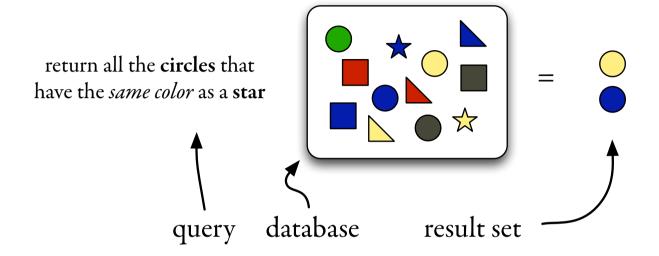
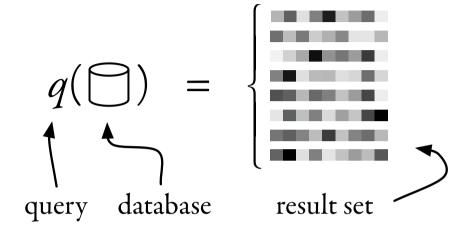
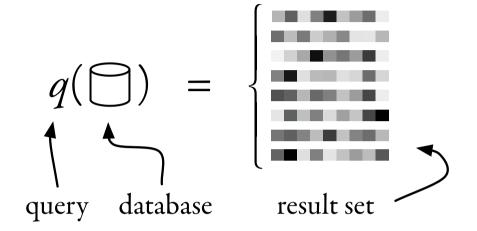
Reasoning on words and trees with data

On decidable automata on data words and data trees in relation to satisfiability of LTL and XPath.

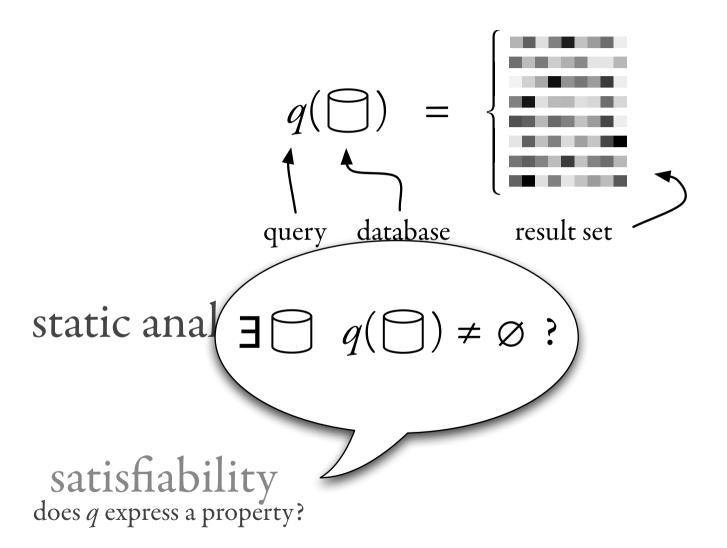
Diego Figueira

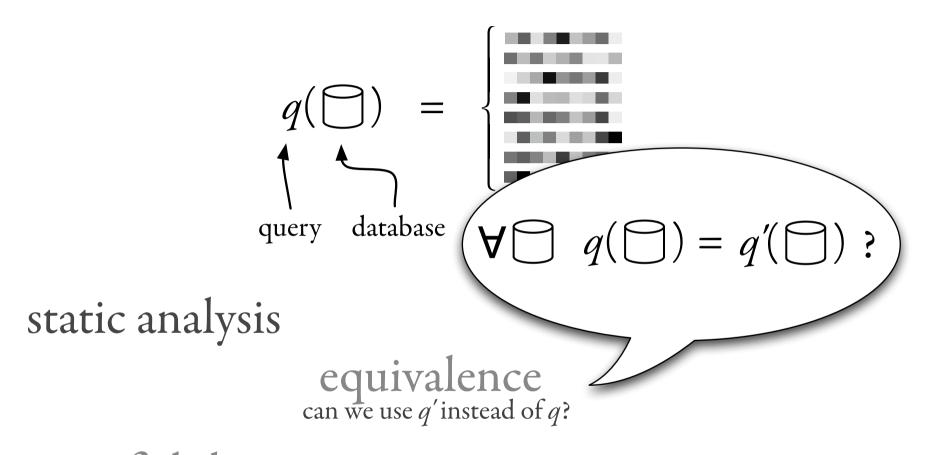




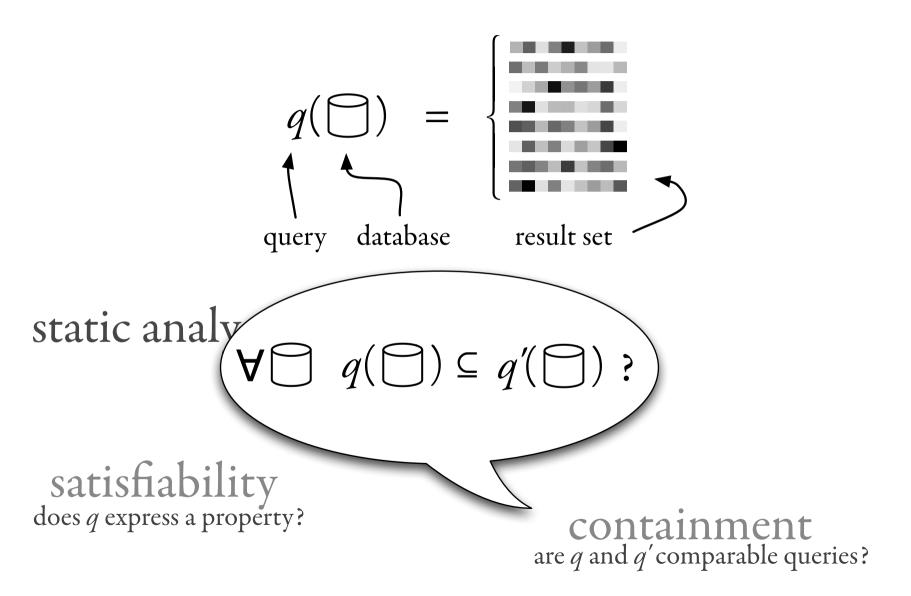


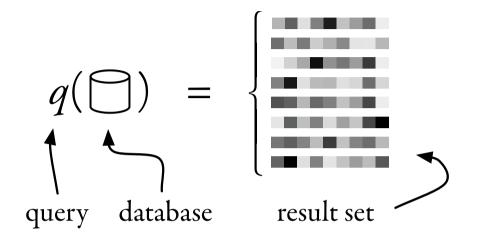
static analysis





satisfiability does q express a property?



