Efficient Timed Reachability Analysis using Clock Difference Diagrams

Gerd Behrmann
Kim G. Larsen
Justin Pearson
Carsten Weise
Wang Yi

BRICS Report Series RS-98-47

ISSN 0909-0878 December 1998

Efficient Timed Reachability Analysis using Clock Difference Diagrams

Gerd Behrmann[1] Kim G. Larsen[1] Justin Pearson[2] Carsten Weise[1]
Wang Yi[2]
[1]BRICS*, Aalborg University, Denmark
[2]Department of Computer Systems, Uppsala University, Sweden

Abstract

One of the major problems in applying automatic verification tools to industrial-size systems is the excessive amount of memory required during the state-space exploration of a model. In the setting of real-time, this problem of state-explosion requires extra attention as information must be kept not only on the discrete control structure but also on the values of continuous clock variables.

In this paper, we present Clock Difference Diagrams, CDD's, a BDD-like data-structure for representing and effectively manipulating certain non-convex subsets of the Euclidean space, notably those encountered during verification of timed automata.

A version of the real-time verification tool UPPAAL using CDD's as a compact data-structure for storing explored symbolic states has been implemented. Our experimental results demonstrate significant space-savings: for 8 industrial examples, the savings are between 46% and 99% with moderate increase in runtime.

We further report on how the symbolic state-space exploration itself may be carried out using CDD's.

1 Motivation

In the last few years a number of verification tools have been developed for real-time systems (e.g. [HHW95, DY95, BLLPW96]). The verification engines of most tools in this category are based on reachability analysis of timed automata following the pioneering work of Alur and Dill [AD94]. A timed automaton is an extension of a finite automaton with a finite set of real-valued clock-variables. Whereas the initial decidability results are based on a partitioning of the infinite state-space of a timed automaton into finitely many equivalence classes (so-called *regions*), tools such as KRONOS and UPPAAL are based on more efficient data structures and algorithms for representing and manipulating timing constraints over clock variables. The abstract reachability algorithm applied in these tools is shown in Figure 1. The algorithm checks whether a timed automaton may reach a state satisfying a given state formula ϕ. It explores the state space of the automaton in terms of *symbolic states* of the form (l, D), where l is a control–node and D is a constraint system over clock variables $\{X_1, \ldots, X_n\}$. More precisely, D consists of a conjunction of simple clock constraints of the form $X_i \, op \, c$, $-X_i \, op \, c$ and $X_i - X_j \, op \, c$, where c is an integer constant and $op \in \{<, \leq\}$. The subsets of $\mathbb{R}^n$ which may be described by clock constraint systems are called *zones*. Zones are convex polyhedra, where all edge-points are integer valued, and where border lines may or may not belong to the set (depending on a constraint being strict or not).

*BRICS: Basic Research in Computer Science, Centre of the Danish National Research Foundation

```
PASSED:= {}
WAIT:= {($l_0, D_0$)}
repeat
        begin
            get $(l, D)$ from WAIT
            if $(l, D) \models \phi$ then return "YES"
            else if $D \not\subseteq D'$ for all $(l, D') \in$ PASSED then
                    begin
                        add $(l, D)$ to PASSED          $(*)$
                        NEXT:=$\{(l_s, D_s) : (l, D) \rightsquigarrow (l_s, D_s) \wedge D_s \neq \emptyset\}$
                        for all $(l_{s'}, D_{s'})$ in NEXT do
                                put $(l_{s'}, D_{s'})$ to WAIT
                    end
        end
until WAIT={}
return "NO"
```

Figure 1: An algorithm for symbolic reachability analysis.

We observe that several operations of the algorithm are critical for efficient implementation. In particular the algorithm depends heavily on operations for checking set inclusion and emptiness. In the computation of the set NEXT, operations for intersection, forward time projection (future) and projection in one dimension (clock reset) are required. A well-known data-structure for representing clock constraint systems is that of *Difference Bounded Matrices*, DBM, [Dill87], giving for each pair of clocks[1] the upper bound on their difference. All operations required in the reachability analysis in Figure 1 can be easily implemented on DBM's with satisfactory efficiency. In particular, the various operations may benefit from a *canonical* DBM representation with tightest bounds on all clock differences computed by solving a shortest path problem. However, computation of this canonical form should be postponed as much as possible, as it is the most costly operation on DBM's with time-complexity $O(n^3)$ (n being the number of clocks).

DBM's obviously consume space of order $O(n^2)$. Alternatively, one may represent a clock constraint system by choosing a minimal subset from the constraints of the DBM in canonical form. This *minimal form* [LPW95] is preferable when adding a symbolic state to the main global data-structure PASSED as in practice the space-requirement is only linear in the number of clocks.

Considering once again the reachability algorithm in Figure 1, we see that a symbolic state (l, D) from the waiting-list WAIT is freed from being explored (the inner box) provided some symbolic state (l, D') already in PASSED 'covers' it (i.e. $D \subseteq D'$). Though clearly a sound rule and provably sufficient for termination of the algorithm, exploration of (l, D) may be avoided under less strict conditions. In particular, it suffices for (l, D) to be 'covered' collectively by the symbolic states in PASSED with location l, i.e.:

$$D \subseteq \bigcup \{D' \mid (l, D') \in \text{PASSED}\} \qquad (1)$$

However, this requires handling of unions of zones, which complicates things considerably. Using DBM's, finite unions of zones – which we will call *federations* in the following – may be represented by a list of all the DBM's of the union. However, the more "non-convex" the zone becomes, the more DBM's will be needed. In particular, this representation makes the inclusion-check of (1) computational expensive.

