Approximate Reachability

for Linear Systems

Ashish Tiwari

Tiwari@csl.sri.com

Computer Science Laboratory

SRI International

Menlo Park CA 94025

http://www.csl.sri.com/~tiwari

Outline

1. Motivation
2. Hybrid-Abstraction
3. New Idea
4. Example
5. Conclusion
6. Future Work

Motivation: The Analysis Problem

- Given a hybrid automaton
- And a safety property
- Show that the property is true of the model

Some Approaches

- Simulation and testing
- Decidability Results: Reachability algorithms for subclasses of hybrid automata
  - Timed Automata
  - Initialized Multirate Automata
  - Initialized Rectangular Automata
  - Linear Vector Fields specified using matrix A, where A is either nilpotent, diagonalizable with rational eigenvalues, or has only imaginary eigenvalues
  - O-minimal automata
- Abstraction Techniques: Compute approximations and work on them

Motivation

Our approach: Abstract and explore the abstraction s.t.

- Avoid hard quantifier-elimination problems
- Go beyond decidability limits
- Guarantee soundness
- Keep techniques compositional

Tradeoff: Quality of approximation vs Computational effort

Abstraction: The General Idea

1. Partition the state space: each member of the partition is a state of a new system
   - defines an abstraction map from concrete states onto a smaller set of “abstract states”
   - one abstract state corresponds to a set of concrete states
2. Quotient the original system modulo the partition:
   - there is a transition from one abstract state to another if there is a possible concrete one between the corresponding sets
   - the abstract system is a conservative approximation of the original system
3. Analyze the abstract system

Hybrid Abstraction

Hybrid Automata

A tuple $(Q, X, \mathbf{S}_0, F, Inv, R)$:

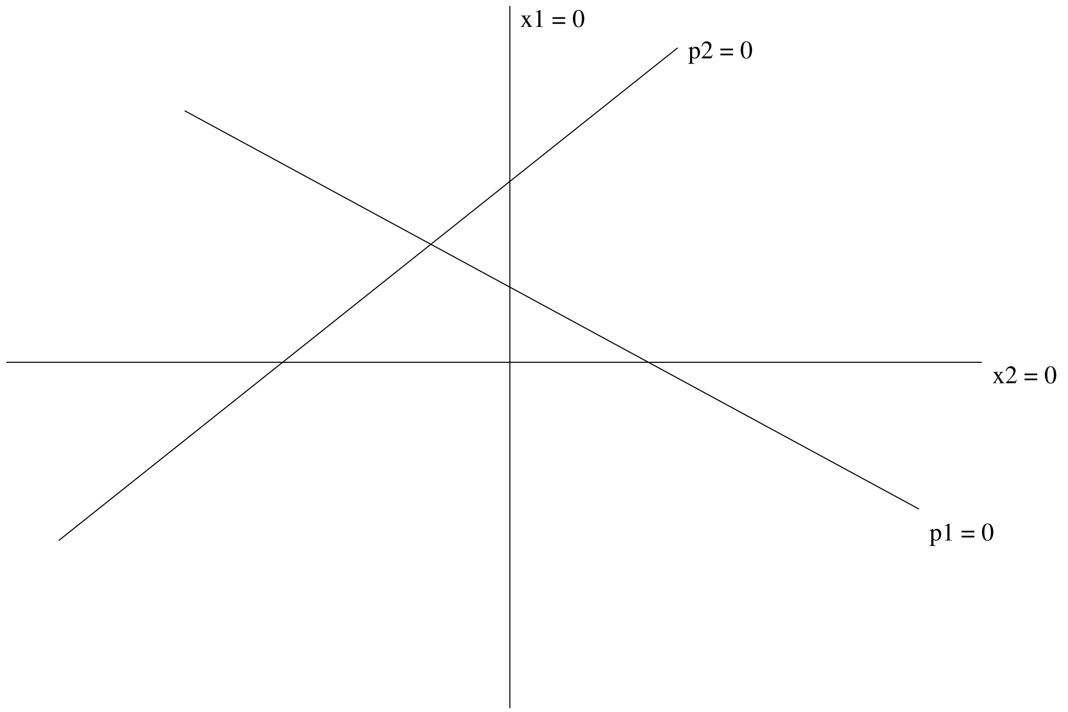
- Q: finite set of discrete variables
- X: finite set of continuous variables
- $\mathbf{X} = \mathbb{R}^{|X|}$, $\mathbf{Q}$ = set of all valuations for Q
- $\mathbf{S} = \mathbf{Q} \times \mathbf{X}$
- $\mathbf{S}_0 \subseteq \mathbf{S}$ is the set of initial states
- $F : \mathbf{Q} \mapsto (\mathbf{X} \mapsto \mathbb{R}^{|X|})$ specifies the rate of flow, $\dot{x} = F(q)(x)$
- $Inv : \mathbf{Q} \mapsto 2^{\mathbb{R}^{|X|}}$ gives the invariant set
- $R \subseteq \mathbf{Q} \times 2^{\mathbf{X}} \mapsto \mathbf{Q} \times 2^{\mathbf{X}}$ captures discontinuous state changes

