Oded Maler

CNRS-VERIMAG Grenoble, France

September 2006

Executive Summary

Describe our (me and colleagues) efforts over the last decade to push the capabilities of timed automata technology beyond toy problems

Try to justify the waste of such public resources and lifetimes by the importance of timed models, which goes much beyond the verification of real-time software (and verification in general).

With contributions of A. Pnueli, J. Sifakis, S. Yovine, E. Asarin, M. Bozga, C. Daws, S. Tripakis, Y. Abdeddaim, O. Bournez, M. Mahfoudh, P. Niebert, R. Ben Salah and S. Cotton

Partially sponsored by the European project AMETIST (Advanced Methods for Timed Systems, 2002-2005)



- Introduction: the importance of the timed level of abstraction
- A crash course in timed automata
- Attack 1: Numerical Decision Diagrams
- Attack 2: Timed Polyhedra
- Attack 3: Getting rid of Zones
- Attack 4: SAT
- Attack 5: Abstraction
- Attack 6: Interleaving
- Conclusions(?)

Levels of Abstraction in Dynamic Description

It is well known that the same phenomenon can be described at different levels of abstraction

The more detailed level should give better predictions but would be computationally harder to analyze (and will require more detailed observations).

The trick of science/math has always been to find the level which is sufficiently refined to give meaningful results and sufficiently abstract to be tractable computationally

Physics, chemistry, biology, physiology, psychology, sociology, economy, ...

From Grenoble to Nancy: Continuous View

Let $x = (x_1, x_2, x_3)$ be a real-valued vector representing the location of my center of mass in a coordinate system adapted to the surface of the earth

The trip is specified as a 3-dimensional signal x(t)



Such behaviors (signals, trajectories) are generated by differential equations (or hybrid automata)

Oded Maler

From Grenoble to Nancy: Discrete View

The trip is described as a sequence of states and transitions:

Grenoble
$$\xrightarrow{\text{bus}}$$
 Lyon $\xrightarrow{\text{plane}}$ Metz $\xrightarrow{\text{bus}}$ Nancy

Transitions are considered as atomic, instantaneous events

Such behaviors are generated by automata, transition systems, discreteevent systems, petri nets, process algebra, and worse

Sometimes we want to keep some of the continuous information, to express the fact that things take time

From Grenoble to Nancy: Timed View

The process of moving from one place to another is abstracted from it numerical details, but the **time** from initiation and termination is maintained

Grenoble
$$\xrightarrow{\text{bus}}$$
 on.bus $\xrightarrow{50}$ Lyon $\xrightarrow{\text{plane}}$ on.plane $\xrightarrow{70}$ Metz $\xrightarrow{\text{bus}}$ on.bus $\xrightarrow{25}$ Nancy



Mathematically Speaking

Discrete behaviors are viewed as **sequences** of events without metric timing information, only order or partial-order between the events.

A timed behavior involves the embedding of the sequence into the real time axis.

a, b, a, b, a, b, a, b



Timed Dynamical Systems

What is the appropriate dynamical system model for the intermediate timed level?

We do not need arbitrary continuous variables

We need discrete states that tell us where we are (in the abstract state space)

We need additional information that tell us how long we have been in this or that state

This additional information is encoded using "clock" variables

Timed Automata are *n***-Tuples...**

A timed automaton is $\mathcal{A} = (Q, C, I, \Delta)$ where...

The above is a sad fact that dooms timed automata into the formal verification circles and prevents it from being comprehensible to those who really need it

I'll try to avoid this as much as possible by giving intuitive explanations (hope you will not be offended)

Adding Time to Automata

Consider two processes that take 3 and 2 times units, respectively, after they start. We model the passage of 1 unit of time by a special tick transition.



