# Decidability and Symbolic Verification

#### Kim G. Larsen Aalborg University, DENMARK





# **Overview**

- Decidability
  - Region Construction
  - Reachability & Bisimulation Checking
- Symbolic Verification
  - On-the-fly Exploration
  - Zones and Difference Bounded Matrices (DBM)
  - Clock Difference Diagrams (CDD)
- Verification Options



# **Reachability**?



# **The Region Abstraction**



- "compatibility" between regions and constraints
- "compatibility" between regions and time elapsing
  - $\rightsquigarrow$  an equivalence of finite index
    - a time-abstract bisimulation



# **Time Abstracted Bisimulation**

This is a relation between • and • such that:



... and vice-versa (swap • and •).





#### **Regions** – From Infinite to Finite



L

Kim Larsen [6]

# **Region Graph**

It "mimicks" the behaviours of the clocks.



#### Region Automaton = Finite Bisimulation Quotiont









# **Region Automaton**



LARGE: exponential in the number of clocks and in the constants (if encoded in binary). The number of regions is

$$\prod_{x \in X} (2M_x + 2) \cdot |X!| \cdot 2^{|X|}$$





# **Fundamental Results**

Reachability



- Model-checking
  - UNDECIDABLE ■ TCTL ③ <sup>PSPACE-C</sup> ; MTL ⊗



- Bisimulation, Simulation
  - Timed <sup>(C)</sup> <sup>EXPTIME-</sup>, ; Untimed <sup>(C)</sup>
- Trace-inclusion
  - : Untimed 😳 PSPACE-c Timed 🙁 UNDECIDABLE

# UPPAAL

# Modeling & Specification





# **Train Crossing**



# **Train Crossing**



#### **Declarations**

🚖 C:/Documents and Settings/I	(im/Desktop/uppaal-3.4.7/demo/train-gate.xml - UPPAAL	
File Templates View Queries (	Options Help	
🔩 📹 💾 🔍 🤅	$\langle \langle \langle \langle \rangle \rangle \rangle = \langle \langle \rangle \rangle \langle \langle \rangle \rangle \langle \rangle \rangle$	
System Editor Simulator Verifie	r	
Drag out       /*         Itain-gate       * For more details about this example, see         Global declarations       * "Automatic Verification of Real-Time Communicating Systems by Constraint Solvin         Train       * by Wang Yi, Paul Pettersson and Mats Daniels. In Proceedings of the 7th Internations         Train       * Conference on Formal Description Techniques, pages 223-238, North-Holland. 1994         Process assignments       * /		by Constraint Solving", gs of the 7th International , North-Holland. 1994.
System definition	<pre>const N 5; // # trains + 1 int[0,N] el; chan appr, stop, go, leave; chan empty, notempty, hd, add, rem;</pre>	Constants Bounded integers
Irain-gate Global declarations □S3 Train ↓ Declarations	clock x;	Channels Clocks Arrays
<ul> <li>Irain-gate</li> <li>Global declarations</li> <li>□- S Train</li> <li>Declarations</li> <li>①- S Gate</li> <li>□- S IntQueue</li> <li>Declarations</li> </ul>	<pre>int[0,N] list[N], len, i;</pre>	Types Functions
<ul> <li>Process assignments</li> <li>System definition</li> </ul>	Trainl:=Train(el, 1); Train2:=Train(el, 2); Train3:=Train(el, 3); Train4:=Train(el, 4);	Templates Processes Systems
IntQueue     Declarations     Declarations     System definition	system Trainl, Train2, Train3, Train4, Beijing, 2011, Queue; Kim Larsen [15]	

# **UPPAAL Help**



# **Logical Specifications**

- Validation Properties
  - Possibly: E<> P
- Safety Properties
  - Invariant: A[] P
  - Pos. Inv.: E[] *P*
- Liveness Properties
  - Eventually: A<> P
  - Leadsto:  $P \rightarrow Q$
- Bounded Liveness
  - Leads to within:  $P \rightarrow_{\leq t} Q$

The expressions *P* and *Q* must be type safe, side effect free, and evaluate to a boolean.

Only references to integer variables, constants, clocks, are allowed (and arrays of these).



# **Symbolic Verification**

# The UPPAAL Verification Engine





#### **Regions** – From Infinite to Finite



The number of regions is  $n! \cdot 2^n \cdot \prod_{x \in C} (2c_x + 2)$ .

ARTIST Design PhD School, Beijing, 2011

Kim Larsen [19]



#### **Zones** – From Finite to Efficiency





# **Zones** – Operations



# Symbolic Transitions





**ARTIST** Design PhD School, Beijing, 2011

Init -> Final ?