Partitioning the State Space

- The discrete component, $\mathbf{Q}$, of the state space, $\mathbf{Q} \times \mathbf{X}$, is abstracted independently
  - Assume it is not abstracted
- The continuous component, $\mathbf{X}$, of the state space is partitioned based on the signs of finite number of functions over the continuous state variables
  - Let $f_1, f_2, \ldots, f_m$ be functions from $\mathbf{X}$ to $\mathbb{R}$
  - Let $q_{f1}, q_{f2}, \ldots, q_{fm}$ be m corresponding discrete variables interpreted over $\{pos, neg, zero\}$ s.t.
    - ⋆ $q_{f1} = pos$ iff $f_1 > 0$
    - ⋆ $q_{f1} = neg$ iff $f_1 < 0$
    - ⋆ $q_{f1} = zero$ iff $f_1 = 0$

Abstraction Algorithm: 1

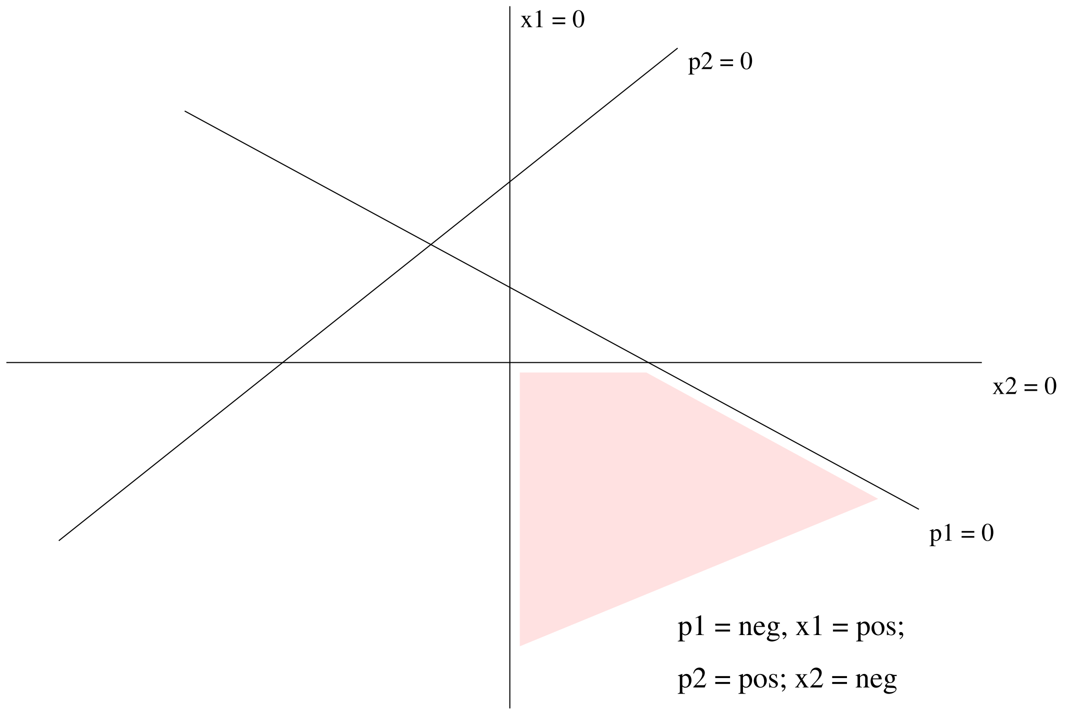
Consider a continuous dynamical system with two state variables. Concrete state space: $\mathbb{R}^2$