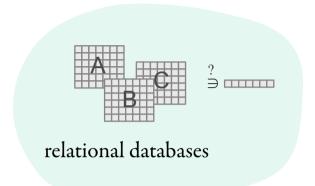


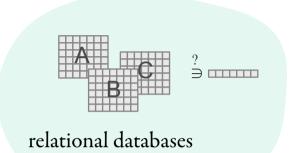
static analysis

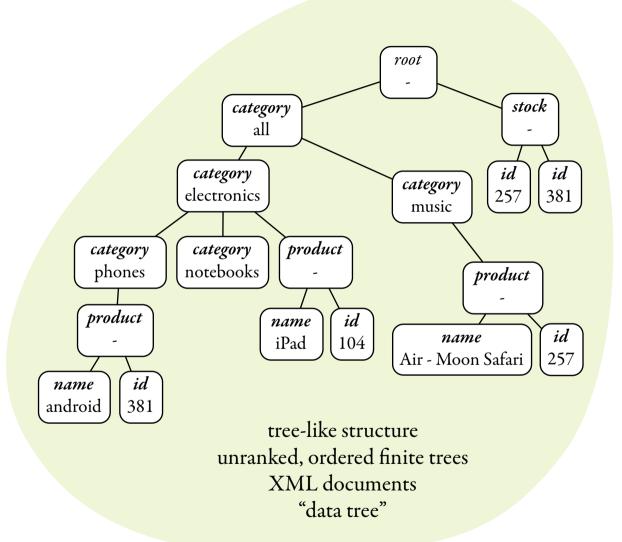
equivalence can we use q'instead of q?

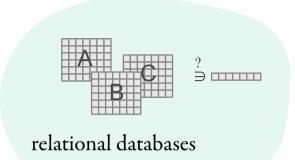
satisfiability does q express a property?

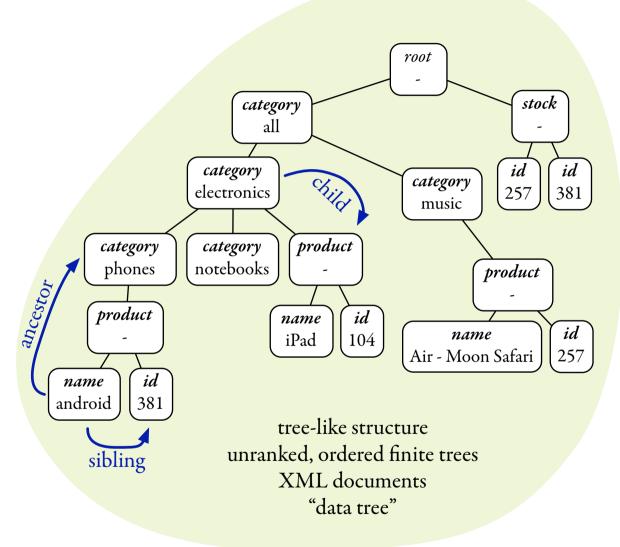
containment are q and q' comparable queries?

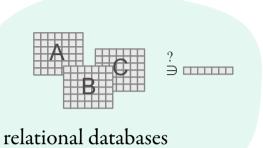


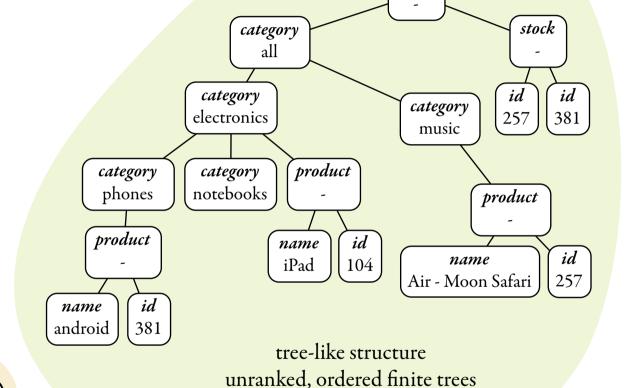






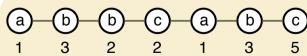




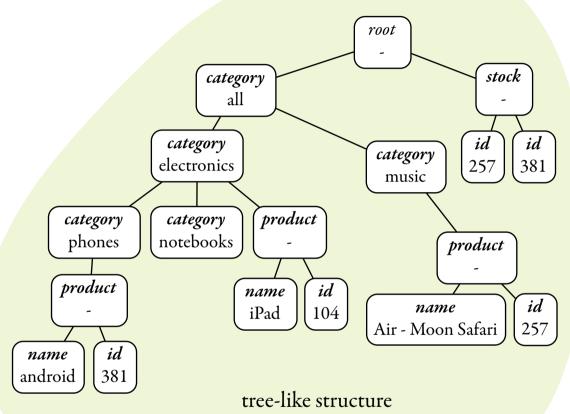


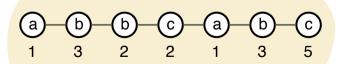
XML documents "data tree"

root



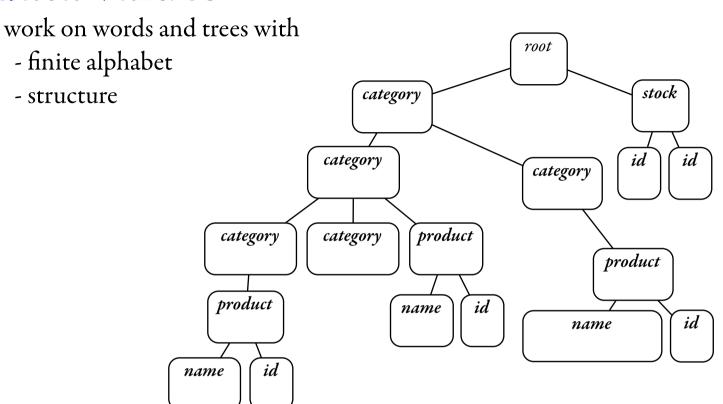
special case of a data tree temporal databases "data word"





special case of a data tree temporal databases "data word" tree-like structure
unranked, ordered finite trees
XML documents
"data tree"

work on words and trees with root- finite alphabet stock category - structure category category product category category product product idname idname idname

