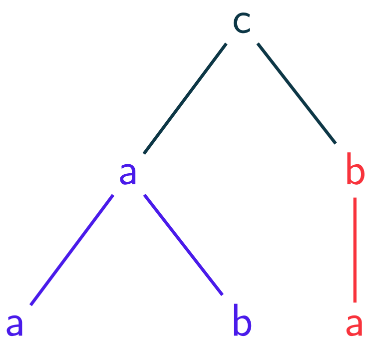
Progress on weak validation of streamed trees

Corentin Barloy

University of Lille

Processing streamed trees

XML encoding of trees:



$\langle c \rangle$

$\quad\langle a \rangle$

$\qquad\langle a \rangle\langle / a \rangle$

$\qquad\langle b \rangle\langle / b \rangle$

$\quad\langle / a \rangle$

$\quad\langle b \rangle$

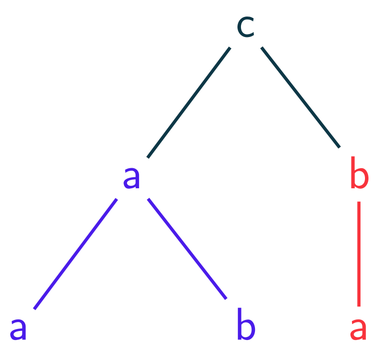
$\qquad\langle a \rangle\langle / a \rangle$

$\quad\langle / b \rangle$

$\langle / c \rangle$

Processing streamed trees

XML encoding of trees:



$\langle c \rangle$

$\quad \langle a \rangle$

$\quad\quad \langle a \rangle \langle / a \rangle$

$\quad\quad \langle b \rangle \langle / b \rangle$

$\quad \langle / a \rangle$

$\quad \langle b \rangle$

$\quad\quad \langle a \rangle \langle / a \rangle$

$\quad \langle / b \rangle$

$\langle / c \rangle$

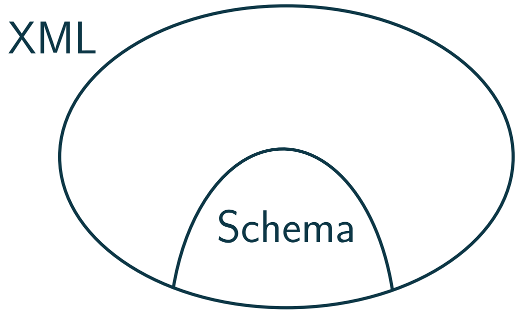
Validation: Does a XML document belongs to a given regular language of trees?

Weak validation

Weak validation: the document is well-formed.

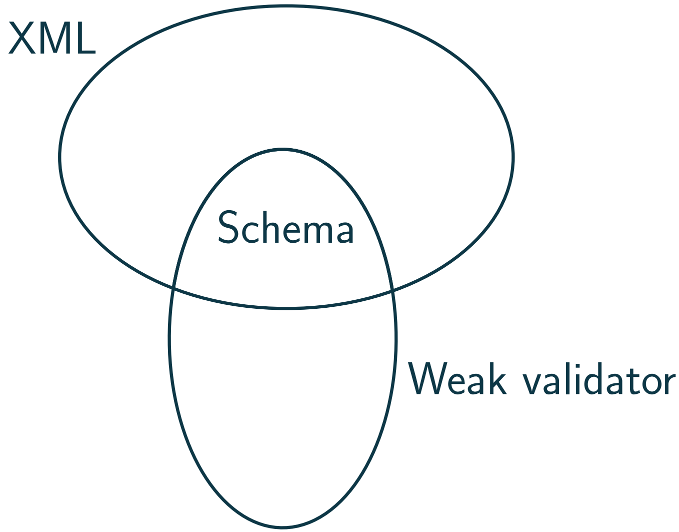
Weak validation

Weak validation: the document is well-formed.