[1]For uniformity, we assume a special clock X_0 which is always zero. Thus $X_i \, op \, c$ and $-X_i \, op \, c$ can be rewritten as the differences $X_i - X_0 \, op \, c$ and $X_0 - X_i \, op \, c$.

In this paper, we introduce a more efficient BDD-like data-structure for federations, *Clock Difference Diagrams*, CDD's. A CDD is a directed acyclic graph, where inner nodes are associated with a given pair of clocks and outgoing arcs state bounds on their difference. This data-structure contains DBM's as a special case and offers simple boolean set-operations and easy inclusion- and emptiness-checking. Using CDD's, the PASSED-list may be implemented as a collection of symbolic states of the form (l, F), where F is a CDD representing the union of all zones for which the location l has been explored[2]. Thus, the more liberal termination condition of (1) may be applied, potentially leading to faster termination of the reachability algorithm. As any BDD-like data-structure, CDD's eliminate redundancies via sharing of substructures. Thus, the CDD representation of F is likely to be much smaller than the explicit DBM-list representation. Furthermore, sharing of identical substructures between CDD's from *different* symbolic states may be obtained for free, opening for even more efficient storage-usage.

Having implemented a CDD-package and used it in modifying UPPAAL, we report on some very encouraging experimental results. For 8 industrial examples found in the literature, significant space-savings are obtained: the savings are between 46% and 99% with moderate increase in run-time (in average an increase of 17%).

To make the reachability algorithm of Figure 1 fully symbolic, it remains to show how to compute the successor set NEXT based on CDD's. In particular, algorithms are needed for computing forward projection in time and clock-reset for this data-structure. Similar to the canonical form for DBM's these operation are obtained via a *canonical* CDD form, where bounds on all arcs are as tight as possible.

Related Work

The work in [Bal96] and [WTD95] represent early attempts of applying BDD-technology to the verification of continuous real-time systems. In [Bal96], DBM's themselves are coded as BDD's. However, unions of DBM's are avoided and replaced by convex hulls leading to an approximation algorithm. In [WTD95], BDD's are applied to a symbolic representation of the discrete control part, whereas the continuous part is dealt with using DBM's.

The Numerical Decision Diagrams of [ABKMPR97, BMPY97] offer a canonical representation of unions of zones, essentially via a BDD-encoding of the collection of regions covered by the union. [CC95] offers a similar BDD-encoding in the simple case of one-clock automata. In both cases, the encodings are extremely sensitive to the size of the in-going constants. As we will indicate, NDD's may be seen as degenerate CDD's requiring very fine granularity.

CDD's are in the spirit of Interval Decision Diagrams of [ST98]. In [Strehl'98], IDD's are used for analysis in a discrete, one-clock setting. Whereas IDD's nodes are associated with independent real-valued variables, CDD-nodes – being associated with differences – are highly dependent. Thus, the subset- and emptiness checking algorithms for CDD's are substantially different. Also, the canonical form requires additional attention, as bounds on different arcs along a path may interact.

The CDD datastructure was first introduced in [LPWW98], where a thorough study of various possible normalforms is given.

2 Timed Automata

Timed automata were first introduced in [AD94] and have since then established themselves as a standard model for real–time systems. We assume familiarity with this model and only give a brief review in order to fix the terminology and notation used in this paper.

[2]Thus D is simply unioned with F, when a new symbolic state (l, D) is added to the PASSED-list (cf. Fig. 1, line $(*)$).

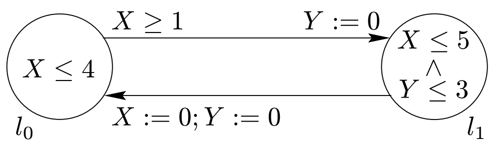

Figure 2: A Timed Automaton.

Consider the timed automaton of Figure 2. It has two control nodes l_0 and l_1 and two real–valued clocks X and Y. A *state* of the automaton is of the form (l, s, t), where l is a control node, and s and t are non–negative reals giving the value of the two clocks X and Y. A control node is labelled with a condition (the invariant) on the clock values that must be satisfied for states involving this node. Assuming that the automaton starts to operate in the state $(l_0, 0, 0)$, it may stay in node l_0 as long as the invariant $X \leq 4$ of l_0 is satisfied. During this time the values of the clocks increase synchronously. Thus from the initial state, all states of the form (l_0, t, t), where $t \leq 4$, are reachable. The edges of a timed automaton may be decorated with a condition (guard) on the clock values that must be satisfied in order to be enabled. Thus, only for the states (l_0, t, t), where $1 \leq t \leq 4$, is the edge from l_0 to l_1 enabled. Additionally, edges may be labelled with simple assignments reseting clocks. E.g. when following the edge from l_0 to l_1 the clock Y is reset to 0 leading to states of the form $(l_1, t, 0)$, where $1 \leq t \leq 4$.

A timed automaton is a standard finite-state automaton extended with a finite collection of real-valued clocks $C = \{X_1, \ldots, X_n\}$. We use $\mathcal{B}(C)$ ranged over by g and D to denote the set of clock constraint systems over C.

