Predicting Space Requirements for a Stream Monitor Specification Language

David M. Cerna, Wolfgang Schreiner, and Temur Kutsia

Research Institute for Symbolic Computation (RISC) Johannes Kepler University, Linz, Austria

September 28th, 2016





・ロト ・ 理 ト ・ ヨ ト ・ ヨ ト

Intro	The Problem	Nesting	Inst & Pos	Q Trees	Algor	Conclusion
●○○○	00000	00	000000	0000000000000	000000	0000
Introd	luction					

 LogicGuard: A coordination language for runtime monitoring of network traffic.

◆□▶ ◆□▶ ◆三▶ ◆三▶ 三三 のへで

- Stream monitors are written in a fragment of predicate logic.

Intro	The Problem	Nesting	Inst & Pos	Q Trees	Algor	Conclusion
●○○○	00000	00	000000	0000000000000	000000	0000
Introd	luction					

 LogicGuard: A coordination language for runtime monitoring of network traffic.

- Stream monitors are written in a fragment of predicate logic.
- Monitor instances are evaluated using an operational semantics.
- Violations, monitor instances evaluating to false, are flagged.

Intro	The Problem	Nesting	Inst & Pos	Q Trees	Algor	Conclusion
●○○○	00000	00	000000	0000000000000	000000	0000
Introd	luction					

- LogicGuard: A coordination language for runtime monitoring of network traffic.
- Stream monitors are written in a fragment of predicate logic.
- Monitor instances are evaluated using an operational semantics.
- Violations, monitor instances evaluating to false, are flagged.
- Previous work focused on analysis of the stream history [Kutsia and Schreiner 2014], and coarse space complexity results [Cerna et al. 2016].
- $-\,$ We present an algorithm which outperforms the previous results.
- In special cases is computes precisely the space complexity of the instance set, and in general provides acceptable bounds.

・ロト ・ 日 ・ モ ・ ト ・ 日 ・ うへつ



- LogicGuard was developed to monitor boolean event streams from external sources.
- The full language does not have a concept of absolute position and messages can be assigned the same time unit depending on arrival time and coarseness of the system clock.

Intro The Problem Nesting Inst & Pos Q Trees Algor Conclusion 00000 The Logic Cuard Specification Language

The LogicGuard Specification Language

- LogicGuard was developed to monitor boolean event streams from external sources.
- The full language does not have a concept of absolute position and messages can be assigned the same time unit depending on arrival time and coarseness of the system clock.
- The <u>core language</u> [Kutsia and Schreiner 2014] on the other hand has absolute positions allowing for easier analysis of a specification's behaviour.
- Also, LogicGuard allows both value computation and internal stream construction.

Intro The Problem Nesting Inst & Pos Q Trees Algor Conclusion 000000 The Logic Cuard Specification Language

The LogicGuard Specification Language

- LogicGuard was developed to monitor boolean event streams from external sources.
- The full language does not have a concept of absolute position and messages can be assigned the same time unit depending on arrival time and coarseness of the system clock.
- The <u>core language</u> [Kutsia and Schreiner 2014] on the other hand has absolute positions allowing for easier analysis of a specification's behaviour.
- Also, LogicGuard allows both value computation and internal stream construction.
- These feature were removed from the core language to limit the express power.
- Though the core language is expressive enough to approximate the behaviour of such concepts

Example LogicGuard Specification

Nesting

Inst & Pos

```
type tcp; type message; ...
stream<tcp> IP;
stream<message> S = stream<IP> x satisfying start(@x) :
   value[seq,@x,combine]<IP> y
      with x < _ satisfying same(@x,@y) until end(@y) :
      @y ;
monitor<S> M = monitor<S> x satisfying trigger(@x) :
   exists<S> y with x < _ <=# x+5000:
   match(@x,@y);</pre>
```

Q Trees

Conclusion

▲□▶ ▲□▶ ▲□▶ ▲□▶ □ ののの

Intro

0000

The Problem

 Intro
 The Problem
 Nesting
 Inst & Pos
 Q Trees
 Algor
 Conclusion

 Core Language

▲口 ▶ ▲圖 ▶ ▲ 臣 ▶ ▲ 臣 ▶ □

æ

0000	00000		00	000000	000000000000000000000000000000000000000	000000	0000
	\sim	. •					