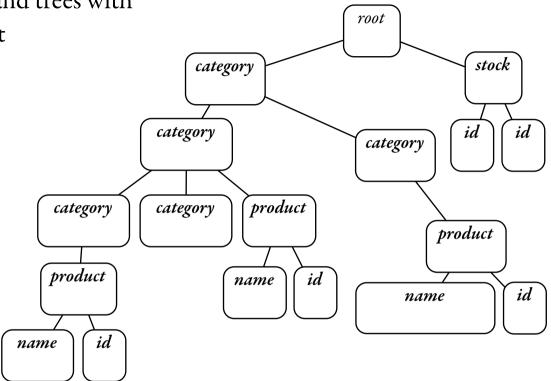


need to perform joins, to test for equality of data values

work on words and trees with

- finite alphabet

- structure



"return all products in stock"

need to perform joins, to test for equality of data values

work on words and trees with root - finite alphabet stock category - structure all category ididcategory 257 electronics 381 music product category category phones notebooks product product idname idname 104 iPad Air - Moon Safari 257 idname android 381

"return all products in stock"

need to perform joins, to test for equality of data values

what's out there?

(alternating) register automata epebble automata

Kaminski, Francez, Neven, Schwentick, Vianu, Tan, Demri, Lazić, Jurdziński, . . .



(alternating) timed automata Alur, Dill, . . .

real time logics

Alur, Henzinger, Harel, Lichtenstein, Pnueli, ...

 FO^2

Bojańczyk, Muscholl, Schwentick, Segoufin, David, . . .

XPath

Benedikt, Fan, Geerts

LTL with registers

Demri, Lazić, Nowak patterns conjunctive queries

David, Björklund, Martens, Schwentick

Hybrid logics

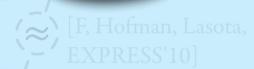
Description Logics

with concrete domains

what's out there?

(alternating) register automata pebble automata

Kaminski, Francez, Neven, Schwentick, Vianu, Tan, Demri, Lazić, Jurdziński, . . .



(alternating) timed automata Alur, Dill, . . .

real time logics

Alur, Henzinger, Harel, Lichtenstein, Pnueli, ...

 FO^2

Bojańczyk, Muscholl, Schwentick, Segoufin, David, . . .

XPath

Benedikt, Fan, Geerts

LTL with registers

Demri, Lazić, Nowak patterns conjunctive queries

David, Björklund, Martens, Schwentick

Hybrid logics

Description Logics

with concrete domains

rLTL data-words

XPath data-trees

data-aware languages automata for word-tree-structured dbs

rLTL data-words





XPath data-trees

data-aware languages automata for word-tree-structured dbs

rLTL XPath data-trees

equivalence satisfiability inclusion

data-aware languages automata for word-tree-structured dbs

rLTL data-words





XPath data-trees

equivalence satisfiability inclusion

data-aware

languages automata

for

word-& structured dbs

rLTL data-words





XPath data-trees

equivalence

satisfiability

inclusion

rLTL(F,X)

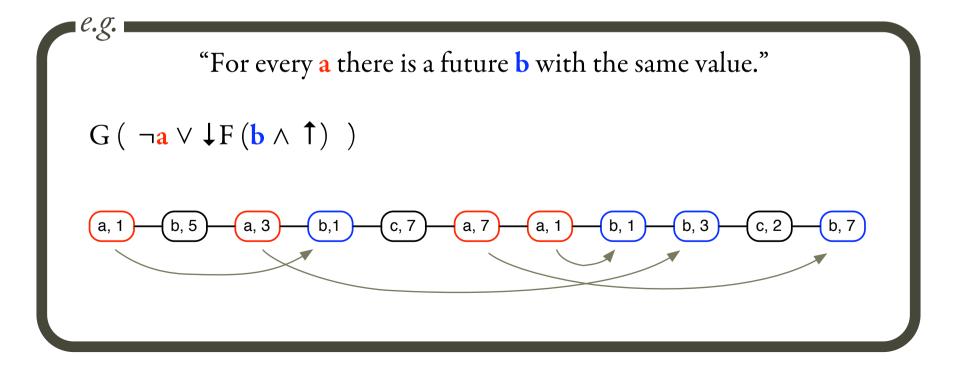
register-LTL over data words

$$\varphi := a \mid F \varphi \mid X \varphi \mid \varphi \lor \varphi \mid \varphi \land \varphi \mid \neg \varphi \mid \downarrow \varphi \mid \uparrow$$

$$\text{tests current datum}$$

$$\text{against the stored}$$

$$\text{stores current datum}$$



lower bounds

We knew...
[Demri / Lazić]

rLTL(X, F) is decidable, non primitive recursive

rLTL(X, F, F⁻¹) is undecidable

lower bounds

[Demri / Lazić]

We knew... Now we know ... The one-step is not necessary to obtain the lower-bounds!

rLTL(X, F) is decidable, non primitive recursive

 $rLTL(X, F, F^{-1})$ is undecidable

lower bounds

[Demri / Lazić]

We knew... Now we know ... The one-step is not necessary to obtain the lower-bounds!