Definition 1 *A timed automaton A over clocks C is a tuple $\langle N, l_0, E, Inv\rangle$ where N is a finite set of nodes (control-nodes), l_0 is the initial node, $E \subseteq N \times \mathcal{B}(C) \times 2^C \times C$ corresponds to the set of edges, and finally, $Inv : N \rightarrow \mathcal{B}(C)$ assigns invariants to nodes. In the case, $\langle l, g, r, l'\rangle \in E$, we write $l \xrightarrow{g,r} l'$.*

Formally, we represent the values of clocks as functions (called clock assignments) from C to the non–negative reals $\mathbb{R}_{\geq}$. We denote by $\mathcal{V}$ the set of clock assignments for C. A semantical *state* of an automaton A is now a pair (l, u), where l is a node of A and u is a clock assignment for C, and the semantics of A is given by a transition system with the following two types of transitions (corresponding to delay–transitions and edge–transitions):

- $(l, u) \longrightarrow (l, u + d)$ if $Inv(l)(u)$ and $Inv(l)(u + d)$
- $(l, u) \longrightarrow (l', u')$ if there exist g, r such that $l \xrightarrow{g,r} l'$, $u \in g$, $u' = [r \mapsto 0]u$, $Inv(l)(u)$ and $Inv(l')(u')$

where for $d \in \mathbb{R}_{\geq}$, $u + d$ denotes the time assignment which maps each clock X in C to the value $u(X) + d$, and for $r \subseteq C$, $[r \mapsto 0]u$ denotes the assignment for C which maps each clock in r to the value 0 and agrees with u over $C \backslash r$. By $u \in g$ we denote that the clock assignment u satisfies the constraint g (in the obvious manner).

Clearly, the semantics of a timed automaton yields an infinite transition system, and is thus not an appropriate basis for decision algorithms. However, efficient algorithms may be obtained using a finite–state *symbolic* semantics based on *symbolic states* of the form (l, D), where $D \in \mathcal{B}(C)$ [HNSY94, YPD94]. The symbolic counterpart to the standard semantics is given by the following two (fairly obvious) types of symbolic transitions:

- $(l, D) \rightsquigarrow \Big(l, (D \wedge Inv(l))^{\uparrow} \wedge Inv(l)\Big)$
- $(l, D) \rightsquigarrow \Big(l', r(g \wedge D \wedge Inv(l)) \wedge Inv(l')\Big)$ if $l \xrightarrow{g,r} l'$

where time progress $D^{\uparrow} = \{u + d \,|\, u \in D \wedge d \in \mathbb{R}_{\geq}\}$ and clock reset $r(D) = \{[r \mapsto 0]u \,|\, u \in D\}$. It may be shown that $\mathcal{B}(C)$ (the set of constraint systems) is closed under these two operations ensuring the well–definedness of the semantics. Moreover, the symbolic semantics corresponds closely to the standard semantics in the sense that, whenever $u \in D$ and $(l, D) \rightsquigarrow (l', D')$ then $(l, u) \longrightarrow (l', u')$ for some $u' \in D'$.

3 Clock Difference Diagrams

While in principal DBM's are an efficient implementation for clock constraint systems, especially when using canonical form only when necessary and minimal form when suitable, they are not very good athandling unions of zones. In this section we will introduce a more efficient data structure for federations: *clock difference diagrams* or short CDD's. A CDD is a directed acyclic graph with two kinds of nodes: inner nodes and terminal nodes. Terminal nodes represent the constants true and false, while inner nodes are associated with a *type* (i.e. a clock pair) and arcs labeled with intervals giving bounds on the clock pair's difference. Figure 3 shows examples of CDD's.

A CDD is a compact representation of a decision tree for federations: take a valuation, and follow the unique path along which the constraints given by type and interval are fulfilled by the valuation. If this process ends at a true node, the valuation belongs to the federtaion represented by this CDD, otherwise not. A CDD itself is not a tree, but a DAG due to sharing of isomorphic subtrees.

A *type* is a pair (i, j) where $1 \leq i < j \leq n$. The set of all types is written $\mathcal{T}$, with typical element t. We assume that $\mathcal{T}$ is equipped with a linear ordering $\sqsubseteq$ and a special bottom element $(0, 0) \in \mathcal{T}$, in the same way as BDD's assume a given ordering on the boolean variables. By $\mathcal{I}$ we denote the set of all non-empty, convex, integer-bounded subsets of the real line. Note that the integer bound may or may not be within the interval. A typical element of $\mathcal{I}$ is denoted I. We write $\mathcal{I}_{\emptyset}$ for the set $\mathcal{I} \cup \{\emptyset\}$.

In order to relate intervals and types to constraint, we introduce the following notation:

- given a type (i, j) and an interval I of the reals, by $I(i, j)$ we denote the clock constraint having type (i, j) which restricts the value of $X_i - X_j$ to the interval I.
- given a clock constraint D and a valuation v, by $D(v)$ we denote the application of D to v, i.e. the boolean value derived from replacing the clocks in D by the values given in v.

Note that typically we will use the notation jointly, i.e. $I(i, j)(v)$ expresses the fact that v fulfills the constraint given by the interval I and the type (i, j).

As an example, if the type is $(2, 1)$ and $I = [3, 5)$, then $I(2, 1)$ would be the constraint $3 \leq X_2 - X_1 < 5$. For v where $v(X_2) = 9$ and $v(X_1) = 5.2$ we would find that $I(2, 1)(v)$ is true, while for v' with $v'(X_2) = 3$ and $v'(X_1) = 4$ we would have $I(2, 1)(v')$ is false.