Partitioned w.r.t signs of four linear forms x_1, x_2, p_1, and p_2.

Abstraction Algorithm: 2

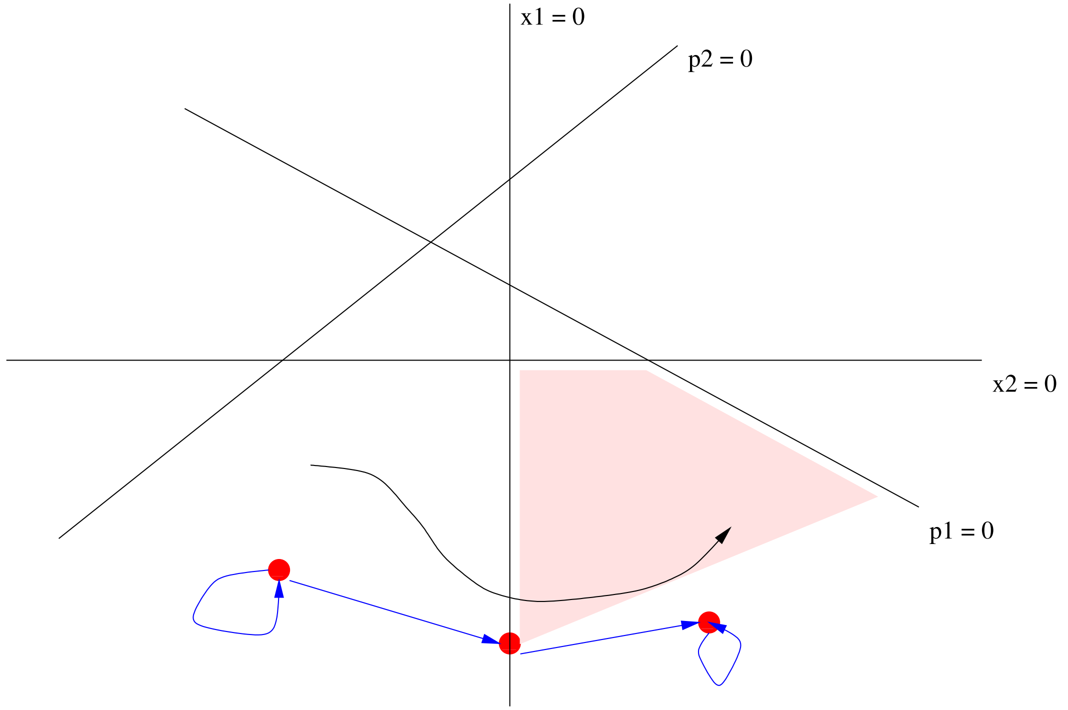
Abstract states correspond to subsets of concrete states.



Total number of abstract states = $3^4 = 81$, but feasible abstract states = $11 + 16 + 6 = 33$

Abstraction Algorithm: 3

Abstract transitions overapproximate concrete transitions.



How to abstract the continuous transitions?.