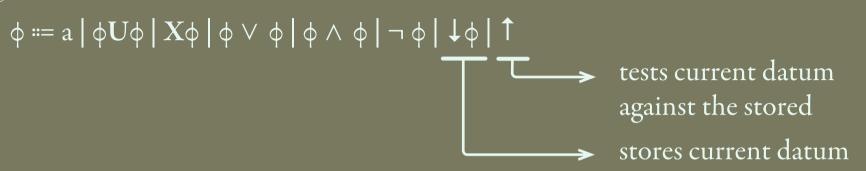
rLTL(X, F) is decidable, non primitive recursive

[F, Segoufin, MFCS'09]

rLTL(X, F, F⁻¹) is undecidable

even for a very simple fragment

register-LTL over data words





register-LTL over data words

$$\phi := a \left| \phi \mathbf{U} \phi \right| \mathbf{X} \phi \left| \phi \lor \phi \right| \phi \land \phi \left| \neg \phi \right| \downarrow \phi \middle| \uparrow \middle| \forall \le \phi \middle| \exists \ge \phi$$
Exists a future data value s.t. ϕ .
For all past data value, ϕ .



register-LTL over data words

$$\phi \coloneqq a \mid \phi \mathbf{U} \phi \mid \mathbf{X} \phi \mid \phi \lor \phi \mid \phi \land \phi \mid \neg \phi \mid \downarrow \phi \mid \uparrow \mid \forall \leq \phi \mid \exists \geq \phi$$

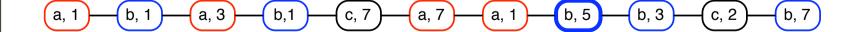
$$\text{Exists a future data value s.t. } \phi.$$

$$\text{For all past data value, } \phi.$$

■ e.g

"There is a b with a data value different from any previous element."

$$F(\neg b \lor \forall \leq (\neg \uparrow))$$



register-LTL over data words

$$\phi := a \left| \phi \mathbf{U} \phi \right| \mathbf{X} \phi \left| \phi \lor \phi \right| \phi \land \phi \left| \neg \phi \right| \mathbf{1} \phi \right| \uparrow \left| \forall \leq \phi \right| \exists \geq \phi$$
Exists a future data value s.t. ϕ .
For all past data value, ϕ .



"There exists a data value that is contained in a b but not in an a."

$$\exists \geq (F(b \land \uparrow) \land \neg F(a \land \uparrow))$$

register-LTL over data words

$$\phi := a \left| \phi \mathbf{U} \phi \right| \mathbf{X} \phi \left| \phi \vee \phi \right| \phi \wedge \phi \left| \neg \phi \right| \mathbf{J} \phi \right| \uparrow \left| \forall \leq \phi \right| \exists \geq \phi$$
Exists a future data value s.t. ϕ .

For all past data value, ϕ .

The satisfiability problem for **positive** $rLTL(U,X) + \forall \leq + \exists \geq$ is **decidable**.

register-LTL over data words

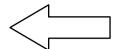
$$\phi := a \left| \phi \mathbf{U} \phi \right| \mathbf{X} \phi \left| \phi \vee \phi \right| \phi \wedge \phi \left| \neg \phi \right| \mathbf{1} \phi \left| \uparrow \right| \forall \leq \phi \mid \exists \geq \phi$$
Exists a future data value s.t. ϕ .

For all past data value, ϕ .

The satisfiability problem for **positive** $rLTL(U,X) + \forall \leq + \exists \geq$ is **decidable**.

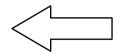
The satisfiability problem for $rLTL(U,X) + \forall \leq + \exists \geq$ is **undecidable**.





satisfiability of rLTL(U,X)

- Decidable emptiness problem
- With non-primitive recursive complexity
- © Closed under complementation, intersection, union

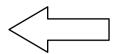


satisfiability of rLTL(U,X)

- Decidable emptiness problem
- With non-primitive recursive complexity
- © Closed under complementation, intersection, union

ARA + guess + spread

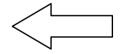
- © Still decidable emptiness problem
- No longer closed under complementation
- Can't be closed under complement preserving decidability



satisfiability of rLTL(U,X)

- Decidable emptiness problem
- With non-primitive recursive complexity
- Closed under complementation, intersection, union

$$ARA + guess + spread$$
 $satisfiability of positive rLTL(U,X) + $\forall \leq + \exists \geq$$



- Still decidable emptiness problem
- ② No longer closed under complementation
- Can't be closed under complement preserving decidability



Alternating register automata one-way 1 register

states:

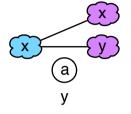


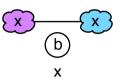


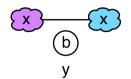
initial state:

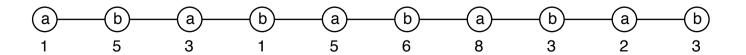
final states:

transitions











Alternating register automata one-way 1 register

states:



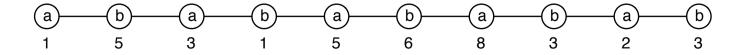


initial state:



final states:







Alternating register automata one-way
1 register

states:

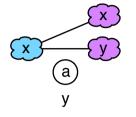


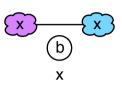


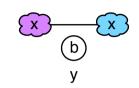
initial state:

final states:

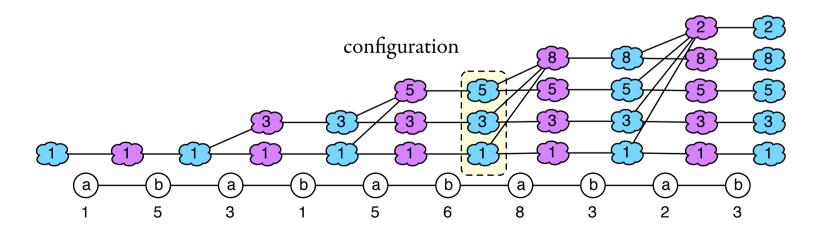
transitions







A configuration = a set of running threads (state, datum)



the string is $(ab)^*$, and all the a's have different data values



Alternating register automata one-way

states:



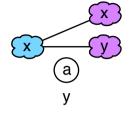


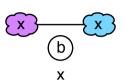
initial state:

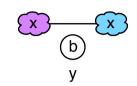


1 register

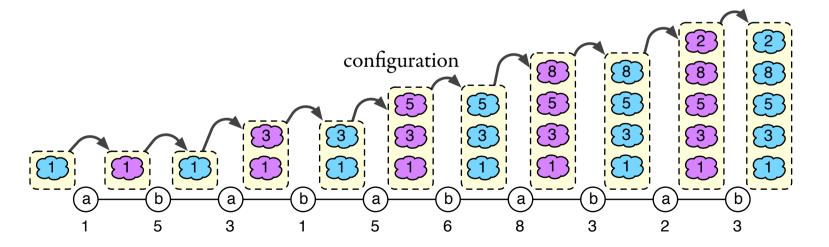
transitions







A configuration = a set of running threads (state, datum)



the string is $(ab)^*$, and all the a's have different data values



Alternating register automata one-way 1 register

states:

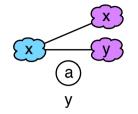


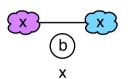


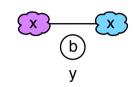
initial state:

final states:

transitions



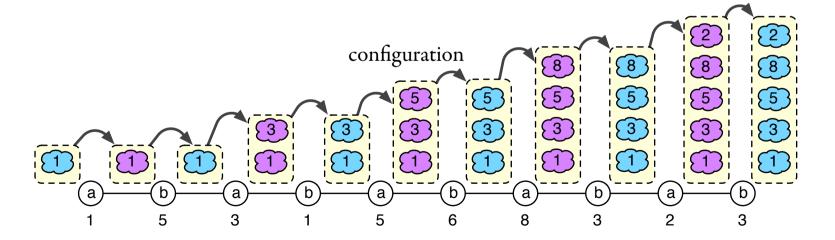




guess(q)spread(q)

Nondeterministically guess a data value and store it in the register.

For every thread (q,d), create a new one (q',d).



the string is $(ab)^*$, and all the a's have different data values

examples

"There exists a data value that is contained in a b but not in an a."

- 1. Guess a data value d
- 2. Create two threads that test
 - a. there is a future position **b** with datum **d**
 - b. there is not a future position a with datum d

examples

"There exists a data value that is contained in a b but not in an a."

- 1. Guess a data value d
- 2. Create two threads that test
 - a. there is a future position **b** with datum **d**
 - b. there is not a future position a with datum d

"There is a **b** with a data value different from any previous element."

Create two threads

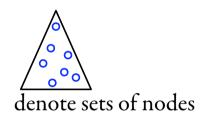
- 1. One that at every position create a thread that
 - a. stores current data value, and
 - b. advances until the last position using some state q
- 2. Other that
 - a. guesses a b position of the word
 - b. spreads all threads with *q* to threads with *p*

(state *p* checks that the stored data value is different from the current value)

XPath

node expressions

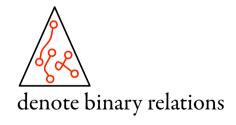
$$\varphi, \psi ::= \mathbf{a} \mid \neg \varphi \mid \varphi \land \psi \mid \alpha = \beta \mid \alpha \neq \beta \mid \alpha? \quad \mathbf{a} \in A$$



path expressions

$$\alpha, \beta ::= \varepsilon \mid \alpha \beta \mid \alpha \cup \beta \mid \alpha [\varphi] \mid o$$

$$o \in \{ \downarrow, \vdots, \rightarrow, \cdots, \uparrow, \uparrow, \leftarrow, \leftarrow, \cdots \}$$



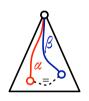
'→': one-step

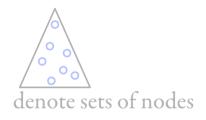
'... : multistep (transitive closure of '---')

XPath

node expressions

$$\varphi, \psi ::= \mathbf{a} \mid \neg \varphi \mid \varphi \land \psi \mid \alpha = \beta \mid \alpha \neq \beta \mid \alpha? \quad \mathbf{a} \in A$$

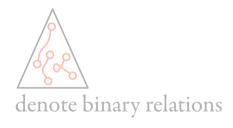




path expressions

$$\alpha, \beta ::= \varepsilon \mid \alpha \beta \mid \alpha \cup \beta \mid \alpha [\varphi] \mid o$$

$$o \in \{ \downarrow, \vdots, \rightarrow, \cdots, \uparrow, \vdots, \leftarrow, \leftarrow, \cdots \}$$



 \longrightarrow : one-step

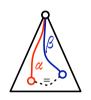
'... ': multistep (transitive closure of '---')

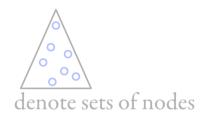
XPath

we note "XPath(=,
$$\rightarrow$$
, \cdots , \uparrow)"

node expressions

$$\varphi, \psi ::= \mathbf{a} \mid \neg \varphi \mid \varphi \land \psi \mid \alpha = \beta \mid \alpha \neq \beta \mid \alpha? \quad \mathbf{a} \in A$$

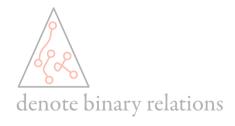




path expressions

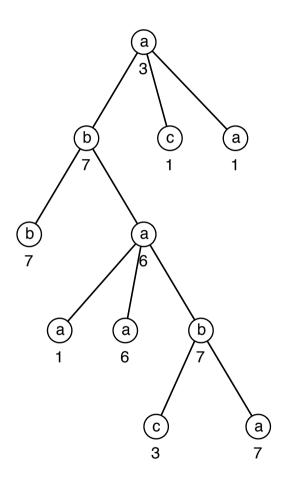
$$\alpha, \beta ::= \varepsilon \mid \alpha \beta \mid \alpha \cup \beta \mid \alpha [\varphi] \mid o$$

$$o \in \{ \downarrow, \vdots, \rightarrow, \cdots, \uparrow, \uparrow, \leftarrow, \leftarrow, \cdots \}$$