155

**Init** -> Final ?



INITIAL Passed :=  $\emptyset$ ; Waiting := {(n<sub>0</sub>,Z<sub>0</sub>)}

REPEAT pick (n,Z) in Waiting if (n,Z) = Final return true for all  $(n,Z) \rightarrow (n',Z')$ : if for some (n',Z'')  $Z' \subseteq Z''$  continue else add (n',Z') to Waiting move (n,Z) to Passed

UNTIL Waiting =  $\emptyset$  return false



Init -> Final ?



AGTIST Design PhD School, Beijing, 2011

**Init** -> Final ?



INITIAL Passed :=  $\emptyset$ ; Waiting :=  $\{(n_0, Z_0)\}$ 

pick (n,Z) in Waiting if (n,Z) = Final return true for all  $(n,Z) \rightarrow (n',Z')$ : if for some (n',Z'')  $Z' \subseteq Z''$  continue else add (n',Z') to Waiting move (n,Z) to Passed

UNTIL Waiting =  $\emptyset$ return false



**Init** -> Final ?



INITIAL Passed :=  $\emptyset$ ; Waiting :=  $\{(n_0, Z_0)\}$ 

pick (n,Z) in Waiting if (n,Z) = Final return true for all  $(n,Z) \rightarrow (n',Z')$ : if for some (n',Z'')  $Z' \subseteq Z''$  continue else add (n',Z') to Waiting move (n,Z) to Passed

UNTIL Waiting =  $\emptyset$ return false



Init -> Final ?















ARTIST Design PhD School, Beijing, 2011



ARTIST Design PhD School, Beijing, 2011



ARTIST Design PhD School, Beijing, 2011










### **Symbolic Exploration**



### **Symbolic Exploration**



# **Datastructures for Zones**

- Difference Bounded Matrices (DBMs)
- Minimal Constraint Form [RTSS97]



 Clock Difference Diagrams [CAV99]

# Inclusion Checking (DBMs)

#### Bellman 1958, Dill 1989



### Future (DBMs)





### Reset (DBMs)



# **Clock Difference Diagrams**





### **Earlier Termination**

Init -> Final ?





### **Earlier Termination**

**Init** -> Final ?



ARTIST Design PhD School, Beijing, 2011

Kim Larsen [45]

### **Clock Difference Diagrams**

= Binary Decision Diagrams + Difference Bounded Matrices

**CAV99** 



- Nodes labeled with differences
- Maximal sharing of substructures (also across different CDDs)
- Maximal intervals
- Linear-time algorithms for set-theoretic operations.
- NDD's Maler et. al
- DDD's Møller, Lichtenberg

### **Clock Difference Diagrams I**

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

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

# **Disjoint & Ordered**

For each inner node n, the following must hold:

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

# Reduced

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

- it has *maximal sharing*: for all  $n, m \in V$ , whenever succ(n) = 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  $I_1 \cup I_2 \notin \mathcal{I}$

### **Clock Difference Diagrams**





## Makenode

Let t be a type and  $S = \{(I_1, n_1), \dots, (I_k, n_K)\}$ a successor set. We want to extend a given CDD C = (V, type, succ) with a node n with these attributes.



#### SPACE PERFORMANCE



# Union

```
union(n_1, n_2)
     if n_1 = \text{True or } n_2 = \text{True then return True}
     elseif n_1 = False then return n_2
     elseif n_2 = False then return n_1
     else
          if type(n_1) = type(n_2) then
                return MN(type(n_1), {(I_1 \cap I_2,
                          union(n'_1, n'_2) | n_1 \xrightarrow{I_1} n'_1, n_2 \xrightarrow{I_2} n'_2, I_1 \cap I_2 \neq \emptyset }
          elseif type(n_1) \sqsubseteq type(n_2) then
                return MN(type(n_1), {(I_1, union(n'_1, n_2)) | n_1 \xrightarrow{I_1} n'_1})
          elseif type(n_2) \sqsubseteq type(n_1) then
                return MN(type(n_2), {(I_2, union(n_1, n'_2)) | n_2 \xrightarrow{I_2} n'_2})
          endif
1
     endif
```

# Complement

```
\begin{aligned} \mathsf{complement}(n) \\ & \text{if } n = \mathsf{True \ return \ False} \\ & \text{elseif } n = \mathsf{False \ return \ True} \\ & \text{elseif \ return \ }\mathsf{MN}\Big(\mathsf{type}(n), \{(I, \mathsf{complement}(m)) \mid n \xrightarrow{I} m\}\Big) \\ & \text{endif} \end{aligned}
```



5

# Zones, CDD, Subset

Single zones may be represented as single path CDD's in the obvious manner.