-		\sim		\sim	
		1	In avatuanal		an a inti an
		~			

	Atomic Formulas								
#	Transition	Constraints							
A1	$n(@V) \rightarrow d(c.2(V)))$	$V \in dom(c.2)$							
	Sequential conjunction								
C1	$\mathbf{n}(f_1 \And f_2) \to \mathbf{n}(\mathbf{n}(f_1') \And f_2)$	$f_1 \rightarrow \mathbf{n}(f_1')$							
C2	$n(f_1 \And f_2) \to d(\bot)$	$f_1 ightarrow \mathbf{d}(\bar{\perp})$							
C3	$\mathbf{n}(f_1 \And f_2) \rightarrow \mathbf{n}(f_2')$	$f_1 \rightarrow d(\top), f_2 \rightarrow n(f_2')$							
	Quantification								
Q1	$\forall_{V \in [b_1, b_2]} : f \to \mathbf{d}(\top)$	$p_1 = b_1(c)$, $p_2 = b_2(c)$, $p_1 = \infty \lor p_1 > p_2$							
Q2	$\forall_{V \in \left[b_1, b_2\right]} : f \to f'$	$p_1 = b_1(c) , p_2 = b_2(c), p_1 \neq \infty , p_1 \le p_2, \mathbf{n}(\forall_{V \in [p_1, p_2]} : f) \to f'$							
Q3	$\mathbf{n}(\forall_{V\in[p_1,p_2]}:f)\to\mathbf{n}(\forall_{V\in[p_1,p_2]}:f)$	$p < p_1$							
Q4	$\mathbf{n}(\forall_{V\in[p_1,p_2]}:f)\to f'$	$p_1 \leq p, \ n(\forall_{V \leq p_2}^{l_0} : f) \rightarrow f'$							
Q5	$n(\forall_{V\leq p_2}^I:f)\tod(\bot)$	DF							
Q6	$n(\forall_{V\leq p_2}^I:f)\tod(\top)$	$\neg DF, I'' = \emptyset, p_2 < p$							
Q7	$\mathbf{n}(\forall_{V\leq p_2}^{I}:f)\to\mathbf{n}(\forall_{V\leq p_2}^{I^{\prime\prime}}:f)$	$\neg DF, (I'' \neq \emptyset \lor p \le p_2)$							

 $DF \equiv \exists t \in \mathbb{N}, f \in \mathcal{F}, c \in \mathcal{C} : (t, f, c) \in I' \land \vdash f \rightarrow d(\perp)$

Intro	The Problem	Nesting	Inst & Pos	Q Trees	Algor	Conclusion
0000	●0000	00	000000	00000000000000	000000	
The	Problem					



▲□▶ ▲圖▶ ▲臣▶ ★臣▶ ―臣 … のへで

Intro	The Problem	Nesting	Inst & Pos	Q Trees	Algor	Conclusion
0000	○●○○○	00	000000	000000000000000	000000	0000
The	Problem					



Intro	The Problem	Nesting	Inst & Pos	Q Trees	Algor	Conclusion
0000	00●00	00	000000	000000000000000	000000	0000
The	Problem					



▲口 → ▲圖 → ▲ 臣 → ▲ 臣 → □ 臣 □

Intro	The Problem	Nesting	Inst & Pos	Q Trees	Algor	Conclusion
0000	○○○●○	00	000000	00000000000000	000000	0000
The	Problem					



Intro	The Problem	Nesting	Inst & Pos	Q Trees	Algor	Conclusion
0000	○○○○●	00	000000	00000000000000	000000	0000
The F	roblem					



▲ロト ▲御 ト ▲ 臣 ト ▲ 臣 ト の Q @





- How large does the instance set gets during evaluation.



- Getting back to the example monitor, Notice the nested variables:



- Getting back to the example monitor, Notice the nested variables:



 $-\,$ Getting back to the example monitor, Notice the nested variables:

- In [Cerna et al. 2016] we developed the concept of <u>dominating monitor</u> tranformation to remove nested variables.
- The following relationship holds between monitors and there dominated counterparts:



 $-\,$ Getting back to the example monitor, Notice the nested variables:

- In [Cerna et al. 2016] we developed the concept of <u>dominating monitor</u> tranformation to remove nested variables.
- The following relationship holds between monitors and there dominated counterparts:

Theorem

Let $M \in \mathbb{M}$. Then for all $p, n, S, S' \in \mathbb{N}$ and $s \in \{\top, \bot\}^{\omega}$ such that $T(M) \multimap_{p,s,n} S$ and $T(D(M)) \multimap_{p,s,n} S'$, we have $S \leq S'$.

