Università degli studi di Udine

Hybrid automata, reachability, and Systems Biology

Publisher copyright

(Article begins on next page)

13 May 2024

Advantages and Limitations of Hybrid Automata in Systems Biology[☆]

Dario Campagna

Dept. of Mathematics and Computer Science
University of Perugia
Via Vanvitelli 1, 06123 Perugia, Italy

Carla Piazza

Dept. of Mathematics and Computer Science
University of Udine
Via delle Scienze 206, 33100 Udine, Italy

Abstract

We consider the reachability problem on semi-algebraic hybrid automata. In particular, we deal with the effective cost that has to be afforded to solve reachability through first-order satisfiability.

The analysis we perform with some existing tools shows that even simple examples cannot be efficiently solved. We need approximations to reduce the number of variables in our formulae: this is the main source of time computation growth. We study standard approximation methods based on Taylor polynomials and ad-hoc strategies to solve the problem, we implement them within a software called SAHA-Tool and we show their effectiveness on the Repressilator and Delta-Notch protein signaling case studies.

Key words: Hybrid Automata, Reachability, Semi-Algebraic First-Order Formulæ

Introduction

Since their introduction (see, e.g., [1]), hybrid automata have initiated a new tradition, promising powerful tools for modeling and reasoning about complex engineered or natural systems (see, e.g., [2, 3]).

Intuitively, a hybrid automaton consists of a finite graph, whose nodes are called *locations*, together with a set of continuous variables which evolve according to continuous laws, called *dynamics*, characterising each discrete location. The continuous evolution of the hybrid automaton may change from location to location. Moreover, each location is characterised by an *invariant* condition which defines the allowed values for the continuous variables inside the location. Finally, each graph's *edge* is labelled by both an *activation* condition and a *reset* map. The edge can be crossed only if the continuous variables satisfy the activation condition and after crossing it the continuous variables are set accordingly to the reset map. The double nature, both discrete and continuous, of hybrid automata make them particularly suitable in the modeling of systems exhibiting a mixed behaviour which cannot be characterised in a proper way using either discrete or continuous formalisms.

In this context, one of the basic problems is the *reachability* one which requires to decide whether it is possible to move from a state (a pair consisting of a location together with a set of values for the continuous variables) to another.

Unfortunately, the flexibility and expressive power of hybrid automata soon lead to undecidability and complexity results [4] which cast doubts on their suitability as a general tool that can be algorithmized and efficiently implemented.

In order to control both undecidability and complexity one can either impose syntactic conditions and concentrate on classes of hybrid automata or define semantic approximation techniques.

In [5] the class of *semi-algebraic hybrid automata* has been introduced. The invariants, dynamics, activations, and resets of semi-algebraic automata have to be first-order formulæ over the theory of $(\mathbb{R}, 0, 1, +, *, <)$. On the one hand, such formulæ are decidable [6] and tools such as QEPCAD B [7] can be used to manage them. On the other hand, Taylor polynomials allow to use semi-algebraic formulæ to approximate with arbitrary precision any smooth function. As a consequence of the expressive power of semi-algebraic hybrid automata, the undecidability of the reachability problem for such class can be proved [8]. In particular, in this case, undecidability is a consequence

2

of the fact that we cannot a-priori bound the number of edges we need to cross. Hence, we can see "the glass half full" saying that *bounded* (w.r.t. edge crossing) reachability is computable. Unfortunately, as noticed in [9] such computation results to be too time/space consuming due to the high computational complexity of semi-algebraic decomposition.

In this paper we start from the considerations presented in [9] concerning the effectiveness of bounded reachability computation on semi-algebraic hybrid automata and we show on some examples which kind of approximations are necessary to keep complexity under control. As done in [9] we may distinguish space and time discretizations in our work. As far as space discretizations are concerned, instead of implementing an ad-hoc algorithm, we try to exploit tools which allow approximate computations over the reals such as RSOLVER [10] and ECL^{i}PS^e [11]. Unfortunately, this is not enough: space approximations which *separate* the continuous variables are necessary. We notice that time discretization and Taylor polynomials are essential ingredients in our approach.

The paper is organized as follows. In Section 1 we quickly overview the state of the art. Some basic notions about semi-algebraic hybrid automata and reachability find place in Section 2, while Section 3 is the core part of our work. In Section 4 we present SAHA-TOOL, a software based on our approximated methods. In Section 5 we apply our analysis to the Repressilator and the Delta-Notch protein signaling case studies. Some conclusions are drawn in Section 6.

1. Related Works

As mentioned in the introduction, we can control undecidability and complexity on hybrid automata in two ways: imposing syntactic constraints which limit the expressive power or introducing semantic approximation techniques.

In [12] Alur et al. introduced *multirate automata* as an extensions of *timed automata* [13]. Such hybrid automata are characterised by resets which are either identity or constant function zero. Moreover, their continuous variables evolve like clocks with rational rates. In the same work it has been proved that the reachability problem over multirate automata is not decidable in general. However, imposing a restriction on dynamics called *simplicity condition*, decidability for reachability problem and finite bisimulation are proved. Puri and Varaiya in [14] introduced *rectangular hybrid automata* whose dynamics

3

can be characterised by a differential inclusion. They showed that, under a condition called *initialized condition*, reachability can be decided. Lafferriere, Pappas and Sastry introduced *o-minimal hybrid automata* in [15]. Such class of hybrid automata guarantee finite bisimulation quotient imposing both constant reset condition to all the edges and a unique o-minimal dynamic from each state. In [16] it has been proved that reachability is still decidable on semi-algebraic o-minimal automata when the conditions on the dynamics are relaxed allowing many possible continuous evolutions. Unfortunately, all the above mentioned classes have restrictions on both dynamics and resets and thus they are not suitable to verify properties of many interesting hybrid systems.

As far as approximation techniques are concerned, in [17] Halbwachs et al. suggested convex approximations as a way to verify linear hybrid systems, Dang and Maler proposed to verify hybrid automaton properties via face lifting in [18], Chutinan and Krogh showed in [19] how evolutions of polyhedral-invariant hybrid automata can be approximated using polyhedra, Asarin et al. gave in [20] a technique to approximate reachability analysis of piecewise-linear dynamical systems, Kurzhanski and Varaiya introduced ellipsoidal techniques in [21], Alur et al. proposed in [22] *predicate abstraction* as a technique to perform reachability analysis. Many tools, based on such techniques, have been developed in the last years. In particular, we can recall *HyTech* [23], *d/dt* [24], *Checkmate* [25], *UPPAAL* [26], and *KRONOS* [27]. Unfortunately, all these approximation methods and tools are again defined on restricted classes of hybrid automata. Such classes are clearly larger than the classes on which decidability has been proved. However, it is still necessary to check that the model satisfies all the required conditions before the method can be applied.

Semi-algebraic hybrid automata introduced in [5] intrinsically combine syntactic restrictions and semantics approximations. On the one hand Taylor polynomials can be used to approximate a large class of hybrid automata with semi-algebraic ones. In [28] Lanotte and Tini proposed an approximation technique for hybrid automata that exploits Taylor polynomials to obtain from an hybrid automaton H a polynomial hybrid automaton H' that over-approximate H. On the other hand, cylindrical algebraic decomposition (CAD) algorithms (see, e.g., [29, 30, 31, 32]) can be used to reason on semi-algebraic hybrid automata. Such considerations are also at the basis of the abstractions and analysis techniques presented in [3].

The tool QEPCAD B [7] efficiently implements Collins' CAD-based al-

gorithm [29] for quantifier elimination, transforming any given first-order semi-algebraic formula into an equivalent quantifier-free one and it can easily become the engine of a step-by-step reachability algorithm for semi-algebraic automata. Unfortunately, the computational cost is still too high. QEPCAD B is not the only tool which can be used to manage constraints over the reals. In particular, we recall: RSOLVER [10], a program for solving quantified inequality constraints over the reals based on a *branch-and-prune* algorithm; ECL^{i}PS^e [11], a software system for the development and deployment of constraint programming applications that contains a general interval propagation solver which can be used to solve problems over both integer and real variables; REDLOG [33], a package that extends the computer algebra system REDUCE to a system that provides algorithms for the symbolic manipulation of first-order formulæ with some syntactic restrictions on the quantified variables; CLP(RL) [34], a constraint solving system, implemented on top the computer logic system REDLOG, where the admissible constraints are arbitrary first-order formulæ.

2. Reachability in Semi-Algebraic Hybrid Automata

In this section we introduce the standard syntax and semantics of hybrid automata and describe the reachability problem on semi-algebraic hybrid automata.