This allows us to give the definition of a CDD:

Definition 2 (Clock Difference Diagram) *A* Clock Difference Diagram *(CDD) is a directed acyclic graph consisting of a set of nodes V and two functions* $\mathsf{type} : V \longrightarrow \mathcal{T}$ *and* $\mathsf{succ} : V \longrightarrow 2^{\mathcal{I} \times V}$ *such that*

- *V has exactly two* terminal nodes *called* True *and* False, *where* $\mathsf{type}(\mathsf{True}) = \mathsf{type}(\mathsf{False}) = (0, 0)$ *and* $\mathsf{succ}(\mathsf{True}) = \mathsf{succ}(\mathsf{False}) = \emptyset$.
- *all other nodes $n \in V$ are* inner nodes, *which have attributed a* type $\mathsf{type}(n) \in \mathcal{T}$ *and a finite set of successors* $\mathsf{succ}(n) = \{(I_1, n_1), \ldots, (I_k, n_K)\}$, *where* $(I_i, n_i) \in \mathcal{I} \times V$.

We shall write $n \xrightarrow{I} m$ to indicate that $(I, m) \in \mathsf{succ}(n)$. For each inner node n, the following must hold:

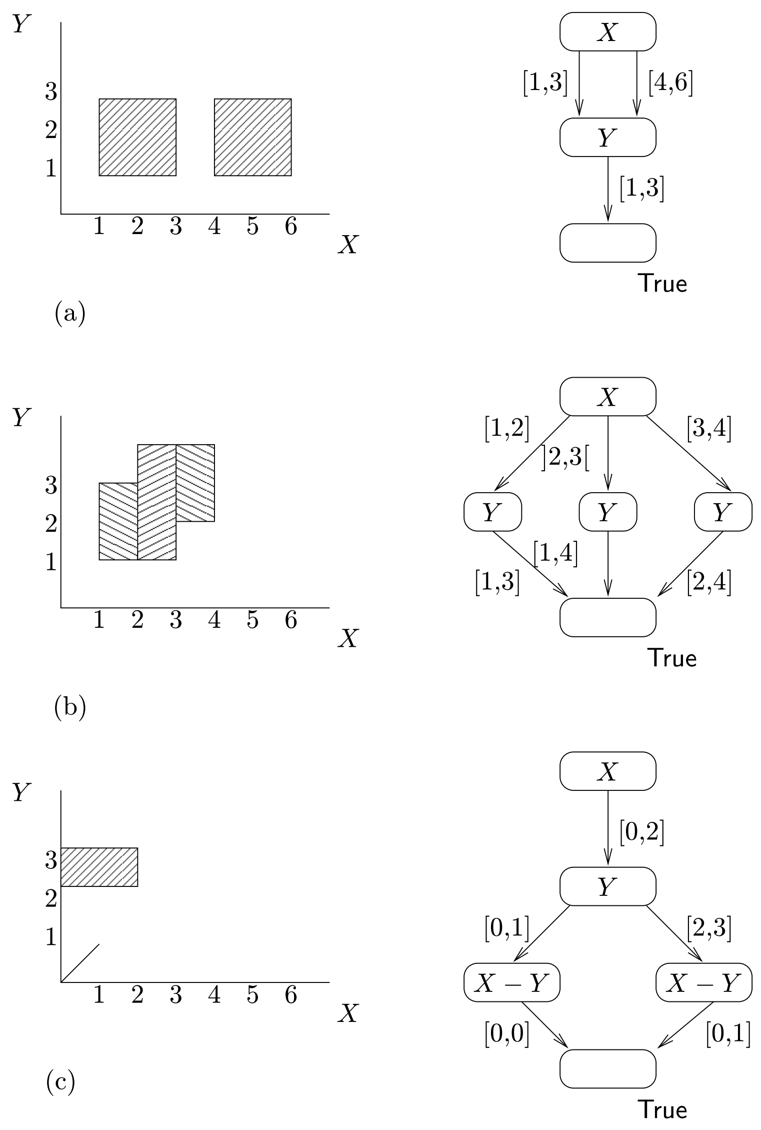

Figure 3: Three example CDD's. Intervals not shown lead implitely to False; e.g. in (a) there are arcs from the X-node to False for the three intervals $]-\infty,1[$, $]3,4[$, and $]6,\infty[$.

- *the successors are* disjoint*: for* $(I, m), (I', m') \in \mathsf{succ}(n)$ *either* $(I, m) = (I', m')$ *or* $I \cap I' = \emptyset$,
- *the successor set is an* $\mathbb{R}$-cover*:* $\bigcup\{I \mid \exists m.n \xrightarrow{I} m\} = \mathbb{R}$,
- *the CDD is* ordered*: for all* m*, whenever* $n \xrightarrow{I} m$ *then* $\mathsf{type}(m) \sqsubseteq \mathsf{type}(n)$

Further, the CDD is assumed to be reduced*, i.e.*

- *it has* maximal sharing*: for all* $n, m \in V$*, whenever* $\mathsf{succ}(n) = \mathsf{succ}(m)$ *then* $n = m$,
- *it has* no trivial edges*: whenever* $n \xrightarrow{I} m$ *then* $I \neq \mathbb{R}$,
- *all intervals are* maximal*: whenever* $n \xrightarrow{I_1} m, n \xrightarrow{I_2} m$ *then* $I_1 = I_2$ *or* $I_1 \cup I_2 \notin \mathcal{I}$

Note that we do not require a special root node. Instead each node can be chosen as the root node, and the sub-DAG underneath this node is interpreted as describing a (possibly non-convex) set of