Intro The Problem Nesting Inst & Pos Q Trees Algor Conclusion Dominating Monitor Example

 $-\,$ The dominating monitor of

$$\begin{aligned} \forall_{0 \leq x} : \forall_{y \in [x+1,x+5]} : ((\forall_{z \in [y,x+3]} : \neg @z \& @z) \& G(x,y)) \\ G(x,y) &= \forall_{w \in [x+2,y+2]} : (\neg @y \& (\forall_{m \in [y,w]} : \neg @x \& @m)) \end{aligned}$$

is the following monitor

$$\forall_{0 \le x} : \forall_{y \in [x+1,x+5]} : ((\forall_{z \in [x+1,x+3]} : \neg @z \& @z) \& G(x,y)) \\ G(x,y) = \forall_{w \in [x+2,x+7]} : (\neg @y \& (\forall_{m \in [x+1,x+7]} : \neg @x \& @m))$$

 $-\,$ The dominating monitor of

$$\begin{aligned} \forall_{0 \leq x} : \forall_{y \in [x+1,x+5]} : ((\forall_{z \in [y,x+3]} : \neg @z \& @z) \& G(x,y)) \\ G(x,y) &= \forall_{w \in [x+2,y+2]} : (\neg @y \& (\forall_{m \in [y,w]} : \neg @x \& @m)) \end{aligned}$$

is the following monitor

$$\forall_{0 \le x} : \forall_{y \in [x+1,x+5]} : ((\forall_{z \in [x+1,x+3]} : \neg @z \& @z) \& G(x,y)) \\ G(x,y) = \forall_{w \in [x+2,x+7]} : (\neg @y \& (\forall_{m \in [x+1,x+7]} : \neg @x \& @m))$$

 We retrain from going into the details of the transformation and will only use dominating monitors for the rest of this talk.



 Runtime representation size equals instances kept in memory while evaluating a monitor.

Intro The Problem Nesting Inst & Pos Q Trees Algor Conclusion Dealing with the Runtime Representation Size

- Runtime representation size equals instances kept in memory while evaluating a monitor.
- Consider the following simple monitor:

$$\forall_{0 \leq x} : \forall_{y \in [x, x+4]} : \forall_{z \in [x, x+4]} : \forall_{r \in [x, x+4]} : @r$$

Intro The Problem Nesting Inst & Pos Q Trees Algor Conclusion Dealing with the Runtime Representation Size

- Runtime representation size equals instances kept in memory while evaluating a monitor.
- Consider the following simple monitor:

$$\forall_{0 \le x} : \forall_{y \in [x, x+4]} : \forall_{z \in [x, x+4]} : \forall_{r \in [x, x+4]} : \mathbb{O}r$$

We can simplify its representation:

 $\left[0,4\right]\left[0,4\right]\left[0,4\right]$

◆□▶ ◆□▶ ◆三▶ ◆三▶ 三三 のへで

Intro The Problem Nesting Inst & Pos Q Trees Algor Conclusion Dealing with the Runtime Representation Size

- Runtime representation size equals instances kept in memory while evaluating a monitor.
- Consider the following simple monitor:

$$\forall_{0 \leq x} : \forall_{y \in [x, x+4]} : \forall_{z \in [x, x+4]} : \forall_{r \in [x, x+4]} : @r$$

We can simplify its representation:

$$\left[0,4\right]\left[0,4\right]\left[0,4\right]$$

(日) (個) (E) (E) (E)

- Now let us consider its behaviour as it is evaluated.

Intro	The Problem	Nesting	Inst & Pos	Q Trees	Algor	Conclusion
0000		00	○●○○○○	00000000000000	000000	0000
Initial	State					

<u>Initial state</u> [0,4][0,4][0,4]:1 [0,4][0,4]:0 [0,4]:0

◆□▶ ◆□▶ ◆三▶ ◆三▶ 三三 のへで

Intro 0000	The Problem	Nesting 00	Inst & Pos ○○●○○○	Q Trees 00000000000000	Algor 000000	Conclusion
Eval	uation					



イロト 不得 とくほ とくほ とうほう

Intro	The Problem	Nesting	Inst & Pos	Q Trees	Algor	Conclusion
0000	00000	00	○○●○○○	00000000000000	000000	
Eval	lation					



- Notice that we did not add new instances.
- How does this relate to true evaluation?

slide 16/43

Intro The Problem Nesting Inst & Pos Q Trees Algor Conclusion

Instance to Position Mapping

 It turns out that there is a mapping from the evaluation of a single instance at various positions to the evaluation of multiple instances at a single position.



Instance to Position Mapping



(日)、

Instance to Position Mapping



 $-\,$ Notice that going to the next position does not change anything

イロト イポト イヨト イヨト

Intro The Problem Nesting One Position Mapping, Next Position



(日)、

Intro The Problem Nesting Inst & Pos Q Trees Algor Conclusio



 Essentially, we only need to look at the behaviour of one instance up to the largest upper bound. This is the key to the algorithm.