We start with some notations and conventions we use on hybrid automata. Capital letters Z_1, Z_2,..., Z_m, Z_1',..., Z_m',..., denote variables ranging over $\mathbb{R}$. Analogously, Z denotes the vector of variables $\langle Z_1, \ldots, Z_d \rangle$ and Z' denotes the vector $\langle Z_1', \ldots, Z_d' \rangle$. The temporal variables $T, T', T'', \ldots$ model time and range over $\mathbb{R}_{\geq 0}$. We use the small letters $p, q, r, s, \ldots$ to denote d-dimensional vectors of real numbers. Occasionally, we may use the notation $\varphi[X_1, \ldots, X_m]$ to stress the fact that the set of free variables of the first-order formula φ is included in the set of variables $\{X_1, \ldots, X_m\}$. By extension, if $\{Z_1, \ldots, Z_n\}$ is a set of variable vectors, $\varphi[Z_1, \ldots, Z_n]$ indicates that the free variables of φ are included in the set of components of $Z_1, \ldots, Z_n$. Moreover, given a formula $\varphi[Z_1, \ldots, Z_i, \ldots, Z_n]$ and a vector p of the same dimension as the variable vector Z_i, the formula obtained by component-wise substitution of Z_i with p is denoted by $\varphi[Z_1, \ldots, Z_{i-1}, p, Z_{i+1}, \ldots, Z_n]$. When in φ the only free variables are the components of Z_i, after the substitution we can determine the truth value of $\varphi[p]$.

Hybrid automata have a mixed discrete and continuous behaviour. The discrete component is represented by a graph, while the continuous one is given as a set of continuous variables. For each node of the discrete graph we have an invariant condition and a dynamic law over the continuous variables. The dynamic law may depend on the initial conditions, i.e., on the values of the continuous variables at the beginning of the evolution in the state. The jumps from one discrete state to another are regulated by activation and reset conditions on the continuous variables.

Definition 1 (Hybrid Automata - Syntax). A *hybrid automaton* $H = (Z, Z', \mathcal{V}, \mathcal{E}, Inv, Dyn, Act, Res)$ of dimension d consists of the following components:

1. $Z = \langle Z_1, \ldots, Z_d \rangle$ and $Z' = \langle Z_1', \ldots, Z_d' \rangle$ are two vectors of variables ranging over the reals $\mathbb{R}$;
2. $\langle \mathcal{V}, \mathcal{E} \rangle$ is a graph. Each element of $\mathcal{V}$ will be dubbed *location*.
3. Each vertex $v \in \mathcal{V}$ is labeled by the formulæ $Inv(v)[Z]$ and
   $Dyn(v)[Z, Z', T] \equiv Z' = f_v(Z, T)$, where $f_v : \mathbb{R}^d \times \mathbb{R}_{\geq 0} \longrightarrow \mathbb{R}^d$;
4. Each edge $e \in \mathcal{E}$ is labeled by the formulæ $Act(e)[Z]$ and $Res(e)[Z, Z']$.

The semantics of hybrid automata regulates the time evolution of the continuous variables.

Definition 2 (Hybrid Automata - Semantics). A *state* ℓ of H is a pair $\langle v, r \rangle$, where $v \in \mathcal{V}$ is a location and $r = \langle r_1, \ldots, r_d \rangle \in \mathbb{R}^{d(H)}$ is an assignment of values for the variables of Z. A state $\langle v, r \rangle$ is said to be *admissible* if $Inv(v)[r]$ is true.

The *continuous reachability transition relation* $\xrightarrow{t}_C$, with $t > 0$ is the transition elapsed time, between admissible states is defined as follows:
$\langle v, r \rangle \xrightarrow{t}_C \langle v, s \rangle$ iff it holds that $s = f_v(r, t)$, and for each $t' \in [0, t]$ the formula $Inv(v)[f_v(r, t')]$ is true.

The *discrete reachability transition relation* $\xrightarrow{e}_D$ between admissible states is defined as follows:
$\langle v, r \rangle \xrightarrow{e}_D \langle u, s \rangle$ iff both $Act(e)[r]$ and $Res(e)[r, s]$ are true.

We use the notation $\ell \to \ell'$ to denote that either $\ell \xrightarrow{t}_C \ell'$ or $\ell \xrightarrow{e}_D \ell'$, for some $t \in \mathbb{R}_{\geq 0}$, $e \in \mathcal{E}$.

A trace is a sequence of continuous and discrete transitions. A point s is reachable from a point r if there is a trace starting from r and ending in s.

Definition 3 (Hybrid Automata - Reachability). A *trace* of H is a sequence of admissible states $[\ell_0, \ell_1, \ldots, \ell_i, \ldots, \ell_n]$ such that $\ell_{i-1} \to \ell_i$ holds for each $1 \leq i \leq n$.

The automaton H *reaches* a point $s \in \mathbb{R}^d$ (in time t) from a point $r \in \mathbb{R}^d$ if there exists a trace $tr = [\ell_0, \ldots, \ell_n]$ of H such that $\ell_0 = \langle v, r \rangle$ and $\ell_n = \langle u, s \rangle$, for some $v, u \in \mathcal{V}$ (and t is the sum of the continuous transitions elapsed times). In such a case, we also say that s is *reachable* from r in H.

A path *ph* over a graph G is a sequence $[v_0, \ldots, v_n]$ of nodes of G such that for each $1 \leq i \leq n$ there is an edge from v_{i-1} to v_i. Given a hybrid automaton H and trace, tr, of H, a *corresponding path* of tr is a path *ph* obtained by considering the discrete transitions occurring in tr.

We are interested in the reachability problem for hybrid automata, to be specific, given a hybrid automaton H, an initial set of points $I \subseteq \mathbb{R}^d$, and a final set of points $F \subseteq \mathbb{R}^d$ we wish to decide whether there exists a point in I from which a point in F is reachable.

An interesting class of hybrid automata is the class of *semi-algebraic hybrid automata* [5].

Definition 4 (Semi-Algebraic Automata). A hybrid automaton is *semi-algebraic* if $Dyn(v)$, $Inv(v)$, $Act(e)$, and $Res(e)$ are formulæ belonging to the first-order theory of $(\mathbb{R}, 0, 1, +, *, <)$ [6], also known as the theory of semi-algebraic sets.

Moreover, we say that H is *continuous* if $\forall v \in \mathcal{V}$ $f_v(Z, T)$ is continuous on $\mathbb{R}^d \times \mathbb{R}_{\geq 0}$ and $f_v(r, 0) = r$, for each $r \in \mathbb{R}^d$.

In the rest of this paper we concentrate on continuous semi-algebraic hybrid automata, avoiding all the technical problems concerning the existence, uniqueness and continuity of dynamics (see [16] for more details).

The reachability problem for such class of automata is semi-decidable and it can be reduced to the satisfiability of a numerable disjunction of formulæ of the form $Reach(ph)[Z, Z']$ [16]. In particular, if H is a semi-algebraic automaton, then $q \in \mathbb{R}^d$ is reachable from $p \in \mathbb{R}^d$ in H through a trace whose corresponding path is *ph* if and only if the formula $Reach(ph)[p, q]$ holds. Unfortunately, as proved in [8], the reachability problem for semi-algebraic automata remains undecidable even if we consider computational models over the reals.

Now let us have a closer look at the first-order formulæ involved in the reachability computation. Inside a discrete location v the following formula expresses that Z reaches Z':

$$Reach(v)[Z, Z'] \quad \equiv \quad Inv(v)[Z] \wedge \exists T \geq 0(Z' = f_v(Z, T) \wedge$$
$$\forall 0 \leq T' \leq T(Inv(v)[f_v(Z, T')]))$$

On the other hand, when we cross an edge $\langle v, u \rangle$ we have to consider the formula:

$$Reach(\langle v, u \rangle)[Z, Z'] \quad \equiv \quad Inv(v)[Z] \wedge Act(\langle v, u \rangle)[Z] \wedge$$
$$Res(\langle v, u \rangle)[Z, Z'] \wedge Inv(u)[Z']$$

Combining the above formulæ, for each path ph we can easily construct the formula $Reach(ph)[Z, Z']$. For instance if we have the path $ph = [v, u]$, then:

$$Reach([v, u])[Z, Z'] \quad \equiv \quad \exists Z'', Z'''(Reach(v)[Z, Z''] \wedge$$
$$Reach(\langle v, u \rangle)[Z'', Z'''] \wedge$$
$$Reach(u)[Z''', Z'])$$

Example 1. Let $H_1 = (Z,\ Z',\ \mathcal{V},\ \mathcal{E},\ Inv,\ Dyn,\ Act,\ Res)$ where:

- $Z,\ Z'$ are variables over $\mathbb{R}$,

- $\mathcal{V} = \{v, u\}$ and $\mathcal{E} = \{e\}$, where e goes from v to u,

- $Inv(v)[Z] \equiv 1 \leq Z \leq 10$ and $Inv(u)[Z] \equiv 10 \leq Z \leq 20$,

- $Dyn(v)[Z, Z', T] \equiv Z' = Z + (2Z^2 + Z)T$ and
  $Dyn(u)[Z, Z', T] \equiv Z' = Z + (3Z^2 + Z)T$,

- $Act(e)[Z] \equiv Z = 10$,

- $Res(e)[Z, Z'] \equiv Z' = Z$.