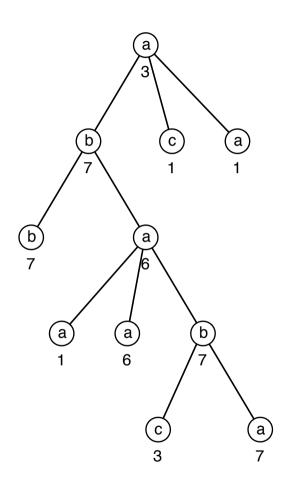
 $' \rightarrow ': one-step$

'··· ': multistep (transitive closure of '→')



$$\c \vdots [a] = \c \vdots [b]$$

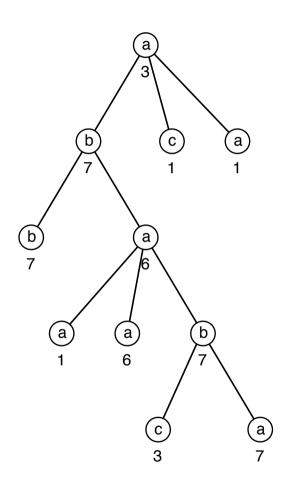
There are an *a* and a *b* with the same datum



$$\vdots [c] \neq \vdots [c]$$

There are an *a* and a *b* with the same datum

There are two *c* with different data values



$$\vdots [a] = \vdots [b]$$

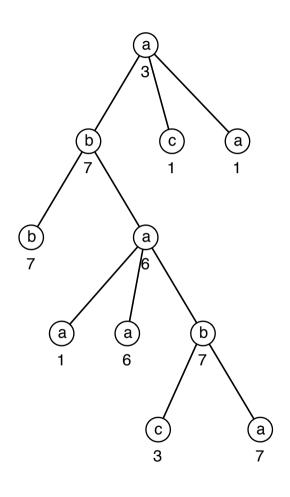
$$\vdots [c] \neq \vdots [c]$$

$$\neg \left(\ \vdots \ [b] \neq \ \vdots \ [b] \ \right)$$

There are an *a* and a *b* with the same datum

There are two *c* with different data values

All *b* have the same data value



$$\vdots [a] = \vdots [b]$$

$$\vdots [c] \neq \vdots [c]$$

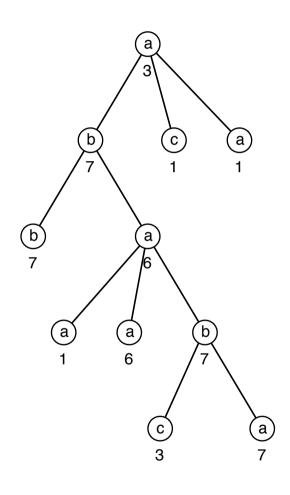
$$\neg \left(\ \ \vdots \ [b] \neq \ \ \vdots \ [b] \ \right)$$

There are an *a* and a *b* with the same datum

There are two *c* with different data values

All *b* have the same data value

There is no data value shared by a *b* and a *c*



$$\vdots [a] = \vdots [b]$$

$$\vdots [c] \neq \vdots [c]$$

$$\neg \left(\ \vdots \ [b] \neq \ \vdots \ [b] \ \right)$$

$$\neg \left(\ \ \vdots \ [a] = \longrightarrow \cdots \right) \ \ \vdots \ [a] \] \)$$

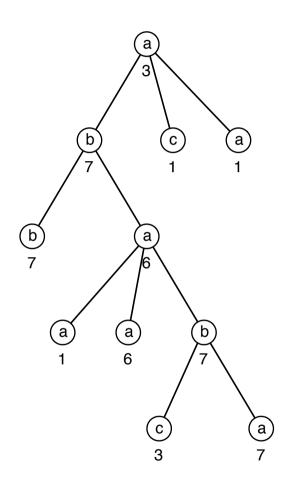
There are an *a* and a *b* with the same datum

There are two *c* with different data values

All *b* have the same data value

There is no data value shared by a *b* and a *c*

The *c* constitutes a primary key



$$\vdots [c] \neq \vdots [c]$$

$$\neg \left(\ \ \ \, \dot{\vdots} \ [b] \neq \ \ \, \dot{\vdots} \ [b] \ \right)$$

$$\neg \left(\ \ \vdots \ [a] = \longrightarrow \cdots \right) \ \ \vdots \ [a] \] \)$$

$$\neg (\vdots [a \land \neg (\epsilon \neq : [b])])$$

There are an *a* and a *b* with the same datum

There are two *c* with different data values

All *b* have the same data value

There is no data value shared by a *b* and a *c*

The *c* constitutes a primary key

Every *a* has an ancestor *b* with different value

Satisfiability of XPath(=, \downarrow , $\stackrel{\downarrow}{,}$, \rightarrow ,..., \uparrow , $\stackrel{\uparrow}{,}$, \leftarrow ,...) is undecidable.

Satisfiability of XPath(=, \downarrow , $\stackrel{\downarrow}{\bullet}$, \rightarrow ,..., \uparrow , $\stackrel{\uparrow}{\bullet}$; \leftarrow ,...) is undecidable.

Satisfiability of XPath(=, \downarrow , \rightarrow , \rightarrow , \rightarrow , \uparrow , \uparrow , \leftarrow , \bullet ...) is undecidable.