Oded Maler

Possible Behaviors of the Processes



 P_1 waits one time unit and then starts:

$$\overline{p}_1 \xrightarrow{tick} \overline{p}_1 \xrightarrow{start1} 0 \xrightarrow{tick} 1 \xrightarrow{tick} 2 \xrightarrow{tick} 3 \xrightarrow{end1} \underline{p}_1$$

Oded Maler

The Two Processes in Parallel



Possible Joint Behaviors

Both processes start at time 2: $(\overline{p}_1, \overline{p}_2) \xrightarrow{tick} (\overline{p}_1, \overline{p}_2) \xrightarrow{tick} (\overline{p}_1, \overline{p}_2) \xrightarrow{start1} (0, \overline{p}_2) \xrightarrow{start2} (0, 0) \xrightarrow{tick} (1, 1) \xrightarrow{tick} (2, 2) \xrightarrow{end2} (2, \underline{p}_2) \xrightarrow{tick} (3, \underline{p}_2) \xrightarrow{end1} (\underline{p}_1, \underline{p}_2)$

 $P_{1} \text{ starts at } 0 \text{ and } P_{2} \text{ starts at } 2:$ $(\overline{p}_{1}, \overline{p}_{2}) \xrightarrow{start1} (0, \overline{p}_{2}) \xrightarrow{tick} (1, \overline{p}_{2}) \xrightarrow{tick} (2, \overline{p}_{2}) \xrightarrow{start2} (2, 0) \xrightarrow{tick} (3, 1) \xrightarrow{end1} (\underline{p}_{1}, 1) \xrightarrow{tick} (\underline{p}_{1}, 2) \xrightarrow{end2} (\underline{p}_{1}, \underline{p}_{2})$ $P_{2} \text{ starts at } 0 \text{ and } P_{1} \text{ starts after } P_{2} \text{ ends:}$ $(\overline{p}_{1}, \overline{p}_{2}) \xrightarrow{start2} (\overline{p}_{1}, 0) \xrightarrow{tick} (\overline{p}_{1}, 1) \xrightarrow{tick} (\overline{p}_{1}, 2) \xrightarrow{end2} (\overline{p}_{1}, \underline{p}_{2}) \xrightarrow{start1} (0, \underline{p}_{2}) \xrightarrow{tick} (1, \underline{p}_{2}) \xrightarrow{tick} (2, \underline{p}_{2}) \xrightarrow{tick} (3, \underline{p}_{2}) \xrightarrow{end1}$

 $(\underline{p}_1, \underline{p}_2)$

Interleaving:

 $(\overline{p}_1, \overline{p}_2) \stackrel{start1}{\longrightarrow} (0, \overline{p}_2) \stackrel{start2}{\longrightarrow} (0, 0) = (\overline{p}_1, \overline{p}_2) \stackrel{start2}{\longrightarrow} (\overline{p}_2, 0) \stackrel{start1}{\longrightarrow} (0, 0)$

Using Clock Variables





Oded Maler

Clock Variables: the Composition



The Notion of a State

Warning: in automata augmented with variables, the **state** is encoded in both the discrete state (location) and the values of the variables.

The merging into (p_1, p_2) is misleading: via different paths you reach different clock valuations.



The Joy of Clock Variables

They allow succinct and natural representation of the system.

Transitions are labeled by guards and resets.

Different clocks represent the time elapsed since certain events.

In the worst-case, however, one needs to expand the automaton by adding clock values to states.

You can use **symbolic** rather than enumerative encoding of the set of reachable states.

You can work in **dense** time without committing a-priori to time granularity.

Symbolic Representation

Assume the two processes with durations d_1 and d_2 such that $d_1 < d_2$ and that p_2 starts 2 time units after p_1 .



The set of clock values that can be reached at state (p_1, p_2) is $\{(2,0), (3,1), (4,2), \dots, (d_1, d_1 - 2)\}$ and its size depends on d_1 .

It can be, however, represented by a fixed size formula $X_1 - X_2 = 2 \wedge X_1 \leq d_1$

From Discrete to Dense Time

So far we have assumed a fixed time granularity Δ associated with a *tick*.

Discrete time flows in Δ quanta by the *tick* transitions. These transitions induce self-loops on the states of all automata.

Other transitions can be taken only at time points $n\Delta$, $n \in \mathbb{N}$.

By considering clocks as continuous variables we can use time-passage of **arbitrary** length.

Time passage, instead of being represented by tick transitions, can be modeled by all active clocks advancing with **derivative** 1 when the automaton stays in a state.

The timed automaton is viewed as a simple kind of a hybrid automaton whose evolution alternates between passage of time and discrete transitions.