It may be advantageous to represent zones using a *minimal* set of constraints (see [RTSS97]).

subset(D, n) // D const. sys., n CDD-node if D = false or n = True then return true elseif n = False then return false else return  $\bigwedge_{n \to m}$  subset $(D \land I_n), m)$ endif where  $I_n$  is the constraint  $X_i - X_j \in I$  if type(n) = (i, j).

#### TIME PERFORMANCE



# Related & Recent Work

- DDD: Andersen et al.
- NDD: Asarin, Bozga, Kerbrat, Maler, Pnueli, Rasse.
- IDD: Strehl, Thiele.

- Recent work on fully symbolic engine for TA:
  - Georges Morbe, Florian Pigorsch and Christoph Scholl: Fully Symbolic Model Checking for Timed Automata. CAV 2011.



# **Verification Options**





# **Verification Options**

♣ C:/Documents and Settings/kgl/Desktop/KIM/UPPAAL/UPPA.							
File Edit View Tools	Options He	elp					
Editor Simulator Verifier	Search Order State Space Reduction State Space Representation Diagnostic Trace Extrapolation		* * * *	<b>&gt;</b> `			
A[] (RobotA.a <= E[] (( bodenA ==	Hash tab V Reuse	le size	•	bodenC =			
E<> ( (bodenA > E<> not deadloc)	5)    (bo :	odenB > 5)	(bo	denC > 5			

**Search Order** Depth First **Breadth First State Space Reduction** None Conservative Aggressive **State Space Representation** DBM **Compact Form Under Approximation Over Approximation Diagnostic Trace** Some Shortest Fastest

#### Extrapolation Hash Table size Reuse

Kim Larsen [59]



### **State Space Reduction**



Cycles:

Only symbolic states involving loop-entry points need to be saved on Passed list



S



## **Over/Under** Approximation



**Declared State Space** 

 $\begin{array}{l} \mathsf{G}{\in}\;\mathsf{U}\;\Rightarrow\mathsf{G}{\in}\;\mathsf{R}\\ \neg(\mathsf{G}{\in}\;\mathsf{O})\Rightarrow\neg(\mathsf{G}{\in}\;\mathsf{R}) \end{array}$ 

5

### Over-approximation Convex Hull





TACAS04: An EXACT method performing as well as Convex Hull has been developed based on abstractions taking max constants into account distinguishing between clocks, locations and  $\leq \& \geq$ 

### Under-approximation *Bitstate Hashing*







### Under-approximation *Bitstate Hashing*



### **Extrapolation**





5

### **Forward Symbolic Exploration**



### Abstractions

$$a: \mathcal{P}(R^X_{\geq 0}) \hookrightarrow \mathcal{P}(R^X_{\geq 0})$$
 such that  $W \subseteq a(W)$ 

$$\frac{(\ell, W) \Rightarrow (\ell', W')}{(\ell, W) \Rightarrow_{a} (\ell', a(W'))} \quad \text{if } W = a(W)$$

We want  $\Rightarrow_a$  to be:

- sound & complete wrt reachability
- finite
- easy to compute
- as coarse as possible



### **Abstraction by Extrapolation**

[Daws, Tripakis 98]

Let *k* be the largest constant appearing in the TA



## **Location Dependency**

[Behrmann, Bouyer, Fleury, Larsen 03]



$$k_x = 5 \ k_y = 10^6$$

Will generate all symbolic states of the form

 $(I_2, x \in [0, 14], y \in [5, 14n], y - x \in [5, 14n - 14])$ 

for  $n \leq \! 10^6 \! / 14 ~ !!$ 

But  $y \ge 10^6$  is not RELEVANT in  $I_2$ 

ARTIST Design PhD School, Beijing, 2011

Kim Larsen [70]

### Location Dependent Constants



$$k_x = 5 \ k_y = 10^6$$

$$\begin{array}{rl} k_x^{i} &= 14 & \text{for } i \in \{1, 2, 3, 4\} \\ k_y^{i} &= 5 & \text{for } i \in \{1, 2, 3\} \\ k_y^{4} &= 10^6 \end{array}$$

 $k_j^i$  may be found as solution to simple linear constraints!