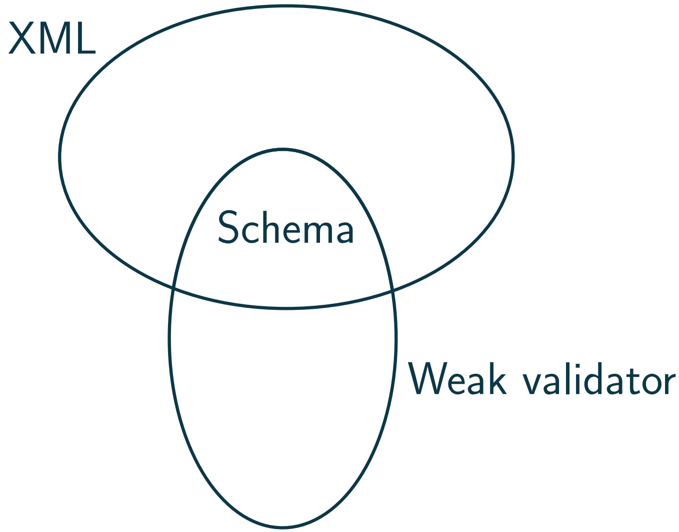
Weak validation

Weak validation: the document is well-formed.



Weak validation

Weak validation: the document is well-formed.



Question

What is the space complexity of weakly validating documents in streaming?

Space complexity

Segoufin, Vianu:

What can be weakly validated in constant space?

Space complexity

Segoufin, Vianu:

What can be weakly validated in constant space?

- Solved for fully recursive DTDs.

Space complexity

Segoufin, Vianu:

What can be weakly validated in constant space?

- Solved for fully recursive DTDs.

B., Paperman, Murlak: Study of Regular Path Queries:

Space complexity

Segoufin, Vianu:

What can be weakly validated in constant space?

- Solved for fully recursive DTDs.

B., Paperman, Murlak: Study of Regular Path Queries:

- Space trichotomy between $\Theta(1)$, $\Theta(\log(h))$, $\Theta(h)$.

Space complexity

Segoufin, Vianu:

What can be weakly validated in constant space?

- Solved for fully recursive DTDs.

B., Paperman, Murlak: Study of Regular Path Queries:

- Space trichotomy between $\Theta(1)$, $\Theta(\log(h))$, $\Theta(h)$.
- Allow vectorisation: rsonpath (Gienieczko, Paperman).

Space complexity

Segoufin, Vianu:

What can be weakly validated in constant space?

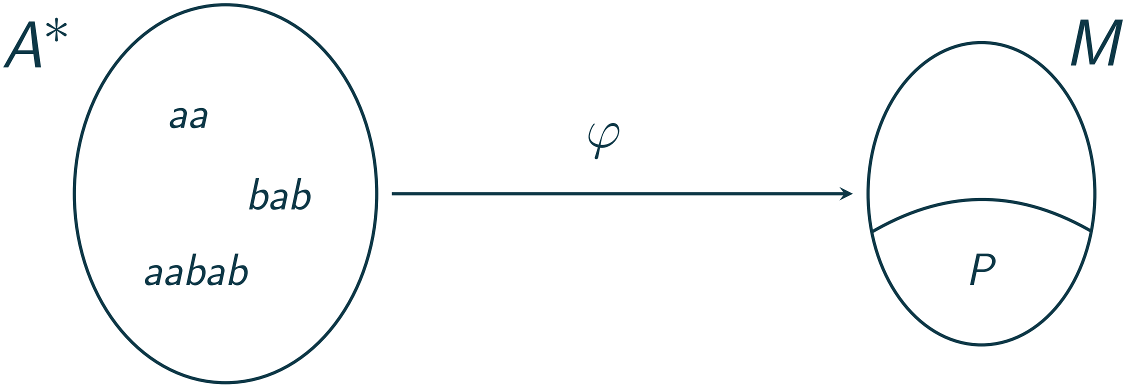
- Solved for fully recursive DTDs.

B., Paperman, Murlak: Study of Regular Path Queries:

- Space trichotomy between $\Theta(1)$, $\Theta(\log(h))$, $\Theta(h)$.
- Allow vectorisation: rsonpath (Gienieczko, Paperman).