Abstracting the Continuous Dynamics

Concrete state space	:	$\mathbf{Q} \times \mathbb{R}^n$
Abstract state space	:	$\mathbf{Q} \times 3^m$

Our approach is based on qualitative reasoning

- For a given mode (continuous dynamics), for each f_i

$$
\begin{array}{lll}
f_i > 0 \ \wedge \ \dot{f}_i \geq 0 & \longrightarrow & f_i' > 0 \\
f_i > 0 \ \wedge \ \dot{f}_i < 0 & \longrightarrow & f_i' > 0 \ \vee \ f_i' = 0
\end{array}
$$

- Do this for all possible modes in the system to get abstractions of the continuous evolutions in all modes

But how do I know the sign of $\dot{f}_1$

Choosing Predicates for Abstraction

- If $\dot{f}_i \in \{f_1, f_2, \ldots, f_m\}$, then we know the sign of $\dot{f}_i$
  - Explicitly add derivatives of f_i's to the set
- If not, we could still deduce the sign of $\dot{f}_i$ from
  - the signs of $\{f_1, f_2, \ldots, f_m\}$
  - state invariants

  More theorem proving proof obligations!
- Add those functions to the set whose sign can be "easily" inferred
  - For linear dynamical systems, forms corresponding to eigenvectors of transpose . . . . . . **NEW**

Choosing Partition Forms

Linear system $\dot{\vec{x}} = A\vec{x}$

Assume linear form $f = \vec{r}^T\vec{x}$

Question: For what $\vec{r}$ is this "good"?

We want $\dot{f} = kf$

i.e., $\vec{r}^T\dot{\vec{x}} = k\vec{r}^T\vec{x}$

i.e., $\vec{r}^T A\vec{x} = k\vec{r}^T\vec{x}$

i.e., $\vec{r}$ is an eigenvector of A^T corresponding to eigenvalue k

We want k to be real

For such f, qualitative abstraction of continuous dynamics is easy

Choosing Partition Forms

Linear system $\dot{\vec{x}} = A\vec{x}$

Assume linear form $f = \vec{r}^T \vec{x}$

Question: For what $\vec{r}$ is this "good"?

We want $\ddot{f} = b\dot{f} + cf$

i.e., $\vec{r}^T \ddot{\vec{x}} = b\vec{r}^T \dot{\vec{x}} + c\vec{r}^T \vec{x}$

i.e., $\vec{r}^T A^2 \vec{x} = b\vec{r}^T A\vec{x} + c\vec{r}^T \vec{x}$

i.e., $\vec{r}$ is in the null space of $(A^T)^2 - bA^T - cI$

Generalize to a primary decomposition for A^T!

For such f, sign of $\ddot{f}$ can be inferred from the signs of f and $\dot{f}$, at least in some cases

Choosing Partition Forms: Primary Decomposition

Say $ch(A) = \Pi_i(a_i\lambda^2 + b_i\lambda + c_i)$, $a_i, b_i, c_i \in \mathbb{R}$

Let $\vec{r} \in Null(a_i(A^T)^2 + b_i A^T + c_i I)$

Then, $f = \vec{r}^T\vec{x}$ is a "good" linear form for abstraction

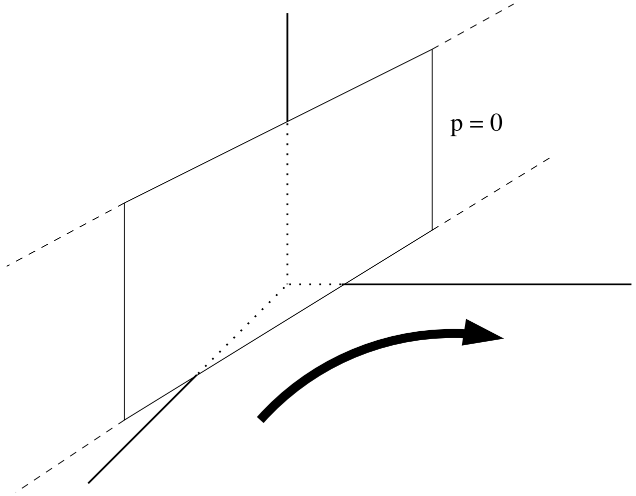
What does "goodness" translate to?