Our contribution:

```
Satisfiability of XPath(=, \ , \ , \ , \ , \ , \ , \ , \ , \ \ ) is undecidable. MFCS'09]
```

Our contribution:

```
Satisfiability of XPath(=, \downarrow, \updownarrow, \rightarrow, \uparrow, \uparrow, \uparrow, \uparrow) is undecidable. MFCS'09]
```

Our contribution:

```
Satisfiability of XPath(=, \downarrow, \downarrow, \rightarrow, \uparrow, \uparrow, \uparrow) is undecidable. MFCS'09]
```

Satisfiability of XPath(=, \downarrow , \vdots) is ExpTime-complete. [F, PODS'09]

Our contribution:

```
Satisfiability of XPath(=, \downarrow, \downarrow, \rightarrow, \uparrow, \uparrow, \uparrow, \uparrow) is undecidable. MFCS'09]
```

Satisfiability of XPath(=, \downarrow , \updownarrow) is ExpTime-complete. [F, PODS'09]

Satisfiability of XPath(=, \downarrow , $\stackrel{\downarrow}{\bullet}$, \rightarrow ,...) is decidable. [F, ICDT'10]

Our contribution:

```
Satisfiability of XPath(=, \downarrow, \downarrow, \rightarrow, \uparrow, \uparrow, \uparrow) is undecidable. MFCS'09]
```

Satisfiability of XPath(=, \downarrow , $\stackrel{\downarrow}{\bullet}$) is ExpTime-complete. [F, PODS'09]

Satisfiability of XPath(=, \downarrow , $\stackrel{\downarrow}{\bullet}$, \rightarrow ,...) is decidable. [F, ICDT'10]

Satisfiability of XPath(=, \downarrow , $\stackrel{\cdot}{\downarrow}$, \uparrow , $\stackrel{\uparrow}{:}$) is decidable. [F, Segoufin, STACS'11]

Our contribution:

```
Satisfiability of XPath(=, \downarrow, \rightarrow, \rightarrow, \uparrow, \uparrow, \uparrow) is undecidable. MFCS'09]
```

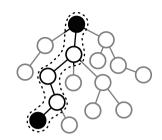
Satisfiability of XPath(=, \downarrow , \vdots) is ExpTime-complete. [F, PODS'09]

Satisfiability of XPath(=, \downarrow , \vdots , \rightarrow , \cdots) is decidable. [F, ICDT'10]

Satisfiability of XPath(=, \downarrow , \updownarrow , \uparrow , \updownarrow) is decidable. [F, Segoufin, STACS'11]

Satisfiability of $XPath(=, \downarrow, \stackrel{\vdots}{\downarrow} \rightarrow, \cdots)$

```
Satisfiability of \mathcal{E}-XPath(=, \downarrow, \vdots \longrightarrow, \cdots) is decidable. [Lazić, Jurdziński, '07] \longrightarrow only tests: \varepsilon = \alpha
```



Satisfiability of
$$\mathcal{E}$$
-XPath(=, \downarrow , $\stackrel{!}{\downarrow} \rightarrow$,...) is decidable. [Lazić, Jurdziński, '07] only tests: $\varepsilon = \alpha$

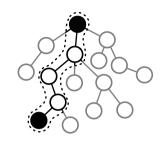


We show

Satisfiability of XPath(=, \downarrow , $\vdots \rightarrow$,...) is decidable. [F, ICDT'10] lower bound: non primitive recursive

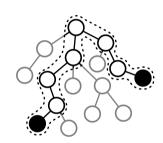


Satisfiability of
$$\mathcal{E}$$
-XPath(=, \downarrow , $\stackrel{!}{\downarrow} \rightarrow$,...) is decidable. [Lazić, Jurdziński, '07] \bullet only tests: $\varepsilon = \alpha$

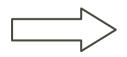


We show

Satisfiability of XPath(=, \downarrow , $\vdots \rightarrow$,...) is decidable. [F, ICDT'10] lower bound: non primitive recursive







ATRA
+ guess + spread

extension of ARA on trees alternating automata
1 register top-down, unranked decidable emptiness pb

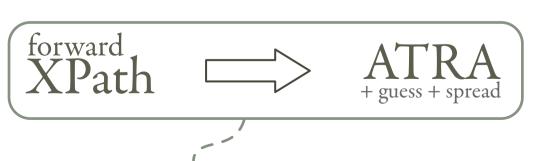
Satisfiability of
$$\mathcal{E}$$
-XPath(=, \downarrow , \vdots \longrightarrow , \cdots) is decidable. [Lazić, Jurdziński, '07] only tests: $\varepsilon = \alpha$



We show

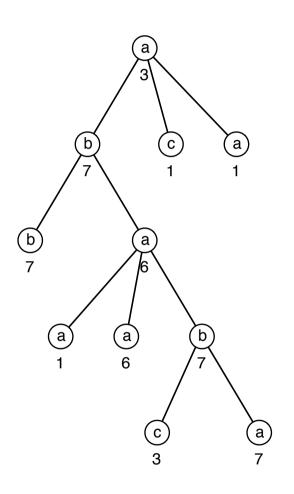
Satisfiability of XPath(=, \downarrow , $\stackrel{\downarrow}{\bullet} \rightarrow$,...) is decidable. [F, ICDT'10] lower bound: non primitive recursive





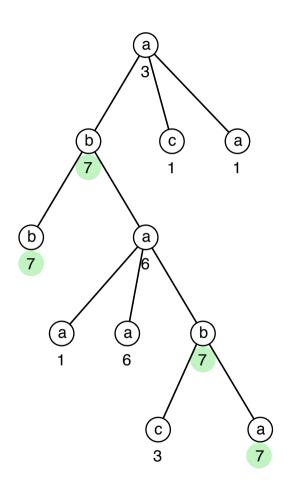
extension of ARA on trees alternating automata
1 register top-down, unranked decidable emptiness pb

difficulty: ATRA are not closed under complementation



$$\vdots [a] = \vdots [b]$$

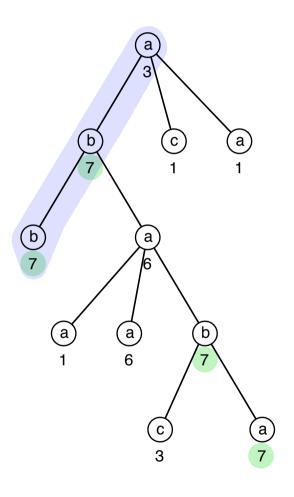
There are an *a* and a *b* with the same datum



$$\vdots [a] = \vdots [b]$$

Guess the data value 7.