・ロト ・ 雪 ト ・ ヨ ト

slide 19/43



The above concept translates to the following algorithm (assuming no variable nesting).

▲□▶ ▲□▶ ▲□▶ ▲□▶ ▲□ ● ● ●

```
function SR(\langle A, a, b, Q \rangle)

if A = \infty then

return \infty

else

return \sum_{i=0}^{A-1}SR(\langle A, a, b, Q \rangle, i)

end if

end function
```



The highlighted object is a representation of a monitor specification.

▲□▶ ▲□▶ ▲□▶ ▲□▶ □ ののの

```
function SR( \langle A, a, b, Q \rangle)

if A = \infty then

return \infty

else

return \sum_{i=0}^{A-1} SR(\langle A, a, b, Q \rangle, i)

end if

end function
```



▲□▶ ▲□▶ ▲□▶ ▲□▶ □ ののの

The highlighted object is the largest upper bound.

```
function SR( \langle A, a, b, Q \rangle)

if A = \infty then

return \infty

else

return \sum_{i=0}^{A-1} SR(\langle A, a, b, Q \rangle, i)

end if

end function
```

Inst & Pos Q Trees

Constructing a Quantifier Tree

Consider the following monitor specification: _

$$M = \forall_{0 \le x} : \forall_{y \in [x+1,x+5]} : ((\forall_{z \in [x+1,x+3]} : \neg @z \& @z) \& G(x,y))$$
$$G(x,y) = \forall_{w \in [x+2,x+7]} : (\neg @y \& (\forall_{m \in [x+1,x+7]} : \neg @x \& @m))$$

- A quantifier tree of *M* is constructed as follows:

 Intro
 The Problem
 Nesting
 Inst & Pos
 Q Trees
 Algor
 Conclusion

 Constructing a Quantifier Tree

◆□▶ ◆圖▶ ◆臣▶ ◆臣▶ ─臣

$$QT(\forall_{0 \le x} : M') = (0, 0, QT(M')) \text{ where } M' = \\ \forall_{y \in [x+1, x+5]} : ((\forall_{z \in [x+1, x+3]} : \neg @z \& @z) \& G(x, y))$$

<0,0,QT(M')>

 Intro
 The Problem
 Nesting
 Inst & Pos
 Q Trees
 Algor
 Conclusion

 Constructing a Quantifier Tree

 $QT(M') = (1, 5, QT(M_1))$ where $M_1 = (\forall_{z \in [x+1, x+3]}: \neg @z \& @z) \& G(x, y)$

▲□▶ ▲□▶ ▲□▶ ▲□▶ □ ののの

<0,0,QT(M')>

 $QT(M_1) = QT(M_l) \cup QT(M_r)$ where $M_l = (\forall_{z \in [x+1,x+3]}: \neg @z \& @z)$ and $M_r = G(x,y)$

◆□▶ ◆□▶ ◆三▶ ◆三▶ 三三 のへで



 Intro
 The Problem
 Nesting
 Inst & Pos
 Q Trees
 Algor
 Conclusion

 Constructing a Quantifier Tree

$QT(M_l) = (1, 5, \emptyset)$



 Intro
 The Problem
 Nesting
 Inst & Pos
 Q Trees
 Algor
 Conclusion

 0000
 00000
 00
 00000
 00000
 00000
 00000

 Constructing a Quantifier Tree

 $QT(M_r) = (2, 7, QT(M'_r))$ where $M'_r = \neg @y \& (\forall_{m \in [x+1,x+7]} : \neg @x \& @m)$



$QT(M'_r) = (1,7,\emptyset)$



- This is the quantifier tree of monitor M.

















- We will refer to this quantifier tree as QT(M). Now we compute

$$\sum_{i=0}^{6} SR(QT(M), i).$$

slide 33/43



- Rather then computing the entire sum

$$\sum_{i=0}^{6} \mathsf{SR}(QT(M), i).$$

We will look into a specific example.