- If $a_i = 0$, then $\dot{f} = (-b_i/c_i)f$, i.e., $f(t) = f(0)e^{\frac{-c_i}{b_i}t}$
- If $b_i = 0$, then $\ddot{f} = (-c_i/a_i)f$, i.e., $f(t)$ oscillates between a minimum and a maximum value
- If $b_i < 0$, then f oscillates while exponentially decaying, still bounded

Illustration

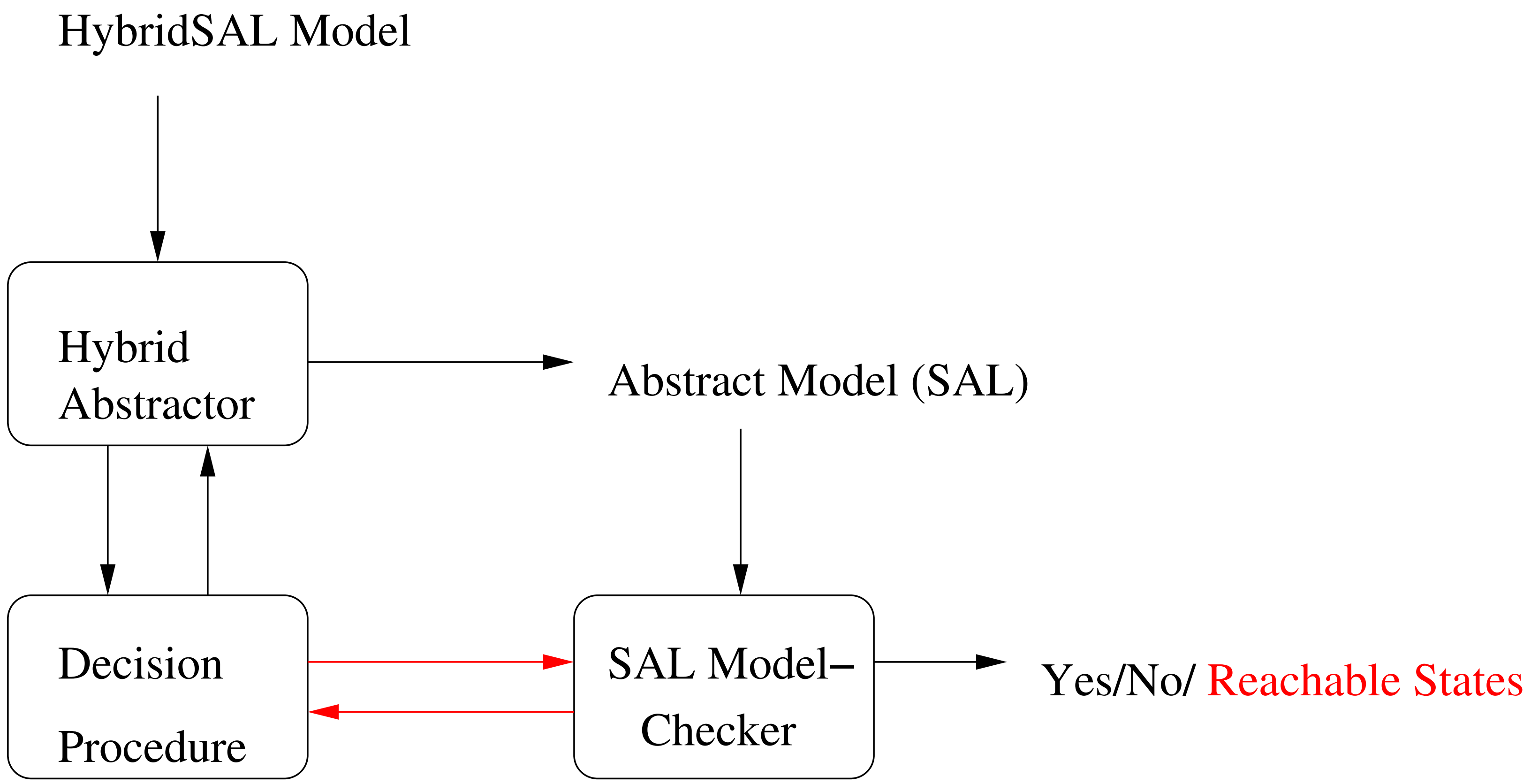
Let $[r_1, r_2, r_3]$ be an eigenvector corresponding to a negative real eigenvalue of a 3 dimensional linear system.