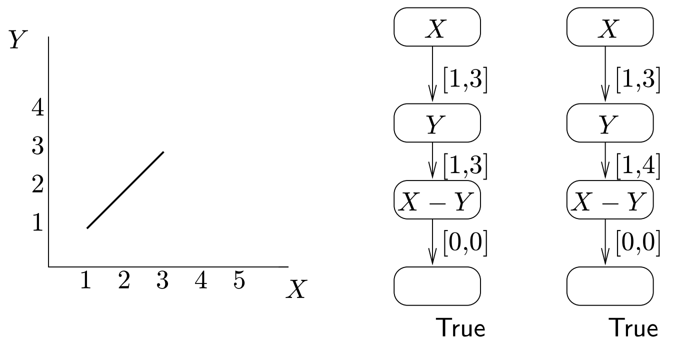

Figure 4: Two reduced CDD's for the same zone

clock valuations. This allows for sharing not only within a representation of one set of valuations, but between all representations. Figure 3 gives some examples of CDD's. The following definition makes precise how to interprete such a DAG:

Definition 3 *Given a CDD* $(V, \mathsf{type}, \mathsf{succ})$, *each node* $n \in V$ *is assigned a semantics* $[\![n]\!] \subseteq \mathcal{V}$, *recursively defined by*

- $[\![\mathsf{False}]\!] := \emptyset,\ [\![\mathsf{True}]\!] := \mathcal{V},$
- $[\![n]\!] := \{v \in \mathcal{V} \mid n \xrightarrow{I} m, I(\mathsf{type}(n))(v) = \mathsf{true}, v \in [\![m]\!]\}$ *where* n *is an inner node*

For BDD's and IDD's, testing for equality can be achieved easily due to their canonicity: the test is reduced to a pure syntactical comparison. However, in the case of CDD's canonicity is not achieved in the same straightforward manner.

To see this, we give an example of two reduced CDD's in Figure 4 describing the same set. The two CDD's are however not isomorphic. The problem with CDD's – in contrast to IDD's – is that the different types of constraints in the nodes are not independent, but influence each other. In the above example obviously $1 \leq X \leq 3$ and $X = Y$ already imply $1 \leq Y \leq 3$. The constraint on Y in the CDD on the right hand side is simply too loose. Therefore a step towards an improved normal form is to require that on all paths, the constraints should be the tightest possible. We turn back to this issue in the final section.

4 Operation on CDD's

4.1 Simple Operations

Three important operations on CDD's, namely union, intersection and complement, can be defined analogously to IDD's. All use a function makenode which for a given type t and a successor set $S = \{(I_1, n_1), \ldots, (I_k, n_K)\}$ will either return the unique node in the given CDD $C = (V, \mathsf{type}, \mathsf{succ})$ having these attributes or, in case no such exists, add a new node to the CDD with the given attributes. This operation – shown in Figure 5 – is important in order to keep reducedness of the CDD. Note that using a hashtable to identify nodes already in V, makenode can be implemented to run in constant time. Then union can be defined as in Figure 6. Intersection is computed by replacing "union" by "intersect" everywhere in Figure 6, and additonally adjusting the base cases. The complement is computed as given in Figure 7.[3]

[3]As for the BDD apply-operator, using a hashed operation-cache is needed to avoid recomputation of the same operation for the same arguments.

```
$\mathsf{makenode}(t, S)$:
    if $(\exists n \in V.\mathsf{type}(n) = t \wedge \mathsf{succ}(n) = S)$ return $n$
    else
        $V := V \cup \{n\}$ // where $n$ is a fresh node
        $\mathsf{type} := \mathsf{type} \cup \{n \mapsto t\}$
        $\mathsf{succ} := \mathsf{succ} \cup \{n \mapsto S\}$
        return $n$
    endif
```

Figure 5: Finding a node for a CDD

```
$\mathsf{union}(n_1, n_2)$
    if $n_1 = \mathsf{True}$ or $n_2 = \mathsf{True}$ then return $\mathsf{True}$
    elseif $n_1 = \mathsf{False}$ then return $n_2$
    elseif $n_2 = \mathsf{False}$ then return $n_1$
    else
        if $\mathsf{type}(n_1) = \mathsf{type}(n_2)$ then
            return $\mathsf{makenode}\big(\mathsf{type}(n_1), \{(I_1 \cap I_2, \mathsf{union}(n_1', n_2')) \mid n_1 \xrightarrow{I_1} n_1', n_2 \xrightarrow{I_2} n_2', I_1 \cap I_2 \neq \emptyset\}\big)$
        elseif $\mathsf{type}(n_1) \sqsubseteq \mathsf{type}(n_2)$ then
            return $\mathsf{makenode}\big(\mathsf{type}(n_1), \{(I_1, \mathsf{union}(n_1', n_2)) \mid n_1 \xrightarrow{I_1} n_1'\}\big)$
        elseif $\mathsf{type}(n_2) \sqsubseteq \mathsf{type}(n_1)$ then
            return $\mathsf{makenode}\big(\mathsf{type}(n_2), \{(I_2, \mathsf{union}(n_1, n_2')) \mid n_2 \xrightarrow{I_2} n_2'\}\big)$
        endif
    endif
```

Figure 6: Union of two CDD's

4.2 From constraint systems to CDD's

The reachability algorithm of UPPAAL currently works with constraint systems (represented either as canonical DBM's or in the minimal form). The desired reachability algorithm will need to combine and compare DBM's obtained from exploration of the timed automaton with CDD's used as a compact representation of the PASSED-list.