The formula for the path $ph = [v, u]$ is the following:

$$Reach([v, u])[Z, Z'] \quad \equiv \quad \exists Z'', Z'''\Big(Inv(v)[Z] \wedge$$
$$\exists T \geq 0\big(Z'' = Z + (2Z^2 + Z)T \wedge$$
$$\forall 0 \leq T' \leq T(Inv(v)[Z + (2Z^2 + Z)T'])\big) \wedge$$
$$Inv(v)[Z''] \wedge Act(e)[Z''] \wedge Res(e)[Z'', Z'''] \wedge$$
$$Inv(u)[Z'''] \wedge \exists T'' \geq 0\big(Z' = Z''' + (3Z'''^2 +$$
$$Z''')T'' \wedge \forall 0 \leq T''' \leq T''$$
$$(Inv(u)[Z''' + (3Z'''^2 + Z''')T'''])\big)\Big)$$

3. Solving the Reachability Problem

In this section we describe some approximation methods for the reachability problem on semi-algebraic hybrid automata. All the computations have been performed on a Dual Core AMD Opteron$^{\text{TM}}$ Processor 275, 2205.042 MHz with 4 GB RAM, running CentOS.

The complexity of the reachability formulæ presented in Section 2 increases with the length of the discrete path. In particular, we can notice that the degree of the involved polynomials and the quantifier alternation remains bounded, while the number of variables linearly increases.

Since the first-order theory of $(\mathbb{R}, 0, 1, +, *, <)$ admits the quantifier elimination, we can try to bound the number of variables in each formulæ. When we apply the quantifier elimination procedure to $Reach(ph)[Z, Z']$ we obtain an equivalent first-order formula $\phi[Z, Z']$ involving only the variables Z and Z'. If we now add a step to the path $ph = [v_1, \ldots, v_n]$, i.e., we consider the path $ph' = [v_1, \ldots, v_n, v_{n+1}]$, we only have to apply quantifier elimination to the formula:

$$\exists Z'', Z'''(\phi[Z, Z''] \wedge Reach(\langle v_n, v_{n+1} \rangle)[Z'', Z'''] \wedge Reach(v_{n+1})[Z''', Z'])$$

Proceeding in this way, it seems that we can keep under control the complexity of our method. Unfortunately, if we try to apply it, exploiting QEPCAD B to obtain quantifier free formulæ at each step, we cannot go far enough, as shown by the following example.

Example 2. Consider the following hybrid automaton.
$H_2 = (Z, Z', \mathcal{V}, \mathcal{E}, Inv, Dyn, Act, Res)$ where:

- $Z = \langle Z_1, Z_2 \rangle$ and $Z' = \langle Z_1', Z_2' \rangle$, where Z_1, Z_2, Z_1', Z_2' variables over $\mathbb{R}$,

- $\mathcal{V} = \{v, u\}$ and $\mathcal{E} = \{e\}$, where e goes from v to u,

- $Inv(v)[Z] \equiv 1 \leq Z_1 \leq 10 \wedge 1 \leq Z_2 \leq 10$ and
  $Inv(u)[Z] \equiv 10 \leq Z_1 \leq 20 \wedge 10 \leq Z_2 \leq 20$,

- $Dyn(v)[Z, Z', T] \equiv Z_1' = Z_1 + (2Z_1^2 + Z_1)T \wedge Z_2' = Z_2 + (2Z_2^2 + Z_2)T$
  and
  $Dyn(u)[Z, Z', T] \equiv Z_1' = Z_1 + (3Z_1^2 + Z_1)T \wedge Z_2' = Z_2 + (3Z_2^2 + Z_2)T$,

- $Act(e)[Z] \equiv Z_1 = 10 \wedge Z_2 = 10$,

9

- $Res(e)[Z, Z'] \equiv Z_1{}' = Z_1 \wedge Z_2{}' = Z_2$.

Suppose we want to apply the method described above with $ph = [v, u]$. First, we use QEPCAD B to compute a quantifier free formula $\phi[Z, Z']$ equivalent to the formula $Reach(v)[Z, Z']$. Then we construct the formula:

$$\exists Z'', Z'''(\phi[Z, Z''] \wedge Reach(\langle v, u \rangle)[Z'', Z'''] \wedge Reach(u)[Z''', Z'])$$

When we try to compute an equivalent quantifier free formula with QEPCAD B we find out that we cannot obtain any result within 20 minutes of CPU time.

Using this method we are able to limit the number of variables in our formulæ, but we have an increasing number of polynomials and constraints in the computed quantifier free formulæ. This is one of the problems of this method, since the complexity of the new constructed formulæ strongly depends on the number of polynomials and constraints occurring in computed quantifier free formulæ. Another problem of the method is that QEPCAD B could not give any result in reasonable time when used on formulæ of the form $Reach(v)[Z, Z']$, i.e., the reachability problem inside a location could be already too complex.

At this point the only possibility we have is that of introducing approximations. A first approximated approach to the reachability problem consists in the application of the above method exploiting RSOLVER instead of QEPCAD B. Acting in this way we hope to solve both the problems mentioned in Example 2. Unfortunately, this approach is less effective than the previous one.

Example 3. Consider the hybrid automaton H_2 of example 2. RSOLVER on the formula $Reach(v)[Z, Z']$ gives the following result:
```
True, volume ~[ 0., 0.]
False, volume ~[ 5905.08179397, 5905.08179397]
⋮
Unknown:
⋮
```

Since the `True` set is empty we do not know which values of Z and Z' satisfy the formula $Reach(v)[Z, Z']$ and we cannot procede with the next step.

RSOLVER on formulæ of the form $Reach(v)[Z, Z']$ gives results that are too approximated for being used. However, we can use it to try to solve the problem related to the number of polynomials and constraints appearing in computed quantifier free formulæ. To do this we apply the previous method exploiting QEPCAD B with the add of an intermediate step that involves the use of RSOLVER.

More precisely, consider the path $ph' = [v_1, \ldots, v_n, v_{n+1}]$ and suppose we have already computed a quantifier free formula $\phi[Z, Z']$ equivalent to $Reach(ph)[Z, Z']$, where $ph = [v_1, \ldots, v_n]$. Using RSOLVER we compute an approximation of the set of values for Z and Z' that satisfy $\phi[Z, Z']$, then we construct a first-order formula $\gamma[Z, Z']$ defining such approximation. Finally, we apply the quantifier elimination procedure to the formula:

$$\exists Z'', Z'''(\gamma[Z, Z''] \wedge Reach(\langle v_n, v_{n+1}\rangle)[Z'', Z'''] \wedge Reach(v_{n+1})[Z''', Z'])$$

It is still not enough, as shown by the following example.

Example 4. Consider again the hybrid automaton H_2 of Example 2. Let $\phi[Z, Z']$ be the quantifier free formula equivalent to $Reach(v)[Z, Z']$ computed by QEPCAD B. RSOLVER on the formula $\phi[Z, Z']$ gives the following result:
```
True, volume ~[ 0., 0.]
False, volume ~[ 5904.9114008, 5904.91140081]
```
⋮
```
Unknown:
```
⋮

As in Example 3 we obtain an empty `True` set and we cannot procede with the next step.

Another approximated approach that we can consider consists in the application of this last described method using ECL^iPS^e instead of RSOLVER to compute the set of values that satisfy a quantifier free formula obtained with QEPCAD B.

Given a quantifier free formula we can define a constraint satisfaction problem with constraint on reals that can be solved by ECL^iPS^e through constraint propagation and search techniques. An answer to a problem on reals is called conditional solution. The number of conditional solutions returned vary according to the level of precision in the search procedure. For instance, given the problem defined from the formula $\phi[Z, Z']$ of Example 4

and using the predicate `locate/2` with final precision 1.0 ECL^{i}PS^e returns 51 answers, if we reduce the final precision to 0.1 we obtain more than 102 answers. Even if we find a way to use the values computed by ECL^{i}PS^e to construct the formula for the successive step, the method would not be effective, because we still have the problem that QEPCAD B could not give any result when used on formulæ of the form $Reach(v)[Z, Z']$.

All the above discussed methods share one problem: the high cost in terms of computation time that has to be afforded to compute a quantifier free formula from a formula of the form $Reach(v)[Z, Z']$ using QEPCAD B. We have to find an approximation strategy to solve this problem in order to obtain an effective method to compute approximated solutions for the reachability problem.