Active Clock Reduction:  $k_i^i = -\infty$ 

ARTIST Design PhD School, Beijing, 2011

Kim Larsen [71]

## **Experiments**

	Constant	Global	Active-clock	Local
	BIG	Method	Reduction	Constants
	$10^{3}$	0.05s/1MB	0.05s/1MB	0.00s/1MB
Naive Example	$10^{4}$	4.78s/3MB	4.83s/3MB	0.00s/1MB
	10 <sup>5</sup>	484s/13MB	480s/13MB	0.00s/1MB
	$10^{6}$	stopped	stopped	0.00s/1MB
	$10^{3}$	3.24s/3MB	3.26s/3MB	0.01s/1MB
Two Processes	$10^{4}$	5981s/9MB	5978s/9MB	0.37s/2MB
	$10^{5}$	stopped	stopped	72s/5MB
	$10^{3}$	0.01s/1MB	0.01s/1MB	0.01s/1MB
Asymmetric	$10^{4}$	2.20s/3MB	2.20s/3MB	0.85s/2MB
Fischer	$10^{5}$	333s/19MB	333s/19MB	160s/13MB
	$10^{6}$	33307s/122MB	33238s/122MB	16330s/65MB
Bang & Olufsen	25000	stopped	159s/243MB	123s/204MB

ARTIST Design PhD School, Beijing, 2011

Kim Larsen [72]


#### Lower and Upper Bounds

ARTIST Design PhD School, Beijing, 2011



Given that  $x \le 10^6$  is an *upper* bound implies that

 $(I,v_x,v_y)$  simulates  $(I,v'_x,v_y)$ 

Kim

whenever 
$$v'_x \ge v_x \ge 10$$
.

For reachability downward closure wrt simulation suffices!

#### Simulation

#### $\preccurlyeq$ is the largest relation satisfying

- 1. if  $(\ell_1, \nu_1) \preccurlyeq (\ell_2, \nu_2)$  then  $\ell_1 = \ell_2$ 2. if  $(\ell_1, \nu_1) \preccurlyeq (\ell_2, \nu_2)$  and  $(\ell_1, \nu_1) \xrightarrow{\longrightarrow} (\ell'_1, \nu'_1)$ , then there exists  $(\ell'_2, \nu'_2)$  such that  $(\ell_2, \nu_2) \xrightarrow{\longrightarrow} (\ell'_2, \nu'_2)$  and  $(\ell'_1, \nu'_1) \preccurlyeq (\ell'_2, \nu'_2)$
- 3. if  $(\ell_1, \nu_1) \preccurlyeq (\ell_2, \nu_2)$  and  $(\ell_1, \nu_1) \xrightarrow{\epsilon(\delta)} (\ell_1, \nu_1 + \delta)$ , then there exists  $\delta'$  such that  $(\ell_2, \nu_2) \xrightarrow{\epsilon(\delta')} (\ell_2, \nu_2 + \delta')$  and  $(\ell_1, \nu_1 + \delta) \preccurlyeq (\ell_2, \nu_2 + \delta')$

#### Proposition

If  $(\ell, \nu_1) \preccurlyeq (\ell, \nu_2)$  and if a discrete state  $\ell'$  is reachable from  $(\ell, \nu_1)$ , then it is also reachable from  $(\ell, \nu_2)$ .

#### **Maximal Bounds**

M(x): the maximum constant k with  $x \sim k$ , L(x): the maximum constant k with  $x\{\geq,>\}k$ , U(x): the maximum constant k with  $x\{\leq,<\}k$ .

$$\nu \equiv_M \nu' \stackrel{\text{def}}{\iff} \forall x \in X : \text{either } \nu(x) = \nu'(x) \text{ or } (\nu(x) > M(x) \text{ and } \nu'(x) > M(x))$$

$$\nu' \prec_{LU} \nu \iff \text{for each clock } x, \begin{cases} \text{either } \nu'(x) = \nu(x) \\ \text{or } L(x) < \nu'(x) < \nu(x) \\ \text{or } U(x) < \nu(x) < \nu'(x) \end{cases}$$