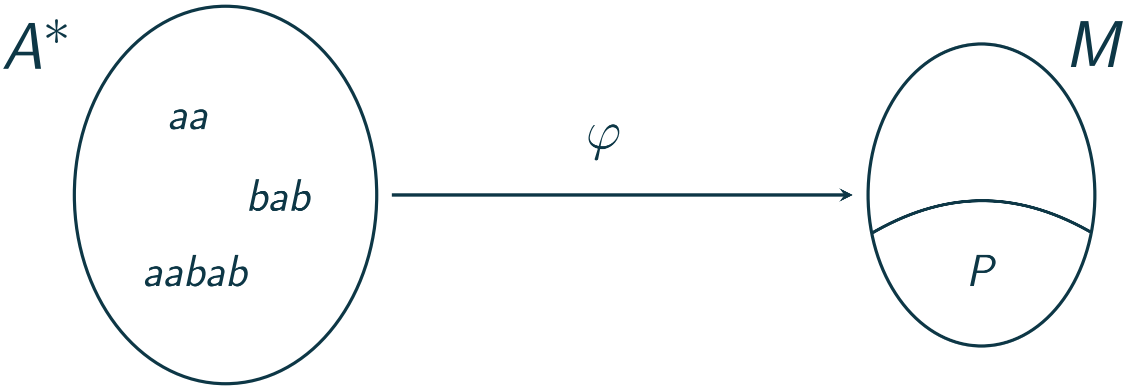
Does the trichotomy holds in general?

Monoïds



$\mathcal{L}$ recognized by M: $\quad \mathcal{L} = \varphi^{-1}(P)$

Monoïds

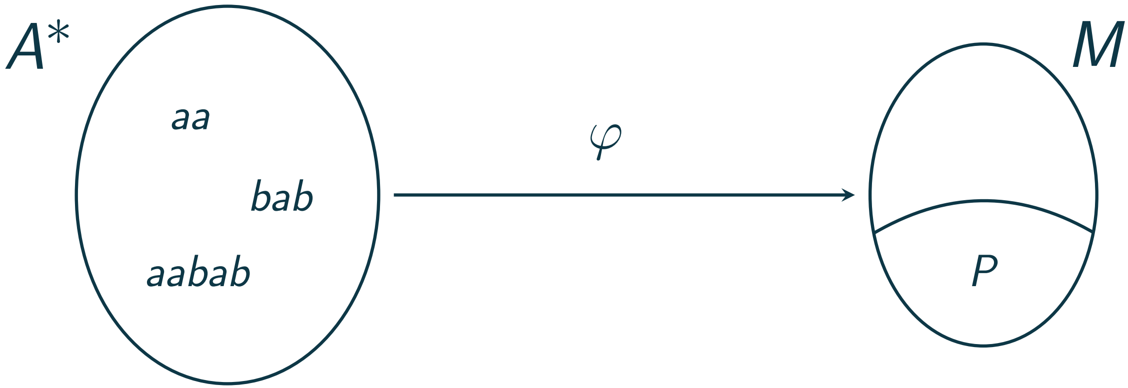


$\mathcal{L}$ recognized by M: $\mathcal{L} = \varphi^{-1}(P)$

Syntactic monoïd:

- Congruence on A^*
- Computed from the minimal automaton
- Smallest monoïd recognizing $\mathcal{L}$

Monoïds



$\mathcal{L}$ recognized by M: $\quad \mathcal{L} = \varphi^{-1}(P)$

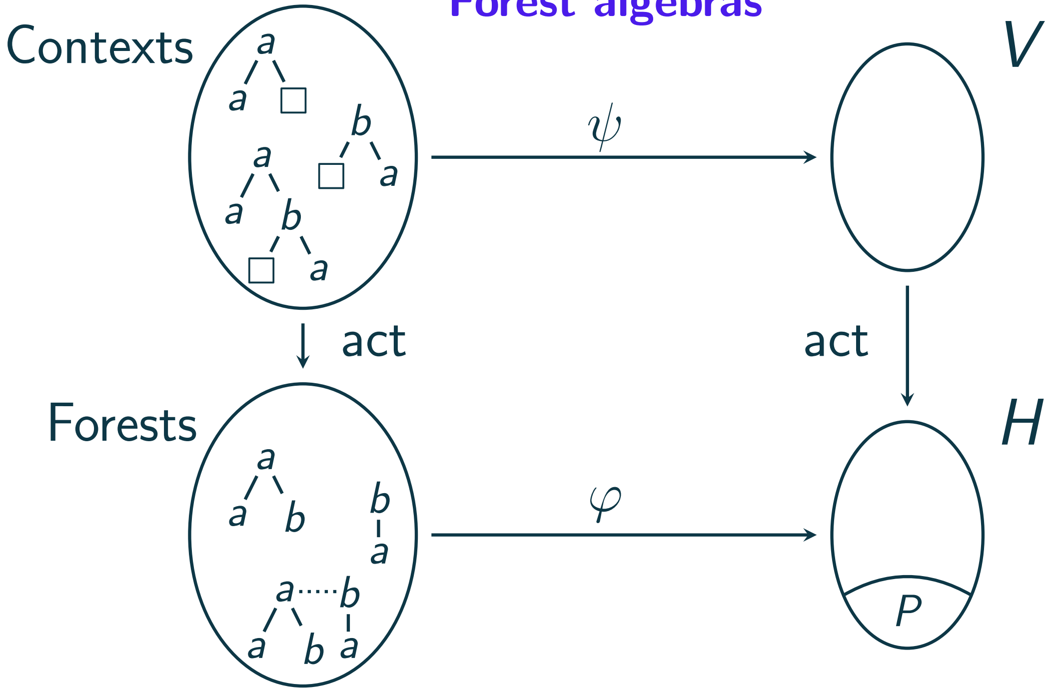
Syntactic monoïd:

- Congruence on A^*
- Computed from the minimal automaton
- Smallest monoïd recognizing $\mathcal{L}$

Varieties: membership to a class can be read on the syntactic monoïd.

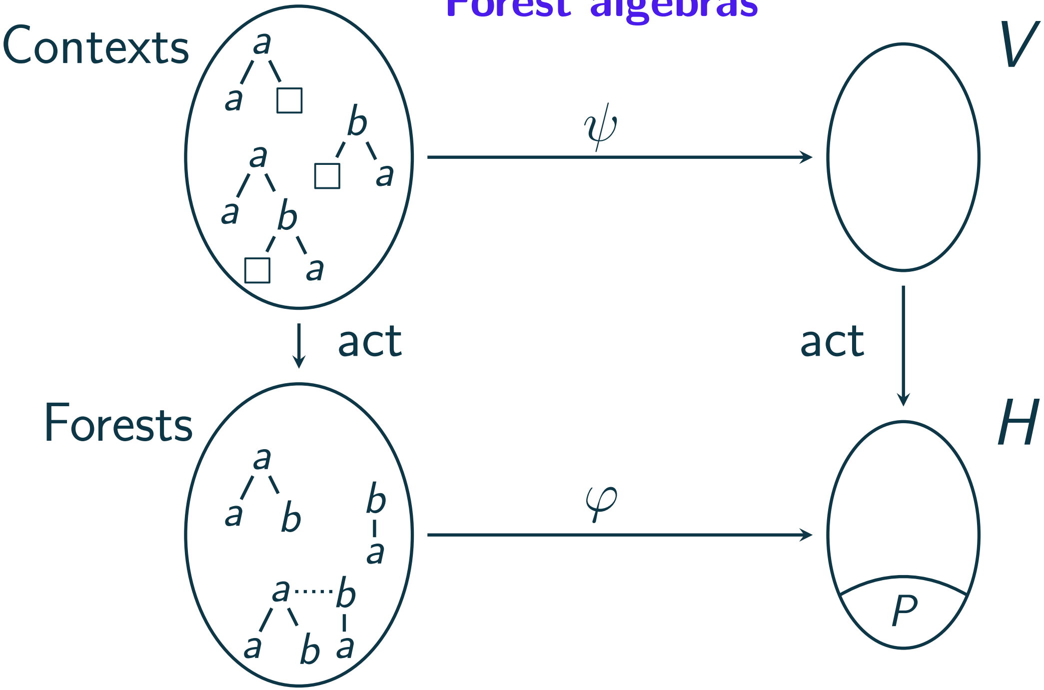
Equational theory.