To achieve this goal we studied a method to over-approximated the set of values reachable inside a discrete location of an automaton with independent dynamics.

Definition 5 (Hybrid Automata with Independent Dynamics). Let H be a continuous semi-algebraic hybrid automaton, let $Z = \langle Z_1, \ldots, Z_d \rangle$ and $Z' = \langle Z_1', \ldots, Z_d' \rangle$. H has independent dynamics if $\forall v \in \mathcal{V}$ the formula $Dyn(v)[Z, Z', T]$ is of the form:

$$Z_1' = f_{v,1}(Z_1, T) \wedge Z_2' = f_{v,2}(Z_2, T) \wedge \ldots \wedge Z_d' = f_{v,d}(Z_d, T)$$

Example 5. The automaton H_2 of Example 2 is a continuous semi-algebraic hybrid automaton with independent dynamics.

Given a discrete location v of an automaton with independent dynamics, we over-approximate the set of values Z' that can be reached inside v after time δ from values Z satisfying $Inv(v)[Z]$ applying the quantifier elimination procedure to the following formula

$$ReachApprox(v)[Z'] \equiv \exists Z(Inv(v)[Z] \wedge Z' = f_v(Z, \delta) \wedge Inv(v)[Z'])$$

This is an over-approximation of the sets of points reachable at time δ, since we did not check that at each time T' between 0 and δ the invariant is satisfied by $f_v(Z, T')$. Notice also that we can replace the condition $Inv(v)[Z]$ with a stronger one if we are interested in a subset of starting points. Using this formula many times we can compute an over-approximation of all the values Z' that can be reached with δ-time steps from values Z satisfying $Inv(v)[Z]$.

After each step of duration δ we can consider the following formula to check if the edge $\langle v, u \rangle$ can be crossed:

$$ReachApprox(\langle v, u \rangle)[Z'] \equiv \exists Z(\phi[Z] \wedge Act(\langle v, u \rangle)[Z] \wedge \\ Res(\langle v, u \rangle)[Z, Z'] \wedge Inv(u)[Z'])$$

where $\phi[Z']$ is a quantifier free formula equivalent to $ReachApprox(v)[Z']$. We apply the quantifier elimination procedure to this formula. If it results to be false, we increase the value of δ to compute another quantifier free formula $\phi[Z']$. Otherwise, we obtain a quantifier free formula $\psi[Z']$ and we can move to the discrete location u where we can apply this procedure considering the formula:

$$ReachApprox(u)[Z'] \equiv \exists Z(\psi[Z] \wedge Z' = f_u(Z, \delta) \wedge Inv(u)[Z'])$$

Proceeding in this way we can keep under control the complexity of our formulæ, avoiding increases in the number of variables and polynomials. Exploiting QEPCAD B to apply this method on the automaton H_2 of Example 2 we can prove that the discrete location u can be reached from v. The result is obtained in 30 milliseconds. Using this method we are able to find approximated solutions for the reachability problem also on automata with independent dynamics with more continuous variables and more complex formulæ than the ones occurring in H_2.

The method can be applied also to automata with non-independent dynamics, but it does not help us, as shown by the following example.

Example 6. Consider the following hybrid automaton which dynamics are non-independent. $H_3 = (Z, Z', \mathcal{V}, \mathcal{E}, Inv, Dyn, Act, Res)$ where:

- $Z = \langle Z_1, Z_2 \rangle$ and $Z' = \langle Z_1', Z_2' \rangle$, where Z_1, Z_2, Z_1', Z_2' variables over $\mathbb{R}$,

- $\mathcal{V} = \{v, u\}$ and $\mathcal{E} = \{e\}$, where e goes from v to u,

- $Inv(v)[Z] \equiv 1 \leq Z_1 \leq 10 \wedge 1 \leq Z_2 \leq 8$ and
  $Inv(u)[Z] \equiv 8 \leq Z_1 \leq 50 \wedge 7 \leq Z_2 \leq 30$,

- $Dyn(v)[Z, Z', T] \equiv Z_1' = Z_1 + (2Z_1^2 + Z_1 Z_2)T \wedge Z_2' = Z_2 + (7Z_2^2 + Z_2 Z_1)T$
  and
  $Dyn(u)[Z, Z', T] \equiv Z_1' = Z_1 + (3Z_1^2 + Z_1 Z_2)T \wedge Z_2' = Z_2 + (4Z_2^2 + Z_2 Z_1)T$,

- $Act(e)[Z] \equiv Z_1 \geq 8 \wedge Z_2 \geq 7$,

- $Res(e)[Z, Z'] \equiv Z_1{}' = Z_1 \wedge Z_2{}' = Z_2$.

Suppose we want to apply the method to find out if the discrete location u is reachable from v. First, we have to apply the quantifier elimination procedure to the formula $ReachApprox(v)[Z']$ with a fixed value δ. When we try to do this using QEPCAD B we are not able to obtain any result within 20 minutes of CPU time because of the presence of non-independent dynamics.

To find approximated solution for the reachability problem on automata with non-independent dynamics, we have to introduce further approximations in our last method. Let H be an automaton with non-independent dynamics. We have that $\forall v \in \mathcal{V}$ the formula $Dyn(v)[Z, Z', T]$ is of the form:

$$Z_1{}' = f_{v,1}(Z, T) \wedge Z_2{}' = f_{v,2}(Z, T) \wedge \ldots \wedge Z_d{}' = f_{v,d}(Z, T)$$

Given the formula $ReachApprox(v)[Z']$ and $\delta > 0$, we compute $\forall = 1, \ldots, d$ the minimum value $(min_i(v))$ and the maximum value $(max_i(v))$ that the function $f_{v,i}(Z, \delta)$ assume in the set defined by the formula $Inv(v)[Z]$. In order to determine an approximation of the values Z' that can be reached after time δ from values Z satisfying $Inv(v)[Z]$, we evaluate the following formula:

$$\bigwedge_{i=1,\ldots,d} min_i(v) \leq Z_i{}' \leq max_i(v) \wedge Inv(v)[Z']$$

If, in the previous method, we use this procedure instead of the quantifier elimination procedure to obtain from a formula $ReachApprox(v)[Z']$ a formula $\phi[Z']$, we have a new method that can find approximated solution to the reachability problem even in presence of non-independent dynamics. We call this method *minimum-Maximum Reachability Approximation* (MIM-RA).

Example 7. Consider the hybrid automaton of Example 6. Suppose we want to apply the MIM-RA method exploiting QEPCAD B to find out if the discrete location u can be reached from v. First, we compute a formula $\phi[Z']$ using the procedure based on the calculus of minimum and maximum of each function in $Dyn(v)[Z, Z', T]$ described above. Then we apply the quantifier elimination procedure to the formula:

$$\exists Z (\phi[Z] \wedge Act(\langle v, u \rangle)[Z] \wedge Res(\langle v, u \rangle)[Z, Z'] \wedge Inv(u)[Z'])$$

14

We obtain a quantifier free formula representing the values for ${Z_1}'$ and ${Z_2}'$ in the discrete location u that can be reached starting from v. We succeed in proving the desired property (result obtained in 55 milliseconds).

We notice that solutions computed with the MIM-RA method are neither over nor under approximations.

Exploiting the MIM-RA method together with QEPCAD B, we could not be able to obtain results on some automata. In particular, on automata whose dynamics are either very complex or not representable in QEPCAD B, e.g., functions where non integer or negative exponents appear. We can solve this problem introducing a further approximation to our method.

Consider a formula $Dyn(v)[Z, Z', T] \equiv \bigwedge_{i=1,\ldots,d} Z_i' = f_{v,i}(Z, T)$. Instead of computing the maximum and the minimum of $f_{v,i}(Z, \delta)$ in the set defined by the formula $Inv(v)[Z]$, we can compute the maximum and the minimum of the linearization of $f_{v,i}$ that is the Taylor polynomial of degree one:

$$Z_i'(\delta) = f_{v,i}(Z(0), 0) + \frac{df_{v,i}}{dT}(Z(0), 0)\delta + R$$

where R is the reminder term. In order to compute the maximum and the minimum at time δ_j (where $\delta_0 = \delta$ and $\delta_j > \delta_{j-1}$) we consider the following expression:

$$Z'(\delta_j) = Z'(\delta_{j-1}) + \frac{df_{v,i}}{dT}(Z(0), \delta_{j-1})\delta_j + R$$

Notice that the derivative $df_{v,i}/dT$ has not to be computed for every time interval. Once computed we can obtain a function that can be used to calculate the values of the derivative for all the different δ_j.