For the following we assume that a constraint system D holds at most one simple constraint for each pair of clocks X_i, X_j (which is obviously true for DBM's and the minimal form). Let $D(i, j)$ be the set of all simple constraints of type (i, j), i.e. those for $X_i - X_j$ and $X_j - X_i$. The constraint system $D(i, j)$ gives an upper and/or a lower bound for $X_i - X_j$. If not present, choose $-\infty$ as lower and $+\infty$ as upper bound. Denote the interval defined thus by $I_{D(i,j)}$.

Further, given an interval $I \in \mathcal{I}$, let $lo(I) := \{r \in \mathbb{R} \mid \forall r' \in I.r < r'\}$ be the set of lower bounds and $hi(I) := \{r \in \mathbb{R} \mid \forall r' \in I.r > r'\}$ the set of upper bounds. Note that always $lo(I), hi(I) \in \mathcal{I}_\emptyset$. Using this notation, a simple algorithm for constructing a CDD from a constraint system can be

```
$\mathsf{complement}(n)$
    if $n = \mathsf{True}$ return $\mathsf{False}$
    elseif $n = \mathsf{False}$ return $\mathsf{True}$
    elseif return $\mathsf{makenode}\big(\mathsf{type}(n), \{(I, \mathsf{complement}(m)) \mid n \xrightarrow{I} m\}\big)$
    endif
```

Figure 7: Computing the complement

given as in Figure 8. Using this, we can easily union zones to a CDD as required in the modified reachability algorithm of UPPAAL (cf. footnote on page 3). Note that for this asymmetric union it is advisible to use the minimal form representation for the zone, as this will lead to a smaller CDD, and subsequently to a faster and less space-consuming union-operation.

```
makeCDD($D$)
    $n := \mathsf{True}$
    for $t \in \mathcal{T} \setminus \{(0,0)\}$ do // use ordering $\sqsubseteq$
        $I := I_{D(t)}$
        if $I \neq \mathbb{R}$ then
            if $lo(I) = \emptyset$ then
                $n := \mathsf{makenode}(t, \{(I,n), (hi(I), \mathsf{False})\})$
            elseif $hi(I) = \emptyset$ then
                $n := \mathsf{makenode}(t, \{(I,n), (lo(I), \mathsf{False})\})$
            else
                $n := \mathsf{makenode}(t, \{(I,n), (hi(I), \mathsf{False}), (lo(I), \mathsf{False})\})$
            endif
        endif
    endfor
    return $n$
```

Figure 8: Generating a CDD from a constraint system

4.3 Crucial Operations

Testing for equality and set-inclusion of CDD's is not easy without utilizing a normal form. Looking at the test given in (1) it is however evident that all we need is to test for inclusion between a zone and a CDD. Such an asymmetric test for a zone Z and a CDD n can be implemented as shown in Figure 9 without need for canonicity.

```
subset($D, n$)
    if $D = \mathsf{false}$ or $n = \mathsf{True}$ then return $\mathsf{true}$
    elseif $n = \mathsf{False}$ then return $\mathsf{false}$
    else return $\bigwedge_{n \xrightarrow{I} m}$ subset($D \wedge I(\mathsf{type}(n)), m$)
    endif
```

Figure 9: Deciding set inclusion for a zone and a CDD

Note that when testing for emptiness of a DBM as in the first if-statement, we need to compute its canonical form. If we know that the DBM is already in canonical form, the algorithm can be improved by passing $D \wedge I(\mathsf{type}(n))$ in canonical form. As $D \wedge I(\mathsf{type}(n))$ adds no more than two constraints to the zone, computation of the canonical form can be done faster than in the general case, which would be necessary in the test $D = \mathsf{true}$.

The above algorithm can also be used to test for emptiness of a CDD using

$$\mathsf{empty}(n) := \mathsf{subset}(\mathsf{true}, \mathsf{complement}(n))$$

where true is the empty set of constraints, fullfilled by every valuation.

As testing for set inclusion $C_1 \subseteq C_2$ of two CDD's C_1, C_2 is equivalent to testing for emptiness of $C_1 \cap \overline{C_2}$, also this check can be done without needing canonicity.

System	Proc.	Clocks	Prop.	Current Time	Current Const.	CDD Time	CDD Nodes	CDD Edges
Philips	4	2	6	0.68	1,072	0.71	130	298
Philips Col	7	5	9	90.38	173,065	124.26	18,495	48,829
B&O	9	3	1	149.04	160,156	170.13	1,222	4,486
BRP	6	4	10	86.18	238,406	124.81	3,886	15,641
PowerDown1	10	2	1	129.70	81,220	132.06	5,215	21,920
PowerDown2	8	1	3	52.11	35,696	53.59	57	114
Dacapo	6	5	4	278.82	243,337	334.35	34,792	98,622
Gearbox	5	5	46	47.58	157,886	64.56	7,277	17,291

Table 1: Performance Statistics

5 Implementation and Experimental Results

We have implemented a CDD-package and used it to obtain a modified, CDD-based reachability algorithm for Uppaal. As indicated in previous sections, the CDD-based reachability algorithm uses DBM's for exploration of symbolic states and CDD's in the representation of the Passed-list.

In this section we present the results of an experiment where both the current version of Uppaal[4] and the CDD-based version of Uppaal were applied to the verification of 8 industrial examples found in the literature. The examples include a gearbox controller [LPY98], various communication protocols used in Philips audio equipment [BPV94, DKRT97, BGK+96], and in B&O audio/video equipment [HSLL97, HLS98], and the start-up algorithm of the DACAPO protocol [LPY97].