The Two Processes as Two Timed Automata

 $\dot{x}_{2} = 1$





Modeling Temporal Uncertainty

The major strength of timed automata is their ability to express **temporal uncertainty**.

"The duration of a task (or the distance between two events) is somewhere in the interval [l, u]"

Using dense time this means **anywhere** in [l, u] not just l or u

Verification can be done with respect to **all** choices of values in the interval

This CS non-determinism is an alternative/complement to probabilistic modeling of uncertainty (for example exponential distribution of durations)

Modeling Temporal Uncertainty with TA

There are different ways to model urgency/non-urgency in TA:

1) Invariants (staying conditions) that the clocks must satisfy in order to remain in a state and "let" time progress.

2) Deadlines on transitions.

Example: a task whose duration is between 3 and 7 time "units":



$$\begin{split} \overline{p}, \bot) &\xrightarrow{\Box i \ominus} (\overline{p}, \bot) \xrightarrow{\Box i \ominus} (p, 0) \xrightarrow{\Box i \ominus} (p, 3.8) \xrightarrow{\Box i \ominus} (\underline{p}, \bot) \\ (\overline{p}, \bot) &\xrightarrow{t_1} (\overline{p}, \bot) \xrightarrow{start} (p, 0) \xrightarrow{t_2} (p, t_2) \xrightarrow{end} (\underline{p}, \bot) \\ & t_1 \in [0, \infty), t_2 \in [3, 7]. \end{split}$$

Verification (Reachability) of Timed Automata





Timed Automata are *n***-Tuples...**

A timed automaton is $\mathcal{A} = (Q, C, I, \Delta) Q$: a set of states, C: a set of clocks,

I: staying condition (invariant), assigning to every q a conjunction I_q of inequalities of the form $c \leq u$, for some clock c and integer u

 Δ : a transition relation consisting of tuples (q,ϕ,ρ,q') where q and q' are states,

 $\rho \subseteq C$ is the set of clocks reset by the transition, and

 ϕ (the transition guard) is a conjunction of formulae of the form $c \ge l$ for some clock c and integer l.

A clock valuation is a function $\mathbf{v} : C \to \mathbb{R}_+ \cup \{0\}$ and a configuration is a pair (q, \mathbf{v}) consisting of a discrete state (location) and a clock valuation.

Runs of Timed Automata

A step of the automaton is one of the following:

- A discrete step: $(q, \mathbf{v}) \xrightarrow{\delta} (q', \mathbf{v}')$, for some transition $\delta = (q, \phi, \rho, q') \in \Delta$, such that \mathbf{v} satisfies ϕ and $\mathbf{v}' = R_{\rho}(\mathbf{v})$.
- A time step: $(q, \mathbf{v}) \xrightarrow{t} (q, \mathbf{v} + t\mathbf{1}), t \in \mathbb{R}_+$ such that $\mathbf{v} + t\mathbf{1}$ satisfies I_q .

A run of the automaton starting from a configuration (q_0, \mathbf{v}_0) is a finite sequence of steps

$$\xi: \quad (q_0, \mathbf{v}_0) \xrightarrow{t_1} (q_1, \mathbf{v}_1) \xrightarrow{t_2} \cdots \xrightarrow{t_n} (q_n, \mathbf{v}_n).$$

Symbolic Reachability Computation

A symbolic state is (q, Z) where q is a discrete state and Z is a zone, a set of clock valuations satisfying a conjunction of inequalities $c_i - c_j \ge d$ or $c_i \ge d$.

Symbolic states are closed under the following operations:

• The time successor of (q, Z), the configurations reachable from (q, Z) by letting time progress without violating the staying condition of q:

 $Post^{t}(q, Z) = \{(q, \mathbf{Z} + r\mathbf{1}) : \mathbf{Z} \in Z, r \ge 0, \mathbf{Z} + r\mathbf{1} \in I_q\}$

• The δ -transition successor of (q, Z) is the configurations reachable from (q, Z) by taking the transition $\delta = (q, \phi, \rho, q') \in \Delta$:

 $Post^{\delta}(q, Z) = \{(q', R_{\rho}(\mathbf{Z})) : \mathbf{Z} \in Z \cap \phi\}$