4. SAHA-Tool

In this section we present SAHA-Tool, a software for computing approximated solution to the reachability problem in semi-algebraic automata based on the MIM-RA method. and it is implemented in *Objective Caml* [35].

The reachability problem in semi-algebraic automata is undecidable even if we consider computational models over the reals, as pointed out in Section 2. However, imposing temporal restriction on traces we can define a reachability problem with time bounds for which we can compute approximated solutions using the MIM-RA method. The *time bounded reachability problem* can be defined as follows: given a hybrid automaton $H = (Z,\ Z',\ \mathcal{V},\ \mathcal{E},$

Inv, *Dyn*, *Act*, *Res*) of dimension m, a start location v, a set of start points defined by the formula $\varphi_v[Z]$, a end location u, a set of end points defined by the formula $\varrho_u[Z]$, a time bound $T \in \mathbb{R}_{\geq 0}$ and a value $\delta \in \mathbb{R}_{>0}$, we want to determine if it exists a state $\langle u, s \rangle$, where s belongs to the set of end points, that is reachable from a state $\langle v, r \rangle$, where r belongs to the set of start points, through a trace in which each continuous transition has duration equal to $k\delta$, where $k \in \mathbb{N}$, and such that the sum of continuous transition durations is less than or equal to T.

Approximated solutions to the time bounded reachability problem in continuous semi-algebraic automata with independent or non-independent dynamics can be computed using Algorithm 1. Starting from input data this algorithm return a Boolean value, represented by the variable *reached*, and a set of points defined by the formula $\vartheta[Z]$. If the value of *reached* is **False** then none of the states $\langle u, s \rangle$ with s belonging to the set defined by ϱ_u is reachable from states $\langle v, r \rangle$ with r belonging to the set defined by φ_v. Otherwise, if the value of *reached* is **True** all the states $\langle u, s' \rangle$ with s' belonging to the set defined by ϑ are reachable.

To compute reachable states Algorithm 1 uses a *First In First Out* queue whose elements are tuples of the form $\langle l, t, \varphi, al \rangle$, where l is a location of H, $t \in \mathbb{R}$ is the remaining time for continuous transitions in l, φ is a formula defining a set of points, al is the array of locations adjacent to l. Initially this queue contains the tuple formed by the start location v, the value T, the formula φ_v and the array of locations adjacent to v. For each element $\langle l, t, \varphi, al \rangle$ of the queue Algorithm 1 computes the states reachable through continuous transitions of duration $d = 0, \delta, 2\delta, \ldots, k\delta, \ldots$, until d is less than or equal to the remaining time. To do this it uses the functions REACHABLE and MINIMUMMAXIMUM. The function MINIMUMMAXIMUM computes a formula $\psi_{min-max}[Z']$ of the form

$$\bigwedge_{i=1,\ldots,m} min_i(l) \leq Z'_i \leq max_i(l)$$

where $min_i(l)$ and $max_i(l)$ are the minimum value and the maximum value assumed by the function $f_{l,i}(Z, d)$ within the set defined by the formula φ, respectively. The function REACHABLE evaluates the formula

$$\psi_{min-max}[Z'] \wedge Inv(l)[Z']$$

and returns the formula $\psi_{reached}[Z]$ representing the set of points reached after time d within the location l. After having computed the points reachable

16

Algorithm 1 Time bounded reachability, independent or non-independent dynamics.

Input: $H = (Z,\ Z',\ \mathcal{V},\ \mathcal{E},\ Inv,\ Dyn,\ Act,\ Res)$, $v \in \mathcal{V}$, φ_v, $u \in \mathcal{V}$, ϱ_u, $T \in \mathbb{R}_{\geq 0}$, $\delta \in \mathbb{R}_{>0}$

Output: $reached \in \{\textbf{True},\textbf{False}\}$, ϑ

  $reached \leftarrow \textbf{False}$

  $adjLocations_v \leftarrow \textsc{adjacentLocations}(v,\ \mathcal{V},\ \mathcal{E})$

  $computationQueue \leftarrow \textsc{createQueue}(\ )$

  $\textsc{enqueue}(computationQueue,\ (v,\ T,\ \varphi_v,\ adjLocations_v))$

  while $\textsc{emptyQueue}(computationQueue) = \textbf{False}$ and $reached = \textbf{False}$ **do**

    $(currentLoc, t, \gamma, adjLocations) \leftarrow \textsc{dequeue}(computationQueue)$

    $outInv \leftarrow \textbf{False}$; $d \leftarrow 0$

    while $t \geq d$ and $reached = \textbf{False}$ and $outInv = \textbf{False}$ **do**

      $\psi_{min-max} \leftarrow \textsc{minimumMaximum}(Z',\ Z,\ Dyn(currentLoc),\ d,\ \gamma)$

      $\psi_{reached} \leftarrow \textsc{reachable}(Z,\ Z',\ \psi_{min-max},\ Inv(currentLoc))$

      if $\psi_{reached} \equiv \bot$ **then**

        $outInv \leftarrow \textbf{True}$

      else

        if $currentLoc = u$ and $\psi_{reached} \wedge \varrho_u \not\equiv \bot$ **then**

          $reached \leftarrow \textbf{True}$; $\vartheta \leftarrow \psi_{reached} \wedge \varrho_u$

        else

          $n \leftarrow \textsc{length}(adjLocations)$; $i \leftarrow 0$

          while $(i < n$ and $reached = \textbf{False})$ **do**

            $l \leftarrow adjLocations[i]$; $e \leftarrow \langle locCorrente, l \rangle$

            $\psi_{edge} \leftarrow \textsc{edge}(Z,\ Z',\ \psi_{reached},\ Act(e),\ Res(e),\ Inv(l))$

            if $\psi_{edge} \not\equiv \bot$ **then**

              if $l = u$ and $\psi_{edge} \wedge \varrho_u \not\equiv \bot$ **then**

                $reached \leftarrow \textbf{True}$; $\vartheta \leftarrow \psi_{edge} \wedge \varrho_u$

              else

                $la_l \leftarrow \textsc{adjLocations}(l,\ \mathcal{V},\ \mathcal{E})$

```
                    if LENGTH(la_l) > 0 or l = u then
                        t' ← t − d
                        ENQUEUE(computationQueue,(l, t', ψ_edge, la_l))
                        i ← i + 1
                    else
                        i ← i + 1
                    end if
                end if
            else
                i ← i + 1
            end if
        end while
        d ← d + δ
    end if
 end if
end while
end while
```

through a continuous transition, if none of the end states has been reached, Algorithm 1 evaluates for each edge $e = \langle l, l' \rangle$, where l' is a location adjacent to l, the formula

$$\exists Z(\psi_{reached}[Z] \wedge Act(e)[Z] \wedge Res(e)[Z, Z'] \wedge Inv(l')[Z'])$$

using the function EDGE that returns the formula ψ_{edge} representing the set of point reached within l' after the discrete transition. If ψ_{edge} is not equivalent to $\bot$ then the edge can be crossed, if the location l' has adjacent locations then the tuple $\langle l', t', \psi_{edge}, al' \rangle$, where $t' = t - d$ and al' is the array of locations adjacent to l', is enqueued. Algorithm 1 terminates when end states are reached or the queue is empty. Theorem 1 and Theorem 2 holds for Algorithm 1.

Theorem 1. *Let $H = (Z, Z', \mathcal{V}, \mathcal{E}, Inv, Dyn, Act, Res)$ be a semi-algebraic hybrid automaton on which Algorithm 1 works. Let d be the dimension of H, $n = |\mathcal{V}|$, T the time bound and δ the value used to determine the duration of continuous transitions. We use Q to indicate the complexity of the formulæ resolver and we assume that in each location of the automaton at least time δ elapses before a discrete transition occurs. The complexity of Algorithm 1*

as a function of Q is

$$O\left(n^{(\lfloor T/\delta \rfloor + 1)} \lfloor T/\delta \rfloor (d+1)Q\right)$$

Theorem 2. *Algorithm 1 is correct with respect to the* MIM-RA *method.*

The time bounded reachability problem can also be considered in hybrid automata whose dynamics are defined by systems of autonomous differential equations. In this case approximated solutions can be computed approximating the solutions of differential equations with the corresponding Taylor polynomial of a certain degree in order to obtain a continuous semi-algebraic automaton, and using a slightly variation of Algorithm 1 that exploit the fact that continuous transitions are transitive when dynamics are defined by systems of autonomous differential equations.