Let $p = r_1 x_1 + r_2 x_2 + r_3 x_3$.

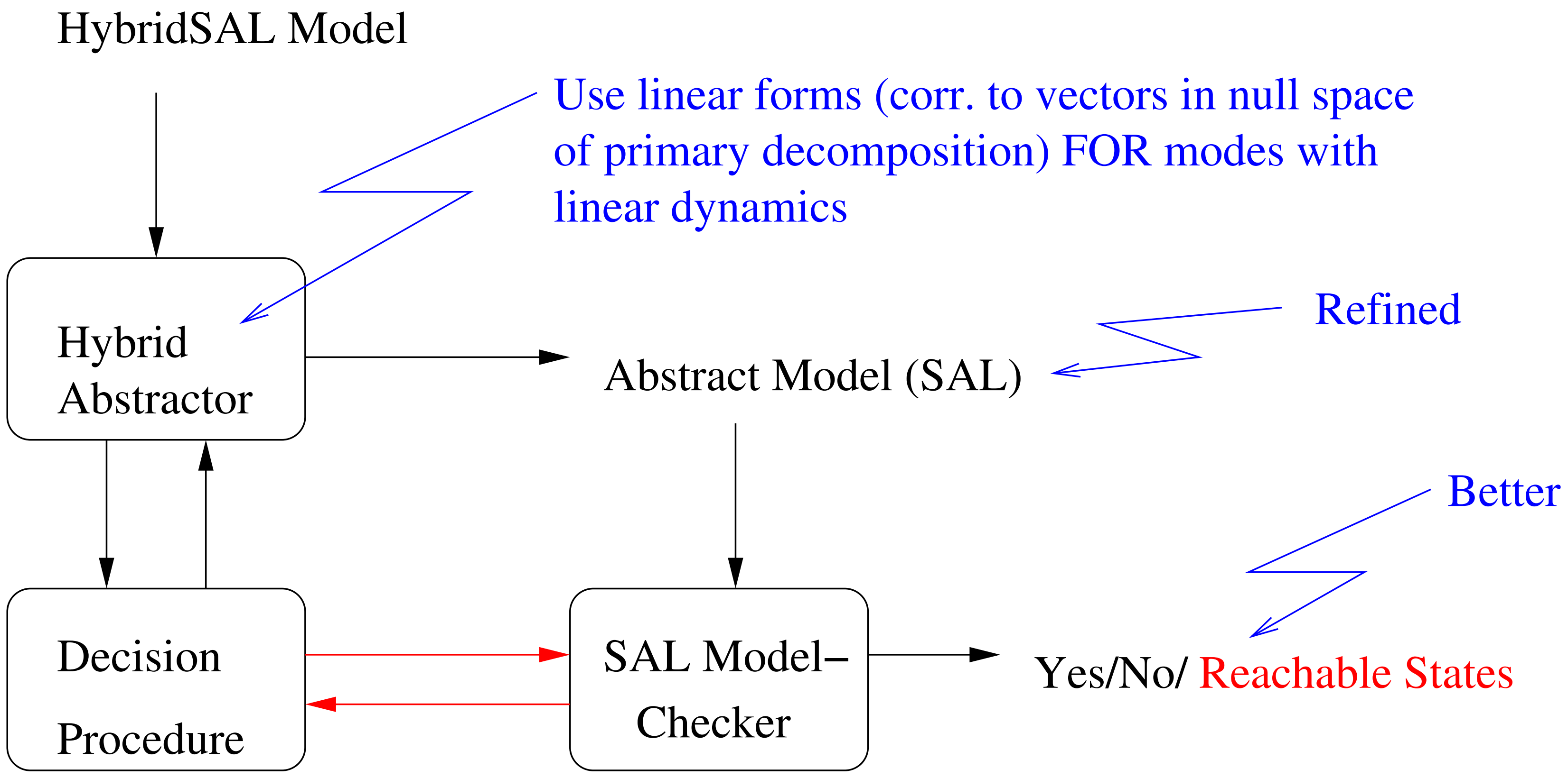


- If initial states are all on one side, then the systems continues to remain on the same side, i.e., always $p \geq 0$ or always $p \leq 0$
- At, say, $x_2 = 1$ and $x_3 = 2$, the value of x_1 can be bounded