• The δ -successor of a time-closed symbolic state (q, Z) is the set of configurations reachable by a δ -transition followed by passage of time:

$$Succ^{\delta}(q,Z) = Post^{t}(Post^{\delta}(q,Z))$$

The Reachability Graph

The basic verification algorithm for TA consists of on-the-fly generation of the reachability (simulation) graph, $S = (N, \rightarrow)$

The nodes are symbolic states computed starting from $Post^t(s, \{0\})$ and applying $Succ^{\delta}$ until termination (guaranteed due to finitely-many zones)

There is a path from (q, Z) to (q', Z') in *S* iff for every $\mathbf{v}' \in Z'$ there exists $\mathbf{v} \in Z$ and a run of \mathcal{A} from (q, \mathbf{v}) to (q', \mathbf{v}') .

Hence the union of all symbolic states in S is exactly the set of reachable configurations.

This is the computation we want to do more efficiently

The Sources of Difficulty

Assume we have n interacting timed automata, each with m states and one clock ranging over [0, d]

The number of states can be up to m^n and the number of zones can be up to $d^n n!$, summing up to $m^n d^n n!$ symbolic states. Each zone takes $O(n^2)$ space

The representation of (convex) zones is fine but there is no nice representation for a union of zones and, even worse, the representation is not symbolic for the discrete states: symbolic states are of the form (q, Z) with q being an explicit n-vector.

Since our our initial motivation came from circuits where the number of discrete states explodes very quickly, we tried BDD-based methods first

BDD: The Principles

Sets of states can be expressed as formulae over the state variables; The transition relation can be expressed this way as well

Based on that you can do breadth-first exploration of the reachable sets, computing a sequence of sets P_0, P_1, \ldots such that P_i consists of sets reachable from P_0 by at most *i* steps

You don't care about disjunctions/non-convexity, everything is a formula

OBDDs provide for a canonical representation of these sets/formulae; If you are lucky they are more succinct than the sets they represent

This is the naive story, there are many details but it seems to work to a certain extent in hardware.

Attack 1: Numerical Decision Diagrams (A. Pnueli, M. Bozga 95-97)

The idea: to have a BDD-like formalism for representing sets of configurations, as formulae of the form $x_1 \wedge c_1 > 3 \wedge (\neg x_2 \vee c_2 < 7)$. The *Succ* operator will be applied to this representation.

First direction: use inequalities of the form $c_i < d$ as nodes in the BDD. The problem is that unlike Boolean variables x_i and x_j which are independent, conditions $c_i < d$ and $c_i < d'$ are not

After some playing we came to the conclusion that if we want canonicity we need to use variables for all the bits in the binary representation of the clock value

Attack 1: Numerical Decision Diagrams (A. Pnueli, M. Bozga 95-97)

A discrete clock range $[0, \ldots, d-1]$ can be encoded using $\log d$ Boolean variables Any subset of these values can be expressed as a Boolean formula over these variables. Adding the state variables we have a canonical representation of sets of configurations

Passage of time is computed as binary addition (or transitive closure of incrementation)



Attack 1: Numerical Decision Diagrams (A. Pnueli, M. Bozga 95-97)

More technical details about variable ordering (bits of clock near the bits of the corresponding state variables, etc.)

Results: managed to verify the STARI circuit 55 clocks and about 2^{18} states

Did not work so good for other cases, sensitivity to the range of the clocks (the number of zones is also sensitive but less)

General problem: binary positional encoding of numbers breaks the topological structure (the Hamming distance between 01111 and 10000 is large while the numbers are close)

Lessons: BDDs are no magic, discrete time is good for many purposes [Asarin Pnueli 98], life is hard

Farn Wang and Dirk Beyer continued to work in this direction

Attack 2: Timed Polyhedra (O. Bournez, M. Mahfoudh 98-00)

Background: still obsessed with the idea of canonical representation of non-convex subsets of \mathbb{R}^n (also for the context of hybrid systems verification)

For griddy (orthogonal, isothetic) polyhedra we found a canonical representation as a XOR of rectangular cones based on some vertices of the polyhedron

Wanted to extend them to timed polyhedra, constructed from the following building blocks



Attack 2: Timed Polyhedra (O. Bournez, M. Mahfoudh 98-00)

The good news: there is a similar canonical representation based on XOR of timed cones (ICALP'00)