Algorithm 1 and its variation have been implemented in the software SAHA-TOOL (Semi Algebraic Hybrid Automata Tool). SAHA-TOOL is a software for computing approximated solution to the time bounded reachability problem in continuous semi-algebraic automata and in automata whose dynamics are defined by systems of autonomous differential equations (for solving the problem in this type of automata it approximates the solutions of differential equations with the corresponding Taylor polynomial of degree one). This software has been implemented using Objective Caml and it uses QEPCAD B to evaluate formulæ and to compute the minimum and the maximum value of a function. To evaluate the quality of the result produced by SAHA-TOOL we compared it with HSOLVER [36], a software for the verification of safety properties of hybrid automata. HSOLVER uses a method of constraint propagation based abstraction refinement and it is implemented on the top of RSOLVER. The comparison showed that the results computed by SAHA-TOOL are in line with the ones computed by HSOLVER.

5. Case Studies

As a simple yet very interesting examples, we consider the Repressilator system and the Delta-Notch protein signaling mechanism.

5.1. The Repressilator Case Study

The Repressilator system constructed by Elowitz and Leibler [37]. It consists of three proteins, namely lacI, tetR, and cI, and the corresponding genes. The protein lacI represses the gene which expresses tetR, tetR represses

the gene which expresses cl, whereas cl represses the gene which expresses lacl, thus completing a feedback system. The dynamics of the network depend on the transcription rates, translation rates, and decay rates. Depending on the values of these rates the system might converge to a stable limit circle or become unstable.

We apply our method to compare the behaviour of two oscillating models proposed for the Repressilator.

First, we consider the hybrid automaton proposed in [38] to model the Repressilator. This hybrid automaton has 8 discrete locations, corresponding to all the possible combinations of genes being either *on* or *off*, and 9 variables. Three of them, A, B, C, represent the quantity of proteins in the system, the other six, $Y_{X,on}$, $Y_{X,off}$, where $X \in \{A, B, C\}$, control activation and deactivation of genes. For each discrete location v, $Inv(v) \equiv true$.

The differential equations governing proteins concentrations in each discrete location are decoupled: for instance, when gene A is *on* its dynamics is $\dot{A} = k_p - k_d A$, where k_p and k_d are costant parameters of the system.

The interactions between repressors and genes are confined to the activation conditions of the automaton transitions. Consider again gene A and suppose to be in a discrete location of the automaton where it is *on*. Then, the differential equation for $\dot{Y}_{A,off}$ is $\dot{Y}_{A,off} = k_b C$, the transition switching this gene *off* has an activation condition equal to $Y_{A,off} \geq 1 \wedge C \geq 1$ and a reset condition equal to $Y_{A,on} = 0 \wedge Y_{A,off} = 0$. The transition that turns gene A *on*, instead, has a constant rate k_u, hence its activation condition is $Y_{A,on} \geq 1$, $\dot{Y}_{A,on} = k_u$ is the differential equation for $\dot{Y}_{A,on}$ and the reset condition is equal to $Y_{A,on} = 0 \wedge Y_{A,off} = 0$, where k_b and k_u are costant parameters of the system.

In order to obtain a continuous semi-algebraic automaton from this hybrid automaton, we have only to define for each discrete location v a formula $Dyn(v)$ satisfying the conditions of Definition 4. To this aim we approximate the solutions of the differential equations in each discrete location with the corresponding Taylor polynomial of degree two. Consider, for instance, the differential equations for A, $Y_{A,off}$, and $Y_{A,on}$ in a discrete location where gene A is *on*, we approximate their solution with the following polynomials:

$$A' = A + (k_p - k_d A)T + (-k_d k_p + k_d^2 A)T^2/2$$
$$Y'_{A,off} = Y_{A,off} + (k_b C)T + (-k_b k_d C)T^2/2 \qquad \text{if gene } C \text{ is } off$$
$$Y'_{A,off} = Y_{A,off} + (k_b C)T + (k_b k_p - k_b k_d C)T^2/2 \qquad \text{if gene } C \text{ is } on$$
$$Y'_{A,on} = Y_{A,on} + k_u T$$

20

The solution of the differential equation for A in a discrete location where the gene A is *off* is approximated with the following polynomial:

$$A' = A + (-k_d A)T + (k_d^2 A)T^2/2$$

The automaton we obtain has non-independent dynamics (see, e.g., the equation for $Y'_{A,off}$), hence we analyse it using the MIM-RA method. Starting from the discrete location where only gene A is active and with fixed values for proteins concentrations we succeed in simulating the automaton and observe an oscillatory behaviour (Figure 1).
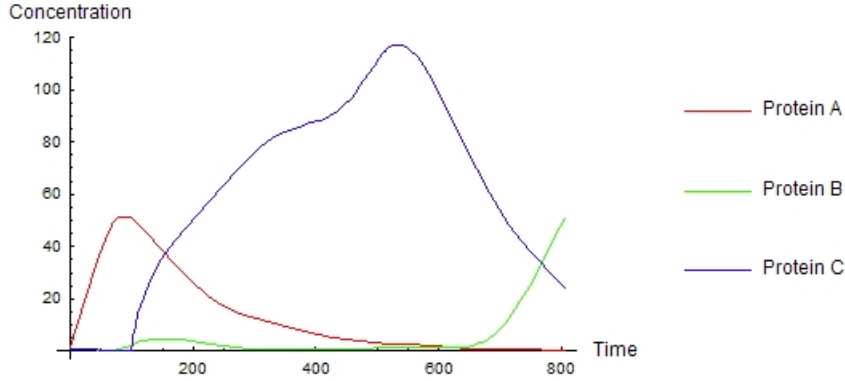

Figure 1: Time trace of the hybrid automata with 8 discrete locations. Parameters are $k_p = 1$, $k_d = 0.01$, $k_b = 1$, $k_u = 0.01$.

The second hybrid automata we consider is the one that can be constructed from the following model for the repressilator written in the S-System equations formalism [39] (see [40])

$$\begin{array}{ll}
\dot{X}_1 = \alpha_1 X_3^{-1} - \beta_1 X_1^{0.5}, & \alpha_1 = 0.2,\ \beta_1 = 1, \\
\dot{X}_2 = \alpha_2 X_1^{-1} - \beta_2 X_2^{0.578151}, & \alpha_2 = 0.2,\ \beta_2 = 1, \\
\dot{X}_3 = \alpha_3 X_2^{-1} - \beta_3 X_3^{0.5}, & \alpha_3 = 0.2,\ \beta_3 = 1.
\end{array}$$

From this model we obtain an hybrid automaton with only one discrete location, no transitions and three variables, X_1, X_2, X_3, representing proteins concentrations. For the unique discrete location v we have $Inv(v) \equiv true$, the dynamics in v are defined by the differential equations of the S-System model.

21

As in the previous case, to obtain a continuous semi-algebraic automaton we approximate the solutions of the differential equations in the discrete location with the corresponding Taylor polynomial of degree two. Consider for instance the differential equation for X_1, we approximate its solution with the following polynomial:

$$\begin{aligned} X_1' \;=\;\; & X_1 + (0.2X_3^{-1} - X_1^{0.5})T+ \\ & (-0.04X_3^{-2}X_2^{-1} + 0.2X_3^{-1.5} - 0.1X_1^{-0.5}X_3^{-1} + 0.5)T^2/2 \end{aligned}$$

The automaton we obtain has non-independent dynamics with real exponents, hence we analyse it using the approximated method based on the computation of minimum and maximum values of the linearization of the functions defining the dynamics. We succeed in the simulation of the automaton, but we do not obtain any interesting result.

The analysis of the two models shows that the one obtained from the S-System does not permit to observe the oscillatory behaviour of the repressilator, this because of the approximations introduced for simulation. The other model, instead, results to be less sensitive to approximation and simulating it we can observe the oscillations in proteins concentrations. This points out that in order to define robust models for biological systems it is important to distinguish from the beginning the discrete from the continuous parts of the systems. Hybrid automata allow to do this and hence to obtain simpler dynamics in each discrete location. Such dynamics are less sensible to the approximations which are necessary to carry out formal analysis.

5.2. The Delta-Notch Protein Signaling

Cellular differentiation, the process by which cells acquire their specialization, such as heart cells, muscle cells, skin cells, and brain cells, is a well studied phenomenon. It occurs numerous times during the development of a multicellular organism as the organism changes from a single zygote to a complex system of tissues and cell types. Genes control cell fate by controlling the type and amount of proteins made in a cell. Proteins in turn affect gene activity by turning "on" or "off" gene expression thereby affecting the production of proteins themselves. Hence, differential gene activity is considered the key to cell differentiation [41] and protein concentrations in a cell are a good measure of gene activity. The idea that lateral signaling between cells through the Delta-Notch protein pathway is responsible for same cell fate decisions has gained wide acceptance.

Delta and Notch are both transmembrane proteins that actively signal only when cells are in direct contact, in a densely packed epidermal layer. Delta is a ligand that binds and activates its receptor Notch in neighboring cells. The activation of Notch in a cell affects the production of Notch ligands (i.e. Delta) both in itself and its neighbors, thus forming a feedback control loop. In the case of lateral inhibition, high Notch levels suppress ligand production in the cell and thus a cell producing more ligands force its neighboring cells to produce less.