Tool Architecture



Tool Architecture



Consequences

- Simple idea that can (and should!) be incorporated in all reachability tools
- Can do non-trivial reachability computation for linear systems with mixed eigenvalues
  - existing decidability results can not handle this class of systems
  - we extract as much information from the system as available, and bridge the gap between the decidable and undecidable problems
  - incrementally getting to the decidability result: the more eigenvalues are negative reals, the better reachability sets you generate
  - In the limit, do we get back the decidability results?

What is really going on?!

Linear system $\dot{\vec{x}} = A\vec{x}$

Say, $\vec{r_1}, \vec{r_2}, \ldots, \vec{r_n}$ are n-eigenvectors of A^T

Define new state variables $y_i \leftarrow \vec{r_i}^T \vec{x}$

Clearly, $\dot{\vec{y}} = D\vec{y}$, where $D = diag(d_1, \ldots, d_n)$.

Reachable states = $\exists t \geq 0.(\wedge_i y_i = y_{i0}e^{d_i t})$

<u>What is implicitly happening in our technique</u>?

Approx. reachable states = $\wedge_i(\exists t \geq 0.y_i = y_{i0}e^{d_i t})$

Approx. reachable states = $\wedge_i(0 < y_i < y_{i0})$ assuming $d_i < 0$

But all is not lost

Some of the lost information can be retrieved using more forms, i.e., refining the abstraction.

Eg. if $x = x_0 e^t$ and $y = y_0 e^{2t}$, then consider the quadratic form $x^2 - y$.

Its derivative, $\frac{d}{dt}(x^2 - y) = 2x\dot{x} - \dot{y} = 2x^2 - 2y = 2(x^2 - y)$.

This gives us (back!) the “connection” between the values of x and y.

This needs to be further formalized and automated.

Observations

Additional interesting consequence:

- Do not need to explicitly compute the real eigenvalues or the eigenvector $\vec{r}$
  - the eigenvalue and eigenvector are easily seen to be algebraic
  - symbolic decision procedures for real closed fields can handle the algebraic expressions representing these eigenvalues
- We assumed linear forms and linear dynamics
- Can we get useful partition forms for nonlinear dynamics using nonlinear forms?

Example: Cruise Control

Example. Consider a cruise control:

$$\begin{aligned} \dot{v} &= a \\ \dot{a} &= -4v + 3vf - 3a + gap \\ g\dot{a}p &= -v + vf \end{aligned}$$

where v, a is the velocity and acceleration of this car, vf is the velocity of car in front, and gap is the distance between the two cars.

Cruise Control: Example Contd

The transpose matrix A^T has one negative real eigenvalue λ.

If $\vec{r} = [r1, r2, r3, r4]^T$ is the eigenvector corresponding to λ, then consider the polynomial

$$p = r1 * v + r2 * vf + r3 * a + r4 * gap$$

Why is this special?

$$\frac{dp}{dt} = \lambda p$$

Since λ is negative, p is monotonically decreasing (to 0). This is sufficient to prove $gap > 0$ (for certain initial conditions and bounds on a, vf, v).

Future Work / Other topics

Composition—

- continuous state space partitioned differently in different modes
- each mode informs the other what it expects to know

Other interesting "seed" forms—

- symbolically integrate forms
- nonlinear systems
- nonlinear forms

Summary

- For abstracting modes with linear continuous dynamics, eigenvectors of the transpose matrix provide crucial "abstraction predicates"
  –if eigenvalue is rational, very easy to incorporate
  –if algebraic, need decision procedure for QF fragment of reals
  –generalizes to primary decomposition of A^T
- This observation integrates seamlessly with the qualitative hybrid abstraction technique presented in HSCC 2002
- Also provides incremental and practical "version" of the decidability results for reachability for linear vector fields