachability problem for H is decidable if and only if the reachability problem for $\tilde{H}$ is decidable. Since the original sets are semialgebraic, they are definable in the o-minimal theory $\mathcal{R}$. The following two facts will suffice to conclude the proof:

(1) the flows of the vectors fields $\widetilde{F}_q$ are definable in an o-minimal theory which extends $\mathcal{R}$.
(2) for any two semialgebraic sets V, W, the formula

$$x \in V \wedge \exists y \in W, \exists t, \exists s : \\ t > s \wedge \phi_q(s, t, x) = y \qquad (2)$$

is equivalent to a formula in $\mathcal{R}$.

The first fact together with Theorem 4 shows that the bisimulation algorithm terminates. The second fact shows that the statement $x \in V \cap Pre_\tau(W)$ can be effectively decided. Therefore the two facts complete the proof.

In order to prove both facts Formula 2 must be analyzed for each type of possible control. From now on the subscript q will be dropped to simplify notation.

If A is nilpotent then each entry of e^{At} is a polynomial in t. Therefore, each entry of $\phi(s, t, x)$ is a polynomial in t and s and hence definable in $\mathcal{R}$.

If A is diagonalizable with real eigenvalues then each entry of e^{At} is a linear combination of the functions $e^{\lambda_i t}$ with λ_i an eigenvalue of A. Since each entry of u is of the form $e^{\mu_j t}$, the entries of $e^{-A\tau} Bu(\tau)$ are linear combinations of $e^{(\mu_j - \lambda_i)t}$. By the assumption $\mu_j - \lambda_i \neq 0$ it results (after integration) that the entries of $\phi(s, t, x)$ are also linear combinations of the functions $e^{\lambda_i t}$, $e^{(\mu_j - \lambda_i)t}$, $e^{\lambda_i s}$, and $e^{(\mu_j - \lambda_i)s}$, and therefore definable in $\mathcal{R}_{\exp}$.

Finally, in the case of imaginary eigenvalues as de-

At

Because of the form of the input u in this case, the entries of $e^{-A\tau}Bu(\tau)$ are linear combinations of the products $\cos(\alpha_j t)\cos(\omega_k t)$, $\cos(\alpha_j t)\sin(\omega_k t)$, $\sin(\alpha_j t)\cos(\omega_k t)$, and $\sin(\alpha_j t)\sin(\omega_k t)$. Using standard product formulas these can be rewritten as linear combinations of $\sin((\alpha_j \pm \omega_k)t)$ and $\cos((\alpha_j \pm \omega_k)t)$. Since $\alpha_j \neq \pm\omega_k$, integration again results in linear combinations of $\sin((\alpha_j \pm \omega_k)t)$ and $\cos((\alpha_j \pm \omega_k)t)$ and therefore $\phi(s,t,x)$ is definable in $\mathcal{R}_{\sin}$.

The analysis above implies that all relevant sets and relevant flows of the hybrid system are definable in the o-minimal theory $\mathcal{R}_{\sin,\exp}$. This concludes the proof of the first fact.

To prove the second fact it suffices to show that the formula $\eta(x,y) \equiv \exists s,t, t>s : \phi(s,t,x) = y$ is equivalent to a formula in $\mathcal{R}$.

In the case of nilpotent A and polynomial inputs, the formula already is in $\mathcal{R}$. Assume now that A is diagonal and the entries of u are exponential functions as indicated above. The formula $\eta(x,y)$ can be written more explicitly (entry by entry) as

$$y_i = \sum_k \psi_{ik}(x) e^{\mu_k t} e^{-\mu_k s} + \sum_l a_{il}(e^{\nu_l t} - e^{\nu_l s})$$

for some linear functions ψ_{ik} of x and rational constants a_{il}, μ_k, and ν_l.