We apply our method to the hybrid automaton developed in [42] to model the Delta-Notch protein signaling mechanism, and compare our results to the ones reported in [42] and [43].

To model the regulation of intercellular Delta and Notch protein concentrations through the feedback network, experimentally observed rules governing the biological phenomenon have to be implemented. First, cells have to be in direct contact for Delta-Notch signaling to occur. This implies that a cell is directly affected by, and directly affects in turn, only immediate neighbors. Second, Notch production is turned on by high Delta levels in the immediate neighborhood of the cell and Delta production is switched on by low Notch concentrations in the same cell. Third, at steady state, a cell with high Delta level must have low Notch level and vice versa. Finally, both Delta and Notch protein concentrations decay exponentially.

In [42] each biological cell is modeled as a hybrid automaton defined by: a set of global invariant conditions which must be always true; a complete graph with four states capturing the property that Notch and Delta protein production can be individually switched on or off at any given time; for each state, a set of local invariant conditions and a set of differential equations determining the flow of the two variables representing Delta and Notch protein concentrations, respectively. Differential equations, invariant conditions and edge activation and reset conditions effectively implement the experimentally observed rules described above.

The hybrid automaton representing the evolution of two cells presented in [42] is the composition of two single cell automata. It is an automaton with four continuous variables and sixteens discrete location, it has the following set of global invariant conditions

$$0 \le d_1, d_2 \le R_D/\lambda_D \quad \wedge \quad 0 \le n_1, n_2 \le R_N/\lambda_N$$
$$\wedge \quad -R_N/\lambda_N \le h_D \le 0 \quad \wedge \quad 0 \le h_N \le R_D/\lambda_D$$

The variables d_1 and d_2 represent the concentration of the Delta protein

in the first and in the second cell, respectively. The variables n_1 and n_2 represent the concentration of the Notch protein in the first and in the second cell, respectively. R_D and R_N are constants representing Delta and Notch production rates, respectively. λ_D and λ_N are the Delta and Notch protein decay constants, respectively. h_D is an unknown switching threshold which determines the Delta protein production. h_N, similar to h_D, is an unknown switching threshold which determines the Notch protein production.

A possible equilibrium for the system is given by the point $d_1^* = 0$, $n_1^* = R_N/\lambda_N$, $d_2^* = R_D/\lambda_D$, $n_2^* = 0$, which belongs to the discrete location v characterized by the following invariant conditions

$$0 \leq d_1 \leq h_N \wedge -h_D \leq n_1 \leq R_N/\lambda_N \wedge h_N \leq d_2 \leq R_D/\lambda_D \wedge 0 \leq n_2 \leq -h_D$$

and by the following differential equations determining the flows of continuous variables

$$\dot{d_1} = -\lambda_D d_1, \ \dot{n_1} = R_N - \lambda_n n_1, \ \dot{d_2} = R_D - \lambda_D d_2, \ \dot{n_2} = -\lambda_N n_2$$

In [42] the hybrid automaton representing the evolution of two cells has been studied using the predicate abstraction methods presented in [44], the analysis performed proved that each point satisfying the condition $d_1 < d_2 \wedge n_1 > n_2$ reaches the equilibrium state belonging to the discrete location v. In [43] the authors applied the method they developed to the analysis of admissible location reachable from v, and found a point in a location u satisfying the condition $d_1 > d_2 \wedge n_1 > n_2$ that with $R_N = R_D = \lambda_N = \lambda_D = 1.0$ and $-h_D = h_N = 0.5$ reaches a point in v such that $d_1 < d_2 \wedge n_1 > n_2$, i.e. able to converge to the equilibrium state in v. The result obtained in [43] does not contradict the result presented in [42], it proves that the two different methods can be combined to obtain better approximations of the set of points from which the equilibrium state in v is reachable.

In order to obtain a continuous semi-algebraic hybrid automaton from the automaton representing the evolution of two cells described in [42], we approximate the solutions of the differential equations in each location with the corresponding Taylor polynomial of degree three, and define for each location w a formula $Dyn(w)$ satisfying the condition of Definition 4. Consider, for instance, the differential equation for n_1 in the location v cited above, $\dot{n_1} = R_N - \lambda_N n_1$, we approximate its solution with the following polynomial

$$n_1' = n_1 + (R_N - \lambda_N)T + (-\lambda_N R_N + \lambda_N^2 n_1)T^2/2 + (\lambda_N^2 R_N - \lambda_N^3 n_1)T^3/6$$

The automaton we obtain is the composition of two semi-algebraic automata for one cell and has independent dynamics, hence we analyze it using the approximated method presented for automata with independent dynamics. Using this method we can prove that each point satisfying the following condition

$$d_1 < d_2 \wedge n_1 < n_2 \wedge d_2 - d_1 > 4/5 \wedge n_2 - n_1 < 1/50$$

belonging to the discrete location q characterized by the following invariant conditions

$$0 \le d_1 \le h_N \wedge -h_D \le n_1 \le R_N/\lambda_N \wedge h_N \le d_2 \le R_D/\lambda_D \wedge -h_D \le n_2 \le 1$$

reaches points satisfying $d_1 < d_2 \wedge n_1 > n_2$ in the location v when $R_N = R_D = \lambda_N = \lambda_D = 1.0$ and $-h_D = h_N = 0.5$, that are points from which the equilibrium state in v is reachable. Our result, which does not contradict the one presented in [42], nonetheless proves that with our method it is possible to find more points from which the equilibrium state in v is reachable.

6. Conclusions

In this paper we presented some experimental results on the reachability problem in semi-algebraic hybrid automata. Our results suggest that even if we try to exploit different techniques and powerful tools, we cannot go *far* enough, without introducing approximations.

However, it is easy to apply some standard, basic, approximation techniques. We showed on the repressilator case study that the approximated results are coherent with the expected behaviour, even when we limit our approximations to the first and second degree, provided that intrinsic discrete nature of the system has been explicitly modeled. In particular, the approximations on the 8-states automaton show the oscillations, while this is not the case if we directly apply our method to the system of differential equations. Intuitively, the system of differential equations implicitly models the discrete nature of the system exploiting more complex dynamics whose simulation requires more sophisticated techniques. The hybrid automaton allows to keep the dynamics more simple and more *robust* to approximations.

References

[1] R. Alur, T. A. Henzinger, P.-H. Ho, Automatic Symbolic Verification of Embedded Systems, in: Proceedings of IEEE Real-Time Systems Symposium, IEEE Computer Society Press, 1993, pp. 2–11.

[2] R. Alur, C. Belta, F. Ivancic, V. Kumar, M. Mintz, G. J. Pappas, H. Rubin, J. Schug, Hybrid Modeling and Simulation of Biomolecular Networks, in: Proceedings of Hybrid Systems: Computation and Control (HSCC'01), Vol. 2034 of LNCS, Springer-Verlag, 2001, pp. 19–32.

[3] R. Ghosh, A. Tiwari, C. Tomlin, Automated Symbolic Reachability Analysis; with Application to Delta-Notch Signaling Automata, in: O. Maler, A. Pnueli (Eds.), Proceedings of Hybrid Systems: Computation and Control (HSCC'03), Vol. 2623 of LNCS, Springer-Verlag, 2003, pp. 233–248.

[4] T. A. Henzinger, P. W. Kopke, A. Puri, P. Varaiya, What's decidable about hybrid automata?, in: Proceedings of the Twenty-Seventh Annual ACM Symposium on the Theory of Computing (STOC '95), ACM Press, New York, NY, USA, 1995, pp. 373–382.

[5] C. Piazza, M. Antoniotti, V. Mysore, A. Policriti, F. Winkler, B. Mishra, Algorithmic Algebraic Model Checking I: Challenges from Systems Biology, in: K. Etessami, S. K. Rajamani (Eds.), Proceedings Computer Aided Verification (CAV'05), no. 3576 in LNCS, Springer-Verlag, 2005, pp. 5–19.

[6] A. Tarski, A Decision Method for Elementary Algebra and Geometry, Univ. California Press, 1951.

[7] C. W. Brown, QEPCAD B: A program for computing with semi-algebraic sets using CADs, ACM SIGSAM Bulletin 37 (4) (2003) 97–108.

[8] V. Mysore, C. Piazza, B. Mishra, Algorithmic Algebraic Model Checking I: Decidability of Semi-Algebraic Model Checking and its Applications to Systems Biology, in: Y. T. D. A. Peled (Ed.), Proceedings Third International Symposium on Automated Technology for Verification and Analysis (ATVA 2005), no. 3707 in LNCS, Springer-Verlag, 2005, pp. 217–233.