System	Current Passed	Current Wait	CDD Passed	CDD Wait
Philips	423	727	422	727
Philips Col.	21,254	52,402	9,526	52,402
B&O	38,351	154,530	17,401	154,530
BRP	34,639	72,329	4,618	72,329
PowerDown1	30,349	67,897	10,666	64,387
PowerDown2	36,255	82,469	36,255	82,469
Dacapo	37,685	172,265	20,056	172,064
GearBox	19,606	37,912	13,782	37,912

Table 2: Generated States

In Table 1 we present the space requirements (in number of constraints for the current implementation and number of CDD nodes and edges for the CDD based implementation) and runtime (in seconds) of the examples on a Sun UltraSPARC 2 equipped with 512 MB of primary memory and two 170 MHz processors. Each example was verified using the current purely DBM-based algorithm of Uppaal (Current), and using the CDD-based modification (CDD). In addition the number of processes, clocks and properties checked for each system is specified. As can be seen, our CDD-based modification of Uppaal leads to truly significant space-savings ranging from 46% to 99% (comparing the number of constraints used in the current Uppaal-version with the total number of CDD-nodes and -edges) with only moderate increase in run-time (in average an increase of 17%).

Table 2 shows for each system the final number of states in the Passed-list and the total number of states that has passed through the Wait-list. The Wait-list number provides a good measure for how frequent exploration of symbolic states in the CDD-based version has been avoided due to use of the supposed less strict termination condition (1). Maybe unexpectedly, we note that the CDD-based version only rarely leads to faster termination. In fact, in all but two cases, the number of explored symbolic states is the same as for the current Uppaal-version. This offers a

[4]More precisely Uppaal version 2.19.2, which is the most recent version of Uppaal currently used in-house.

good explanation for the lack of time-improvement; the moderate increase in run-time may partly be explained by the prototypical nature of our CDD-based version.

6 Towards a fully symbolic timed reachability analysis

The presented CDD-version of Uppaal uses CDD's to store the Passed-list, but zones (i.e. DBM's) in the exploration of the timed automata. The next goal is to use CDD's in the exploration as well, thus treating the continuous part fully symbolic. In combination with a BDD-based approach for the discrete part, this would result in a fully symbolic timed reachability analysis, saving even more space and time.

The central operations when exploring a timed automaton are time progress and clock reset. Using *tightened CDD's*, these operations can be defined along the same lines as for DBM's. A tightened CDD is one where along each path to True all constraints are the the tightest possible. In [LPWW98] we have shown how to effectively transform any given CDD into an equivalent tightened one.

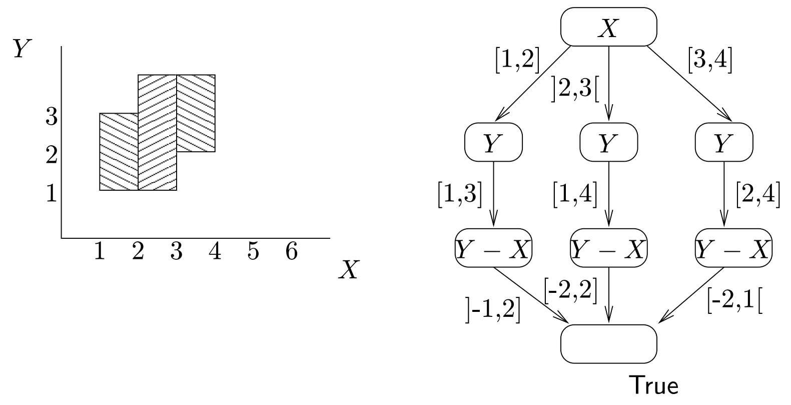

Figure 10: A tightened CDD

Figure 10 shows the tightened CDD-representation for example (b) from Figure 3. Given this tightened version, the time progress operation is obtained by simply removing all upper bounds on the individual clocks. In general, this gives a CDD with overlapping intervals, which however can easily be turned into a CDD obeying our definition. More details on these operations can be found in [LPWW98].

CDD's come equipped with an obvious notion of being *equally fine partitioned*. For equally fine partitioned CDD's we have the following normal form theorem [LPWW98]:

Theorem 1 *Let C_1, C_2 be two CDD's which are tightened and equally fine partitioned. Then $[\![C_1]\!] = [\![C_2]\!]$ iff C_1 and C_2 are graph-isomorphic.*

A drastical way of achieving equally fine partitioned CDD's is to allow only atomic integer-bounded intervals, i.e. intervals of the form $[n, n]$ or $(n, n+1)$. This approach has been taken in [ABKMPR97, BMPY97] demonstrating canonicity. However, this approach is extremely sensitive to the size of the constants in the analysed model. In contrast, for models with large constants our notion of CDD allows for coarser, and hence more space-efficient, representations.

7 Conclusion

In this paper, we have presented Clock Difference Diagrams, CDD's, a BDD-like data-structure for effective representation and manipulation of finite unions of zones. A version of the real-time verification tool UPPAAL using CDD's to store explored symbolic states has been implemented. Our experimental results on 8 industrial examples found in the literature demonstrate significant space-savings (46%–99%) with a moderate increase in run-time (in average 17%). As future work, we want to experimentally pursue the fully symbolic state-space exploration of the last section and [LPWW98].

Acknowledgement

The second author of this paper was introduced to the idea of developing a BDD-like structure with nodes labeled with bounds on clock-differences by Henrik Reif Andersen.

References

[ABKMPR97] E. Asarain, M. Bozga, A. Kerbrat, O. Maler, A. Pnueli, A. Rasse. Data-Structures for the Verification of Timed Automata. In Proc. HART'97, LNCS 1201, pp. 346–360.