Let d_k denote the least common denominator of the μ_k and c_l the least common denominator of the ν_l (both assumed positive). Set $d_0 = \prod d_k$, $c_0 = \prod c_l$, $p_k = \mu_k d_0$, and $q_l = \nu_l c_0$. Then $\eta(x,y)$ is equivalent to

$$\begin{aligned} &\exists \tilde{s}\ \exists \tilde{t} : d_0 \tilde{t} > c_0 \tilde{s} \ \wedge \\ &y_i = \sum_k \psi_{ik}(x) e^{p_k \tilde{t}} e^{-p_k \tilde{s}} + \sum_l a_{il}(e^{q_l \tilde{t}} - e^{q_l \tilde{s}}) \end{aligned}$$

Using the substitutions $z = e^{\tilde{t}}$, $w = e^{-\tilde{s}}$, the previous formula is equivalent to

$$\begin{aligned} &\exists z\ \exists w : z > 0, w > 0 \ \wedge \ w^{c_0} z^{d_0} > 1 \\ &y_i = \sum_k \psi_{ik}(x) z^{p_k} w^{p_k} + \sum_l a_{il}(z^{q_l} - w^{-q_l}) \end{aligned}$$

which can be rewritten in $\mathcal{R}$ by multiplying through by $\prod w^{q_l}$.

In the case of purely imaginary eigenvalues the formula for $\phi(s,t,x)$ can be rewritten in terms of polynomials in $\cos(t)$, $\sin(t)$, $\cos(s)$, and $\sin(s)$ with rational coefficients. The substitutions $z = \cos(t)$, $w = \sin(t)$ and the identity $z^2 + w^2 = 1$ (and similarly for the s variable) lead to the equivalent formula in $\mathcal{R}$. This concludes the proof of the theorem.

the functions of the corresponding type: exponentials in case of real eigenvalues, and sinusoidal in the case of imaginary eigenvalues. In all cases the same "resonance" restrictions apply on the parameters μ and α.

6. CONCLUSIONS

This paper extends the decidability results for linear hybrid systems to include systems with inputs of a special type. The techniques used involved the notion of a bisimulation and elimination of quantifiers.

Areas for future research include extensions of Theorem 8 to differential inclusions of the form

$$\dot{x} = Ax + Bu \qquad u \in U$$

where the input set is a rectangular set. This requires certain monotonicity conditions on the solution integral, which will be valid on a smaller class of linear hybrid systems.

Acknowledgment: This research is supported by the Army Research Office under grants DAAH 04-95-1-0588 and DAAH 04-96-1-0341.

7. REFERENCES

Alur, R. and D.L. Dill (1994). A theory of timed automata. *Theoretical Computer Science* **126**, 183–235.

Alur, R., C. Coucoubetis, N. Halbwachs, T.A. Henzinger, P.H. Ho, X. Nicolin, A. Olivero, J. Sifakis and S. Yovine (1995). The algorithmic analysis of hybrid systems. *Theoretical Computer Science* **138**, 3–34.

Chang, C. C. and H. J. Keisler (1977). *Model Theory.* second ed.. North-Holland.

Henzinger, T.A., P.W. Kopke, A. Puri and P. Varaiya (1995). What's decidable about hybrid automata?. In: *Proceedings of the 27th Annual Symposium on Theory of Computing.* ACM Press. pp. 373–382.

Lafferriere, G., G.J. Pappas and S. Sastry (1998*a*). O-minimal hybrid systems. Technical Report UCB/ERL M98/29. University of California at Berkeley. Berkeley, CA. Submitted to *Mathematics of Control, Signals, and Systems.*

Lafferriere, G., G.J. Pappas and S. Yovine (1998*b*). Decidable hybrid systems. Technical report. University of California at Berkeley. Berkeley, CA. To appear in *Hybrid Systems: Computation and Control, 1999.*

Puri, A. and P. Varaiya (1994). Decidability of hybrid systems with rectangular differential

Tarski, A. (1951). *A decision method for elementary algebra and geometry.* second ed.. University of California Press.

van den Dries, L. (1986). A generalization of the tarski-siedenberg theorem, and some nondefinability results. *Bulletin of the American Mathematical Society* **15**(2), 189–193.

van den Dries, L. and C. Miller (1994). On the real exponential field with restricted analytic functions. *Israel Journal of Mathematics* **85**, 19–56.

Wilkie, A. J. (1996). Model completeness results for expansions of the ordered field of real numbers by restricted pfaffian functions and the exponential function. *Journal of the American Mathematical Society* **9**(4), 1051–1094.