[9] V. Mysore, B. Mishra, Algorithmic Algebraic Model Checking III: Approximate Methods, Electronic Notes in Theoretical Computer Science 149 (1) (2006) 61–77.

[10] S. Ratschan, Efficient Solving of Quantified Inequality Constraints over the Real Numbers, ACM Transactions on Computational Logic 7 (4) (2006) 723–748.

[11] K. Apt, M. Wallace, Constraint Logic Programming using Eclipse, Cambridge University Press, 2006.

[12] R. Alur, C. Courcoubetis, N. Halbwachs, T. A. Henzinger, P.-H. Ho, X. Nicollin, A. Olivero, J. Sifakis, S. Yovine, The algorithmic analysis of hybrid systems, Theoretical Computer Science 138 (1) (1995) 3–34.

[13] R. Alur, D. L. Dill, A theory of timed automata, Theoretical Computer Science 126 (2) (1994) 183–235.

[14] A. Puri, P. Varaiya, Decidability of hybrid systems with rectangular differential inclusions, in: D. L. Dill (Ed.), Proceedings of International Conference on Computer Aided Verification (CAV'94), Vol. 818 of LNCS, Springer-Verlag, 1994, pp. 95–104.

[15] G. Lafferriere, G. J. Pappas, S. Sastry, O-Minimal Hybrid Systems, Mathematics of Control, Signals, and Systems 13 (2000) 1–21.

[16] A. Casagrande, C. Piazza, B. Mishra, Semi-Algebraic Constant Reset Hybrid Automata -SACoRe, in: Proc. of the 44rd Conference on Decision and Control (CDC'05), IEEE Computer Society Press, 2005, pp. 678–683.

[17] N. Halbwachs, Y.-E. Proy, P. Raymond, Verification of linear hybrid systems by means of convex approximations, in: B. Le Charlier (Ed.), Static Analysis Symposium, Springer-Verlag, 1994, pp. 223–237.

[18] T. Dang, O. Maler, Reachability analysis via face lifting, in: T. A. Henzinger, S. Sastry (Eds.), HSCC, Vol. 1386 of LNCS, Springer-Verlag, 1998, pp. 96–109.
URL `citeseer.nj.nec.com/dang98reachability.html`

[19] A. Chutinan, B. Krogh, Verification of Polyhedral-Invariant Hybrid Automata Using Polygonal Flow Pipe Approximations, in: F. W. Vaandrager, J. H. van Schuppen (Eds.), Proceedings of Hybrid Systems: Computation and Control (HSCC'99), Vol. 1569 of LNCS, Springer-Verlag, 1999, pp. 76–90.

[20] E. Asarin, T. Dang, O. Maler, O. Bournez, Approximate Reachability Analysis of Piecewise-Linear Dynamical Systems, in: B. Krogh, N. Lynch (Eds.), Proceedings of Hybrid Systems: Computation and Control (HSCC'00), Vol. 1790 of LNCS, Springer-Verlag, 2000, pp. 20–31.

[21] A. B. Kurzhanski, P. Varaiya, On verification of controlled hybrid dynamics through ellipsoidal techniques, in: Proceedings of the 44rd Conference on Decision and Control and European Control Conference (CDC-ECC'05), IEEE Computer Society Press, Seville, Spain, 2005, pp. 4682–4687.

[22] R. Alur, T. Dang, F. Ivancic, Progress on reachability analysis of hybrid systems using predicate abstraction, in: O. Maler, A. Pnueli (Eds.), HSCC, Vol. 2623 of LNCS, Springer-Verlag, 2003, pp. 4–19.

[23] T. A. Henzinger, P. H. Ho, H. Wong-Toi, HYTECH: A Model Checker for Hybrid Systems, International Journal on Software Tools for Technology Transfer 1 (1–2) (1997) 110–122.

[24] E. Asarin, T. Dang, O. Maler, The d/dt tool for verification of hybrid systems, in: Proceedings of the Forteenth International Conference on Computer Aided Verification (CAV'02), LNCS, Springer-Verlag, London, UK, 2002, pp. 365–370.

[25] B. I. Silva, B. H. Krogh, Formal verification of hybrid system using checkmate: A case study, in: Proceedings of the American Control Conference 2000 (ACC'00), IEEE Computer Society Press, Chicago, Illinois, 2000, pp. 678–683.

[26] J. Bengtsson, K. G. Larsen, F. Larsson, P. Pettersson, W. Yi, Uppaal - a tool suite for automatic verification of real-time systems., in: R. Alur, T. A. Henzinger, E. D. Sontag (Eds.), Hybrid Systems III: Verification

and Control, Proceedings of the DIMACS/SYCON Workshop, Vol. 1066 of LNCS, Springer-Verlag, 1996, pp. 232–243.

[27] C. Daws, A. Olivero, S. Tripakis, S. Yovine, The tool KRONOS, in: R. Alur, T. A. Henzinger, E. D. Sontag (Eds.), Hybrid Systems III: Verification and Control, Proceedings of the DIMACS/SYCON Workshop, Vol. 1066 of LNCS, Springer-Verlag, 1996, pp. 208–219.

[28] R. Lanotte, S. Tini, Taylor approximation for hybrid systems, Inf. Comput. 205 (11) (2007) 1575–1607.

[29] G. E. Collins, Quantifier Elimination for the Elementary Theory of Real Closed Fields by Cylindrical Algebraic Decomposition, in: Proceedings of the Second GI Conference on Automata Theory and Formal Languages, Vol. 33 of LNCS, Springer-Verlag, 1975, pp. 134–183.

[30] D. Grigorév, N. Vorobjov, Counting connected components of a semi-algebraic set in subexponential time, Computational Complexity 2 (2) (1992) 133–186. doi:http://dx.doi.org/10.1007/BF01202001.

[31] J. Renegar, On the computational complexity and geometry of the first-order theory of the reals. Part III: Quantifier elimination, Journal of Symbolic Computation 13 (3) (1992) 329–352.

[32] S. Basu, An improved algorithm for quantifier elimination over real closed fields, in: Proceedings of the Thirty-Eighth Annual Symposium on Foundations of Computer Science (FOCS '97), IEEE Computer Society Press, Washington, DC, USA, 1997, pp. 56–65.

[33] A. Dolzmann, T. Sturm, REDLOG: Computer algebra meets computer logic, SIGSAM Bulletin (ACM Special Interest Group on Symbolic and Algebraic Manipulation) 31 (2) (1997) 2–9.

[34] T. Sturm, Quantifier Elimination-Based Constraint Logic Programming, Tech. Rep. MIP-0202, Fakultät für Mathematik und Informatik, Universität Passau (2002).

[35] X. Leroy, D. Doligez, J. Garrigue, D. Rmy, J. Vouillon, The Objective Caml system release 3.10 Documentation and user's manual (2007).

[36] S. Ratschan, Z. She, Safety verification of hybrid systems by constraint propagation based abstraction refinement, ACM Journal in Embedded Computing Systems 6 (1).

[37] M. Elowitz, S. Leibler, A Synthetic Oscillatory Network of Transcriptional Regulators, Nature 403 (2000) 335–338.

[38] L. Bortolussi, A. Policriti, Hybrid approximation of Stochastic Concurrent Constraint Programming, in: International Federation of Automatic Control World Congress, (IFAC'08), 2008, to appear.

[39] E. O. Voit, Computational Analysis of Biochemical Systems. A Pratical Guide for Biochemists and Molecular Biologists, Cambridge University Press, 2000.

[40] M. Antoniotti, A. Policriti, N. Ugel, B. Mishra, Model Building and Model Checking for Biological Processes, Cell Biochemistry and Biophysics.

[41] L. Wolpert, Principles of Development, Current Biology Ltd e Oxford University Press, 1998.

[42] R. Ghosh, A. Tiwari, C. Tomlin, Automated Symbolic Reachability Analysis; with Application to Delta-Notch Signaling Automata, in: O. Maler, A. Pnueli (Eds.), Proceedings of Hybrid Systems: Computation and Control (HSCC'03), Vol. 2623 of LNCS, Springer-Verlag, 2003, pp. 233–248.

[43] C. Piazza, M. Antoniotti, V. Mysore, A. Policriti, F. Winkler, B. Mishra, Algorithmic Algebraic Model Checking I: Challenges from Systems Biology, in: K. Etessami, S. K. Rajamani (Eds.), Proceedings Computer Aided Verification (CAV'05), no. 3576 in LNCS, Springer-Verlag, 2005, pp. 5–19.

[44] A. Tiwari, G. Khanna, Series of Abstraction for Hybrid Automata, in: C. Tomlin, M. R. Greenstreet (Eds.), HSCC, Vol. 2289 of LNCS, Springer-Verlag, 2002.