The bad news: the representation is enumerative in the cone types; To represent a set satisfying $x_1 < x_2$ you need to specify it as $x_3 < x_1 < x_2 \lor x_1 < x_3 < x_2 \lor x_1 < x_2 < x_3$. Also the number of vertices grows badly with dimension

We tried some symbolic representation with BDD-like structures, but nothing to write home about in performance

Lessons: not all that glitters is gold, maybe the idea of canonical representation and BFS is not always good

Attack 3: No Zones (Y. Abdeddaim, 98-00)

As mentioned earlier, timed automata exhibit dense non-determinism: a transition can be taken at any point in an interval [l, u]

In verification, where the non-determinism is associated with the external uncontrolled world, we need to take all these choices into consideration

In synthesis/optimization where the choice of when to take a transition depends on us, sometimes we need not consider the whole interval but only some points in it that "dominate" the others

This turned out to be the case in optimal scheduling problems where it is sufficient to consider only a small subset of the runs

Deterministic Job-Shop Scheduling: the Problem

$$J^1: (m_1, 4), (m_2, 5) \qquad J^2: (m_1, 3)$$

Determine the execution times of the tasks such that:

The termination time of the last task is minimal

Precedence and resource constraints are satisfied



Sometimes it is better not to start a task although the machine is idle



Modeling with Timed Automata

Each automaton represents the set of all possible behaviors of each task/job in isolation (respecting the precedence constraints)

The Start transitions are issued by the controller/scheduler and the End transitions by the environment

Oded Maler

The Global Automaton

Resource constraints expressed via forbidden states in the product automaton



Optimal scheduling = shortest path problem for timed automata

Finding the Shortest Path

Add an additional clock T which is never reset to zero, hence it measures the absolute time since the beginning

Naive approach: perform zone-based reachability computation on the extended clock space (the graph is acyclic and all paths lead to the final state); Find the minimal value of T over all symbolic states associated with the final state

However, it can be shown that postponing a *start* transition from t to t' is useless if the machine is used by anyone else during [t, t']

Hence the optimum can be found among a finite number of schedules/runs where a transition not taken in a state at the first moment it was enabled will not be taken at that state at all

Oded Maler

Attack 3: No Zones (Y. Abdeddaim, 98-00)



Lessons: there is life after operations research

Attack 4: SAT and Bounded Verification

(P. Niebert, E. Asarin, M. Mahfoudh S. Cotton, 00-06)

Verification for bounded horizon (BMC) is based on a very simple idea. The existence of a run of length k from initial set P to a bad set B can be formulated using a k-unfolding of the transition relation R:

 $\exists x_0, \ldots, x_k P(x_0) \land R(x_0, x_1) \land R(x_1, x_2) \cdots \land R(x_{k-1}, x_k) \land B(x_k)$

The existence of such an assignment can be checked by a constraint solver for the domain. For finite-state systems this reduces to Boolean SAT.

We have shown that for timed automata, path existence can be formulated in **difference logic**, propositional logic plus constraints of the form x - y < c the basic logic for timing issues (distance between events)

Oded Maler

Attack 4: SAT and Bounded Verification

(P. Niebert, E. Asarin, M. Mahfoudh S. Cotton, 00-06)

We (and others) have developed several SAT solvers for this logic using a variety of methods (reduction to SAT, lazy, eager, mixed, preprocessing)

This domain is called today satisfiability modulo theories (SMT)

Our solvers have improved with the years and can solve some really hard problems

We have learned a new fascinating domain

But we never managed to solve even a modest bounded model checking problems for timed automata. A fundamental folk wisdom says that this holds for all asynchronous system

Attack 5: Abstraction (R. Ben Salah, M. Bozga, 02-06)

Principle is simple: the system $S = S_1 ||S_2|| \cdots ||S_n|$ is made of components whose product explodes

Replace each (or some) S_i by and S'_i such that $S'_i < S_i$ in syntax and $S'_i > S_i$ in semantics

Correctness of $S' = S'_1 ||S'_2|| \cdots ||S'_n$ implies correctness of S and may be computationally easier

We developed an automatic methodology to create such abstractions, specialized (bot not restricted to) Boolean circuits with delays