There are an *a* and a *b* with the same datum

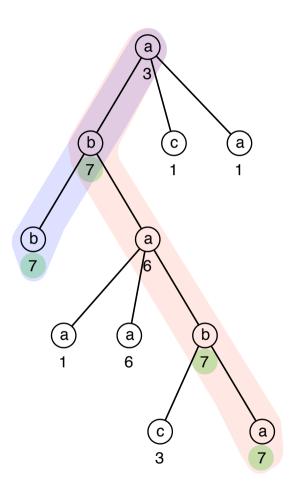


$$\vdots [a] = \vdots [b]$$

There are an *a* and a *b* with the same datum

Guess the data value 7.

Check that it can be accessed with " 🕹 [b] ".



$$\vdots [a] = \vdots [b]$$

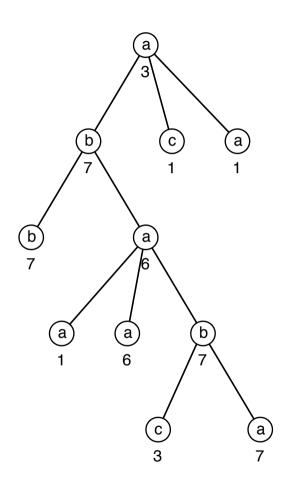
There are an *a* and a *b* with the same datum

Guess the data value 7.

Check that it can be accessed with " . [b] ".

Check that it can be accessed with " : [a] ".

how do we code...?



$$\vdots [a] = \vdots [b]$$

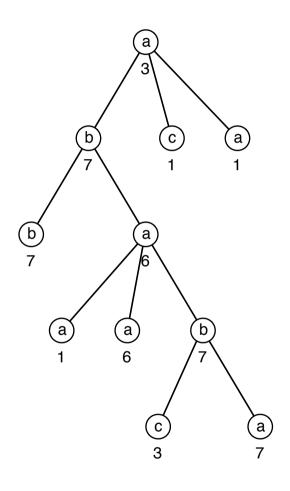
$$\vdots [c] \neq \vdots [c]$$

There are an *a* and a *b* with the same datum

There are two *c* with different data values

All **b** have the same data value

how do we code...?



$$\vdots [a] = \vdots [b]$$

There are an *a* and a *b* with the same datum

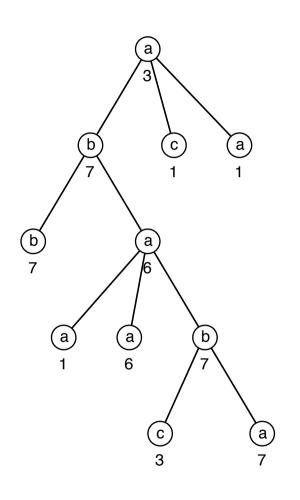
There are two *c* with different data values

All *b* have the same data value

$$\neg \left(\ \ \vdots \ [b] = \ \ \vdots \ [c] \ \right)$$

There is no data value shared by a *b* and a *c*

how do we code...?



$$\vdots [a] = \vdots [b]$$

$$\vdots [c] \neq \vdots [c]$$

$$\neg \left(\ \ \vdots \ [b] \neq \ \ \ \vdots \ [b] \ \right)$$

build automaton A such that

if $T \vDash \varphi$, then A accepts T if A accepts T, then there is T' = f(A,T) such that $T' \vDash \varphi$

XPath: inclusion&equivalence of path expressions

XPath: inclusion&equivalence of path expressions

More domain specific operations number arithmetic string substring

XPath: inclusion&equivalence of path expressions

More domain specific operations number arithmetic string substring

Quest for an expressive formalism logic automaton with low complexity

XPath: inclusion&equivalence of path expressions

More domain specific operations number arithmetic string substring

Quest for an expressive formalism logic automaton with low complexity

restricting the logics/automata restricting the class of models to obtain decidable/tractable procedures

1. Horizontal

```
rLTL + \forall \leq + \exists \geq is decidable
ARA(guess, spread) is decidable
rLTL(F) is non primitive recursive
rLTL(F, F<sup>-1</sup>) is undecidable
```

1. Horizontal

rLTL + $\forall \leq + \exists \geq$ is decidable ARA(guess, spread) is decidable rLTL(F) is non primitive recursive rLTL(F, F⁻¹) is undecidable

2. Downward

 $XPath(=, \downarrow, \vdots)$ is ExpTime-complete

DD automaton is 2ExpTime

1. Horizontal

rLTL + $\forall \leq + \exists \geq$ is decidable ARA(guess, spread) is decidable rLTL(F) is non primitive recursive rLTL(F, F⁻¹) is undecidable

2. Downward

XPath(=, ↓, ₺) is ExpTime-complete

DD automaton is 2ExpTime

3. Forward

XPath(=, \downarrow , \vdots , \rightarrow ,...) is decidable, non primitive recursive

ATRA(guess, spread) is decidable

1. Horizontal

rLTL + $\forall \leq + \exists \geq$ is decidable ARA(guess, spread) is decidable rLTL(F) is non primitive recursive rLTL(F, F⁻¹) is undecidable

2. Downward

 $XPath(=, \downarrow, \ddagger)$ is ExpTime-complete

DD automaton is 2ExpTime

3. Forward

XPath(=, \downarrow , \vdots , \rightarrow ,...) is decidable, non primitive recursive

ATRA(guess, spread) is decidable

4. Vertical

XPath(=, \downarrow , \updownarrow , \uparrow , \updownarrow) is decidable, non primitive recursive

BUDA is decidable

1. Horizontal

rLTL + $\forall \leq + \exists \geq$ is decidable ARA(guess, spread) is decidable rLTL(F) is non primitive recursive rLTL(F, F⁻¹) is undecidable

2. Downward

XPath(=, ↓, ₺) is ExpTime-complete

DD automaton is 2ExpTime

3. Forward

XPath(=, \downarrow , \vdots , \rightarrow ,...) is decidable, non primitive recursive

ATRA(guess, spread) is decidable

4. Vertical

XPath(=, \downarrow , \updownarrow , \uparrow , \updownarrow) is decidable, non primitive recursive

BUDA is decidable