Forest algebras



$\mathcal{L}$ recognized by (V, H): $\mathcal{L} = \varphi^{-1}(P)$

Forest algebras

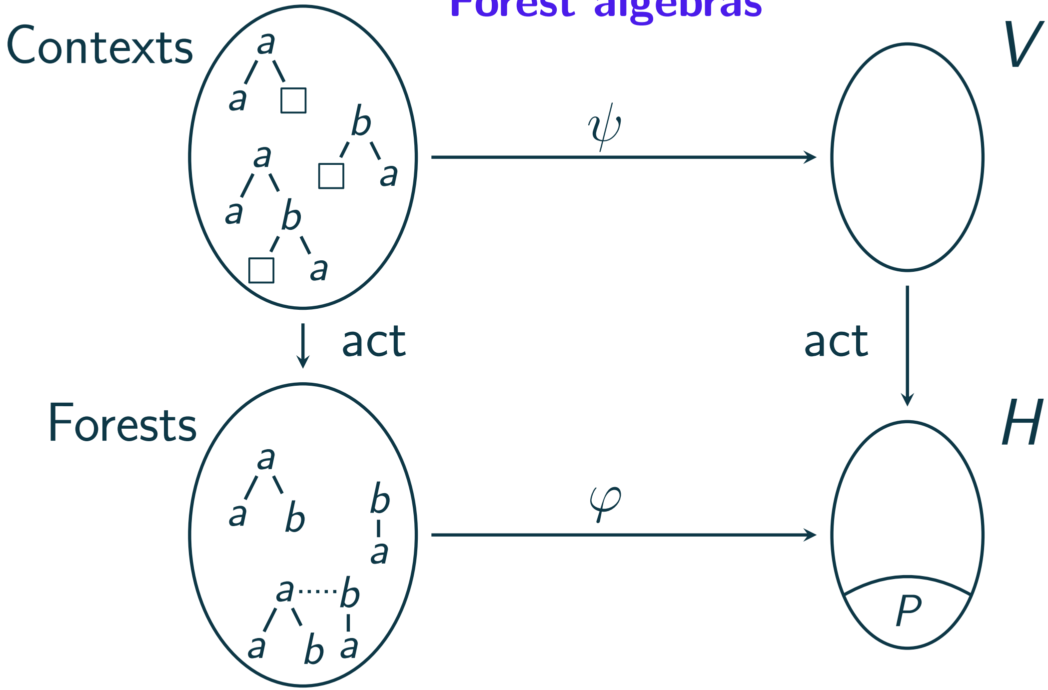


$\mathcal{L}$ recognized by (V, H): $\quad \mathcal{L} = \varphi^{-1}(P)$

Syntactic algebra:

- Congruences on contexts and trees
- Computed form the minimal automaton
- Smallest algebra recognizing $\mathcal{L}$

Forest algebras



$\mathcal{L}$ recognized by (V, H): $\mathcal{L} = \varphi^{-1}(P)$

Syntactic algebra:

- Congruences on contexts and trees
- Computed form the minimal automaton
- Smallest algebra recognizing $\mathcal{L}$

Varieties

Equations

Monoïdal forest algebras

A forest algebra (V, H) is monoïdal iff:

For all $v \in V$, there exists $x, y \in H$ such that for all $h \in H$

$$v \cdot h = x + h + y$$

Monoïdal forest algebras

A forest algebra (V, H) is monoïdal iff:

For all $v \in V$, there exists $x, y \in H$ such that for all $h \in H$

$$v \cdot h = x + h + y$$

Theorem

$\mathcal{L}$ is weakly validatable in $\Theta(1) \iff \mathcal{L}$ is recognized by a monoïdal forest algebra

Monoïdal forest algebras

A forest algebra (V, H) is monoïdal iff:

For all $v \in V$, there exists $x, y \in H$ such that for all $h \in H$

$$v \cdot h = x + h + y$$

Theorem

$\mathcal{L}$ is weakly validatable in $\Theta(1) \iff \mathcal{L}$ is recognized by a monoïdal forest algebra

Problem: Not closed by subalgebras!

The class of monoïdal algebras is not a variety.

Membership cannot be checked on the syntactic algebra.

Monoïdal forest algebras

A forest algebra (V, H) is monoïdal iff:

For all $v \in V$, there exists $x, y \in H$ such that for all $h \in H$

$$v \cdot h = x + h + y$$

Theorem

$\mathcal{L}$ is weakly validatable in $\Theta(1) \iff \mathcal{L}$ is recognized by a monoïdal forest algebra

Problem: Not closed by subalgebras!

The class of monoïdal algebras is not a variety.

Membership cannot be checked on the syntactic algebra.