ARTIST Design PhD School, Beijing, 2011

5

#### **Maximum Bounds Abstraction**



#### ARTIST Design PhD School, Beijing, 2011

5

#### **Extrapolation Using Zones**



ARTIST Design PhD School, Beijing, 2011

#### **Experiments**

		Classical			Loc. dep. Max			Loc. dep. LU			Convex Hull		
		-n1			-n2			-n3			-A		
Fischer	Model	Time	States	Mem	Time	States	Mem	Time	States	Mem	Time	States	Mem
	f5	4.02	82,685	5	0.24	16,980	3	0.03	2,870	3	0.03	3,650	3
	f6	597.04	1,489,230	49	6.67	158,220	7	0.11	11,484	3	0.10	14,658	3
	f7				352.67	1,620,542	46	0.47	44,142	3	0.45	56,252	5
	f8							2.11	$164,\!528$	6	2.08	208,744	12
	f9							8.76	598,662	19	9.11	754,974	39
	f10							37.26	2,136,980	68	39.13	2,676,150	143
	f11							152.44	7,510,382	268			
CSMA/CD	c5	0.55	27,174	3	0.14	10,569	3	0.02	2,027	3	0.03	1,651	3
	c6	19.39	287,109	11	3.63	87,977	5	0.10	6,296	3	0.06	4,986	3
	c7				195.35	813,924	29	0.28	18,205	3	0.22	14,101	4
	c8							0.98	50,058	5	0.66	38,060	7
	c9							2.90	132,623	12	1.89	99,215	17
	c10							8.42	341,452	29	5.48	251,758	49
	c11							24.13	859,265	76	15.66	625,225	138
	c12							68.20	2,122,286	202	43.10	1,525,536	394
	bus	102.28	6,727,443	303	66.54	4,620,666	254	62.01	4,317,920	246	45.08	3,826,742	324
	philips	0.16	12,823	3	0.09	6,763	3	0.09	6,599	3	0.07	5,992	3
	sched	17.01	929,726	76	15.09	700,917	58	12.85	619,351	52	55.41	3,636,576	427

ARTIST Design PhD School, Beijing, 2011

|0055

#### Additional "secrets"

- Sharing among symbolic states
  - Iocation vector / discrete values / zones
- Distributed implementation of UPPAAL
- Symmetry Reduction
- Sweep Line Method
- Guiding wrt Heuristic Value
  - User-supplied / Auto-generated
- Slicing wrt "C" Code

# Leader Election Protocol



Protocol analysed in UPPAAL by Leslie Lamport CHARME'05





























# Flooding



5

# Flooding





ARTIST Design PhD School, Beijing, 2011

# Flooding







#### Forwarding





#### Forwarding





#### Forwarding



























# Claim to be verified Correct leader is known at a node *i* after

 $t(i) = \Delta_{TO} + \Delta_{TDELAY} + d_i \Delta_{MDELAY}$ 

#### A model checking problem

 $IMP \vDash \Box_{>t(i)} I(i) = L(i)$  for all i.

ARTIST Design PhD School, Beijing, 2011

#### Modelling (RT) protocols



#### Modelling the election protocol



#### <u>Static</u> Topology : Node $\times$ Node $\rightarrow$ B

Message src: Node dst: Node leader: Node hopss: N

ARTIST Design PhD School, Beijing, 2011

Kim Larsen [96]

**L D 5** 5

#### **Global Declaration**

```
void setMsq(msq t &msq, id t src, id t dst, id t leader, int[0,N] hops)
 msq.src = src;
                                 const int N = 3;
 msq.dst = dst;
                                 const int MDELAY = 3;
 msq.leader = leader;
                                 const int TDELAY = 5;
 msq.hops = hops;
                                 const int TO = 10;
                                 typedef int[0,N-1] id t;
chan send;
                                 typedef struct
chan receive[N];
msq t shared;
                                   id t src;
                                   id t dst;
const int link[N][N] = {
  \{0,1,1\},\
                                   id t leader;
  \{1,0,1\},\
                                   int[0,N] hops;
  \{1,1,0\}
                                  msq t;
```









# Local Declarations (Node[id])

```
id t leader = id;
                                            int[0,1000] timeout()
int[0,N] hops;
clock x;
                                              if (hops > 0)
int[0,N] i;
                                                return TO + TDELAY + hops * MDELAY;
id t src;
                                              return TO;
void set(id t l, int[0,N] h)
                                            bool worse(const msq t &msq)
  leader = 1;
  hops = h;
                                              return msq.leader > leader || msq.leader
                                                   == leader \&\& msq.hops > hops;
int[0,N] next(int[0,N] i,int[0,N] src)
  while (i < N && (!link[id][i] || i == src))</pre>
  Ł
    i++;
  }
  return i;
                                                                    Kim Larsen [100]
ARTIST Design PhD School, Beijing, 2011
```

#### Demo



ARTIST Design PhD School, Beijing, 2011



#### Optimisations

- Reducing the number of active variables
  - If variable is never used until next reset, then the value does not matter.

- Symmetry of message processes
  - The message processes are symmetric: It does not matter which is used to transfer a message.