Oded Maler

Circuits with Bi-bounded Inertial Delays



Modeling Circuits with Timed Automata

Our modeling approach, based on [Maler and Pnueli 95]: Decompose any gate into an instantaneous Boolean function and a bi-bounded (non-deterministic) inertial delay element

Model every delay element as a timed automaton with 4 states and 1 clock



Composing all these automata we obtain a timed automaton with $O(2^n)$ states and n clocks

Abstraction of Acyclic Circuits

Start with a stable states, primary inputs change only once at start. This induces a noncountable number of possible behaviors

Each behavior admits a finite number of changes and stabilizes in a bounded amount of time. We want to compute the maximal stabilization time, that of the worst behavior

The basic idea: take a sub-circuit on the left, use TA technology to generate an approximate timed model of its output. It is then plugged as an input model to the rest of the circuit.



The Reachability Graph

The reachability graph of a timed automaton can be viewed as an "interpretation" of the automaton:

On on one hand we split some discrete states according to clock values

On the other, we remove transitions that are infeasible due to timing constraints.

By associating with each symoblic state (q, Z) the staying condition Z and with each outgoing transition the intersection of Z with the guard we obtain a TA equivalent to the original one where all states are reachable from the initial state.

The abstraction is done by applying certain transformation to this timed automaton

Oded Maler





The Nature of the Abstraction

First, the obvious thing: hiding internal actions such as excitation and "regrets" of the outputs and all transitions of internal wires.

Relaxation of timing constraints by allowing things to happen at impossible times (but not in impossible orders!)

We project the TA obtained from the reachability graph on a subset of the clocks. The constraints related to the other clocks are removed.

For acyclic circuits it is natural to project only on the auxiliary clock T that measures absolute time. This way we keep the information about the time each transition can be taken (but lose some inter-dependence information).



Oded Maler





Minimization

After minimization we obtain the following small-description abstraction for the observed behavior of the circuit:





Attack 5: Abstraction (R. Ben Salah, M. Bozga, 02-06)

Current status: for acyclic circuits we could treat (under certain choice of parameters that keep the ratio u/(u - l) low) a cascade of up to 22 4-gate circuits.

Still a far cry from static methods used in industry

We have developed a very interesting novel method for abstracting open timed components (the inputs may arrive anytime, not only in time zero)

Unfortunately, the size of the basic component that could be analyzed and abstracted was too small to be useful

Looking for the reasons for that has led us to the last discovery conerning interleaving and convexity

Attack 6: Interleaving (R. Ben Salah, M. Bozga, 06)

There is an additional explosion in TA reachability due to interleaving. At the end of a "diamond" you have two zones: one with $x \leq y$ and one with $y \leq x$



Attack 6: Interleaving (R. Ben Salah, M. Bozga, 06)

Given a run ξ of a timed automaton, we denote by $\langle \xi \rangle$ all runs that make the same transitions (but possibly in another order). In other words, all runs that their local projections do the same transitions as those of ξ

The following result (CONCUR'06) helps to avoid this explosion:

Let Z be a convex timed polyhedron and let q and q' be two global states of \mathcal{A} . Let ξ be a run starting at q and ending in q'. Then the set

$$R_{Z,\langle\xi\rangle} \equiv \bigcup_{\xi'\in\langle\xi\rangle} \{\mathbf{v}': \exists \mathbf{v}\in Z \ (\mathbf{q},\mathbf{v}) \xrightarrow{\xi'} (\mathbf{q}',\mathbf{v}')\}$$
 is convex.

Remark: this result turned out to be implicit in [Rockiki, Myers 94], [Zhao 02] and [Lugiez, Niebert, Zenou 05]





The graph generated by the standard reachability algorithm.









A New Reachability algorithm

Anotate symbolic states with (partially-ordered) path information

Do BFS exploration; Whenever two symblic states have the same set of labels, merge them by taking their convex hull

This way explosion is killed when still small

The results are guaranteed to be exact

Oded Maler





Oded Maler





Oded Maler





Oded Maler





Oded Maler





Oded Maler





Oded Maler

A New Reachability algorithm





Oded Maler

A New Reachability algorithm





Oded Maler

Comparison





Oded Maler

Interim Summary

The road is long

Next hope, to combine the interleaving reduction with the abstraction, hopefully this year

Thank you