Ex: Trees of depth 2

Monoïdal forest algebras

A forest algebra (V, H) is monoïdal iff:

For all $v \in V$, there exists $x, y \in H$ such that for all $h \in H$

$$v \cdot h = x + h + y$$

Theorem

$\mathcal{L}$ is weakly validatable in $\Theta(1) \iff \mathcal{L}$ is recognized by a monoïdal forest algebra

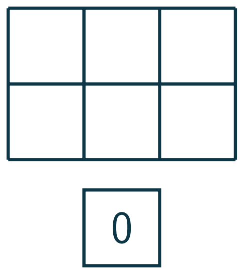
Problem: Not closed by subalgebras!

The class of monoïdal algebras is not a variety.

Membership cannot be checked on the syntactic algebra.

Ex: Trees of depth 2

Solved for algebras with H that is 0-simple :



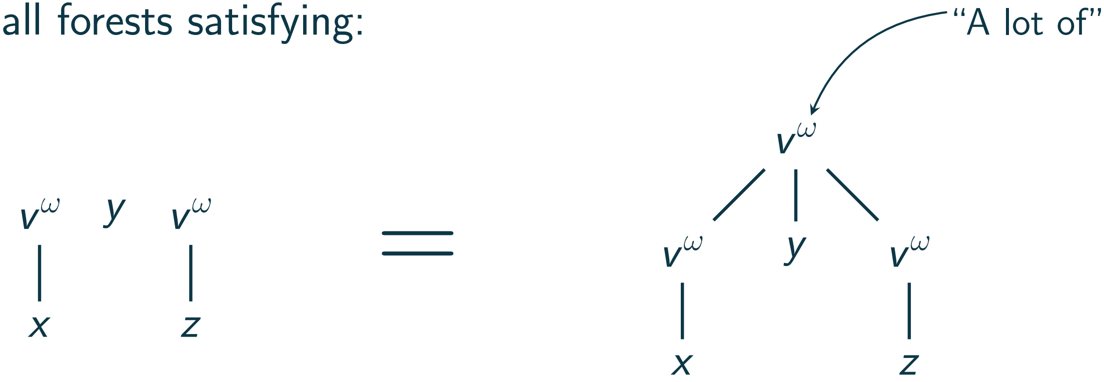
A candidate set of equations

$\mathcal{V}$ is the variety of all forests satisfying:

“A lot of”

A candidate set of equations

$\mathcal{V}$ is the variety of all forests satisfying:



Fact

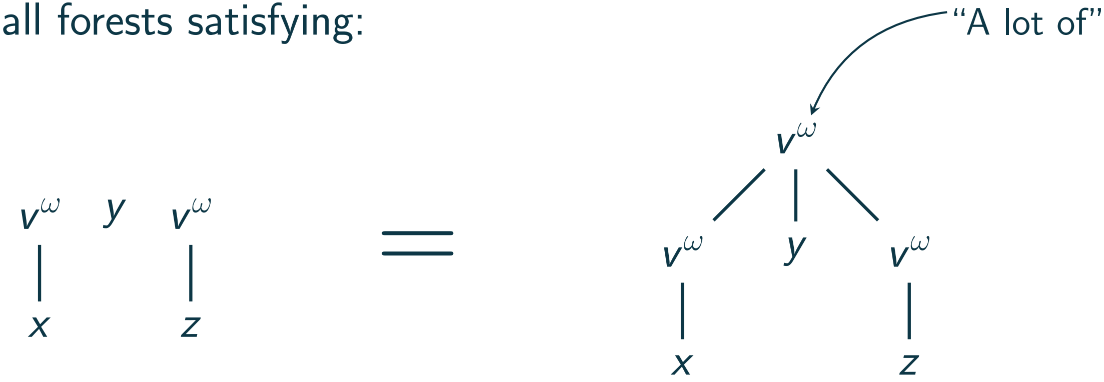
$$\text{Weakly validatable} \subseteq \mathcal{V}$$

Proof

Mimics a pumping argument.

A candidate set of equations

$\mathcal{V}$ is the variety of all forests satisfying:



Fact

$$\text{Weakly validatable} \subseteq \mathcal{V}$$

Proof

Mimics a pumping argument.

Do we need more equations?

Summary

- A space trichotomy for the weak validation problem?
- Can we decide if an action has a monoïdal extension?

Summary

- A space trichotomy for the weak validation problem?
- Can we decide if an action has a monoïdal extension?

Thanks!