- SR(QT(M), 5) We will also ignore the first node $\langle 7, 0, 0, QT(M') \rangle$

- At position 5 the whole interval will unroll.

$$\langle 7, 1, 5, QT(M_1) \rangle = \begin{cases} \langle 7, 1, 1, QT(M_1) \rangle \\ \langle 7, 2, 2, QT(M_1) \rangle \\ \langle 7, 3, 3, QT(M_1) \rangle \\ \langle 7, 4, 4, QT(M_1) \rangle \\ \langle 7, 5, 5, QT(M_1) \rangle \end{cases}$$

The number of generated instances is computed using the following formula:

$$1 + \min\{i, b\} - a$$

◆□▶ ◆□▶ ◆三▶ ◆三▶ 三三 少へ⊙

which in our case is $1+\min\left\{5,5\right\}-1=5.$



- The specific instances don't really matter.
- We can just write the following

 $5 \cdot SR(QT(M_1), 5)$

- But notice that $QT(M_1)$ branches.



Intro The Problem Nesting Inst & Pos Q Trees Algor Conclusion

The Optimization and Branching

- Normally $5 \cdot SR(QT(M_1), 5) =$ $5 \cdot (SR(QT(M_l), 5) + SR(QT(M_r), 5)).$
- However $QT(M_l) = \langle 5, 1, 5, \emptyset \rangle$, the upper bound is equal to the position.
- This means $SR(QT(M_l), 5) = 0$, and we optimize the computation by ignoring it. Thus,

 $5 \cdot SR(QT(M_1), 5) = 5 \cdot (SR(QT(M_l), 5) + SR(QT(M_r), 5)) =$ $5 \cdot SR(QT(M_r), 5)$

▲□▶ ▲□▶ ▲□▶ ▲□▶ □ ののの

Computing the Right Branch

Nesting

- Moving on to $(7, 2, 7, QT(M'_r))$ we compute the interval size as

Inst & Pos

$$1 + \min{\{5,7\}} - 2 = 4$$

Q Trees

Algor

000000

Conclusion

Thus, we get

$$5 \cdot SR(QT(M_r), 5) = 20 \cdot SR(QT(M_r'), 5)$$

- As the last step we get

 $SR(QT(M),5) = 20 \cdot SR(QT(M'_r),5) = 20 \cdot 5 = 100$



Intro	The Problem	Nesting	Inst & Pos	Q Trees	Algor	Conclusion
0000		00	000000	0000000000000	○○○○○●	0000
Algori	thm					

```
- The algorithm is as follows:
```

```
function SR((A, a, b, Q), i)
    cil \leftarrow 1 + \min\{i, b\} - a
    if cil \leq 0 \& b \geq a then
        return 1
    else
        return 0
    end if
    if i > b then
        inst \leftarrow 0
    else
        inst \leftarrow 1
    end if
    for all aqt' = (A', a', b', Q') \in Q do
        if i < A' then
            inst \leftarrow inst + cil \cdot SR(aqt', i)
        end if
    end for
    return inst
end function
```

- It has a running time of O(n) in terms of formula size.

★白▶ ★課▶ ★注▶ ★注▶ 一注

 Intro
 The Problem
 Nesting
 Inst & Pos
 Q Trees
 Algor
 Conclusion

 Experimental Results: Artificial

- We ran the algorithm on the following monitor specifications:

$$\begin{array}{ll} \forall_{0 \leq x} : \forall_{y \in [x, x+80]} : \forall_{z \in [x, x+80]} : @z & (1a) \\ \forall_{0 \leq x} : \forall_{y \in [x, x+80]} : \forall_{z \in [x, y]} : @z & (1b) \\ \forall_{0 \leq x} : \forall_{y \in [x, x+40]} : \forall_{z \in [x, x+80]} : @z & (2a) \\ \forall_{0 \leq x} : \forall_{y \in [x, x+40]} : \forall_{z \in [x, y+40]} : @z & (2b) \end{array}$$



(日)、

 $\exists \rightarrow$

 Intro
 The Problem
 Nesting
 Inst & Pos
 Q Trees
 Algor
 Conclusion

 00000
 000000
 000000
 000000
 000000
 000000
 000000

 Experimental Results: Realistic

 We ran the algorithm on the following monitor written in the full specification language:

exists<S> z with y < _ <=# y+4000: IsEven(#z);</pre>





- The Run time representation size for general monitor specifications is bounded by our algorithm.
 - Dealing with nested variables would provide precise results for all monitor specifications
 - Currently we are investigating the implications of these results for writing monitor specifications.

(日) (個) (E) (E) (E)

- Looking for more optimal ways of writing monitor specifications.
- The next measure we are going to tackle concerning logical guard is the number of stream accesses per message.

0000 00000 00 000000 000000 000 0	Intro	The Problem	Nesting	Inst & Pos	Q Trees	Algor	Conclusion
							0000

Thank you for your time.

◆□▶ ◆□▶ ◆臣▶ ◆臣▶ 三臣 - のへぐ