[AD94] R. Alur and D. Dill. Automata for Modelling Real-Time Systems. In *Proc. of ICALP'90*, volume 443 of *Lecture Notes in Computer Science*, 1990.

[Bal96] Felice Balarin. *Approximate Reachability Analysis of Timed Automata*. Proc. Real-Time Systems Symposium, Washington, DC, December 1996, pp. 52–61.

[BGK+96] Bengtsson, Griffioen, Kristoffersen, Larsen, Larsson, Pettersson, and Yi. Verification of an audio protocol with bus collision using uppaal. In *Proceedings of CAV'96*, volume 1102, 1996.

[BLLPW96] Johan Bengtsson, Kim G. Larsen, Fredrik Larsson, Paul Pettersson, and Wang Yi. UPPAAL in 1995. In *Proc. of the 2*nd *Workshop on Tools and Algorithms for the Construction and Analysis of Systems*, number 1055 in Lecture Notes in Computer Science, pages 431–434. Springer–Verlag, March 1996.

[BMPY97] M. Bozga, O. Maler, A. Pnueli, and S. Yovine. Some progress in the symbolic verification of timed automata. *Lecture Notes in Computer Science*, 1254:179–190, 1997. Proceedings of CAV'97.

[BPV94] D. Bosscher, I. Polak, and F. Vaandrager. Verification of an Audio-control Protocol. In *Proceedings of Formal Techniques in Real-Time and Fault-Tolerant Systems*, volume 863 of *Lecture Notes in Computer Science*, 1994.

[CC95] S.V. Campos and E.M. Clarke. Real-time symbolic model checking for discrete time models. In C. Rattray T. Rus, editor, *AMAST Series in Computing: Theories and Experiences for Real-Time System Development*, 1995.

[Dill87] D.L. Dill. *Timing Assumptions and Verification of Finite-State Concurrent Systems.* in: J. Sifakis (Ed.), Automatic Verification Methods for Finite State Systems. LNCS 407, Springer Berlin 1989, pp. 197-212.

[DKRT97] D'Arginio, Katoen, Ruys, and Tretmans. Bounded retransmission protocol must be on time ! In *Proceedings of TACAS'97*, volume 1217, 1997.

[DY95] C. Daws and S. Yovine. Two examples of verification of multirate timed automata with KRONOS. In *Proc. of the 16*th *IEEE Real-Time Systems Symposium*, pages 66–75, December 1995.

[HNSY94] Thomas. A. Henzinger, Xavier Nicollin, Joseph Sifakis, and Sergio Yovine. Symbolic Model Checking for Real-Time Systems. *Information and Computation*, 111(2):193–244, 1994.

[HHW95] Thomas A. Henzinger, Pei-Hsin Ho, and Howard Wong-Toi. A Users Guide to HYTECH. Technical report, Department of Computer Science, Cornell University, 1995.

[HLS98] K. Havelund, K. G. Larsen, and A. Skou. Formal Verification of an Audio/Video Power Controller using the Real-Time Model Checker UPPAAL. Technical report made for Bang& Olufsen, 1998.

[HSLL97] K. Havelund, A. Skou, K.G. Larsen, and K. Lund. Formal Modelling and Analysis of an Audio/Video Protocol: An Industrial Case Study using UPPAAL. In *In Proceedings of the 18th IEEE Real-Time System Symposium*, 1997.

[LPW95] Kim G. Larsen, Paul Pettersson and Wang Yi. *Compositional and Symbolic Model-Checking of Real-Time Systems.* In Proceedings of the 16th IEEE Real-Time Systems Symposium, Pisa, Italy, 5-7 December, 1995.

[LPWW98] K.G. Larsen, C. Weise, W. Yi, J. Pearson. *Clock Difference Diagrams.* DoCS Technical Report No.98/99, Uppsala University, Sweden, presented at the Nordic Workshop on Programming Theory, Turku, Finland, November 1998.

[LPY97] H. Lönn, P. Pettersson, and W. Yi. Formal Verification of a TDMA Protocol Start-Up Mechanism. In *Proceedings of 1997 IEEE Pacific Rim International Symposium on Fault-Tolerant Systems*, pages 235–242, 1997.

[LPY98] M. Lindahl, P. Pettersson, and W. Yi. Formal design and analysis of a gear controller. In Bernhard Steffen, editor, *Proceedings of 4th International Workshop on Tools and Algorithms for the Construction and Analysis of Systems*, volume 1384 of *Lecture Notes in Computer Science*, pages 281–297. Springer Verlag, 1998.

[YPD94] Wang Yi, Paul Pettersson, and Mats Daniels. Automatic Verification of Real-Time Communicating Systems By Constraint-Solving. In *Proc. of the 7*th *International Conference on Formal Description Techniques*, 1994.

[ST98] Karsten Strehl, Lothar Thiele. Symbolic Model Checking of Process Networks Using Interval Diagram Techniques. Proceedings of the IEEE/ACM International Conference on Computer-Aided Design (ICCAD-98), San Jose, California, pp. 686–692, 1998.

[Strehl'98] Karsten Strehl. Using Interval Diagram Techniques for the Symbolic Verification of Timed Automata. Technical Report TIK-53, Institut für Technische Informatik und Kommunikationsnetze (TIK), ETH Zürich, July 1998.

[WTD95] Howard Wong-Toi and David L. Dill. *Verification of real-time systems by successive over and under approximation.* International Conference on Computer-Aided Verification, July 1995.

Recent BRICS Report Series Publications