Regular languages of thin trees

Mikołaj Bojańczyk   Tomasz Idziaszek   Michał Skrzypczak

University of Warsaw

STACS 2013, Kiel

The setting — regular languages

- Finite words — minimal det. automata, semigroups.

The setting — regular languages

- Finite words — minimal det. automata, semigroups.
- Infinite words — det. automata, Wilke algebras.

The setting — regular languages

- Finite words — minimal det. automata, semigroups.
- Infinite words — det. automata, Wilke algebras.
- Finite trees — minimal det. automata, forest algebras.

The setting — regular languages

- Finite words — minimal det. automata, semigroups.
- Infinite words — det. automata, Wilke algebras.
- Finite trees — minimal det. automata, forest algebras.
- Infinite trees — non-det. automata, (hyper)clones?

The setting — regular languages

- Finite words — minimal det. automata, semigroups.
- Infinite words — det. automata, Wilke algebras.
- Finite trees — minimal det. automata, forest algebras.
- Infinite trees — non-det. automata, (hyper)clones?

In all the cases: Monadic Second-Order logic.

The setting — regular languages

- Finite words — minimal det. automata, semigroups.
- Infinite words — det. automata, Wilke algebras.
- Finite trees — minimal det. automata, forest algebras.
- Infinite trees — non-det. automata, (hyper)clones?

In all the cases: Monadic Second-Order logic.

Motivating problem

Except decidability, little is known about regular languages of infinite trees.
No *simple* algebras for infinite trees.

Outline

- Thin trees: structures in-between words and trees.
- Thin forest algebra: Wilke algebra $\oplus$ forest algebra.
- Effective characterisations.
- Topological properties: thin trees are much poorer than all trees.

Outline

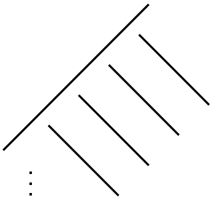
- Thin trees: structures in-between words and trees.
- Thin forest algebra: Wilke algebra $\oplus$ forest algebra.
- Effective characterisations.
- Topological properties: thin trees are much poorer than all trees.

- Finite alphabet A.
- Infinite, finitely branching, labelled trees t (leafs allowed).
- Regular languages L (those definable in MSO logic).
- Also, weak regular languages (definable in weak MSO logic).

Definition

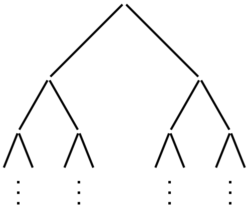
A tree is *thin* if it has only countably many infinite branches.

THIN



A *comb*

THICK



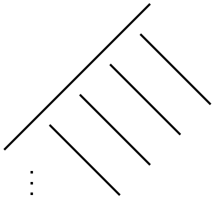
The full binary tree

Lemma (Cantor, Bendixson [1882])

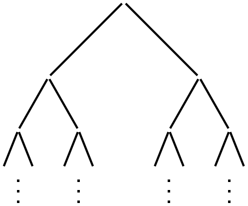
A tree is either:

- **thin** — has countably many infinite branches,
- **thick** — contains a full binary tree as a minor.

THIN



THICK



Lemma (Cantor, Bendixson [1882])

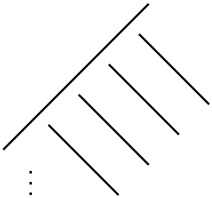
A tree is either:

- **thin** — has countably many infinite branches,
- **thick** — contains a full binary tree as a minor.

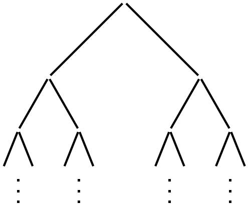
Corollary

Being a thin tree is MSO-definable.

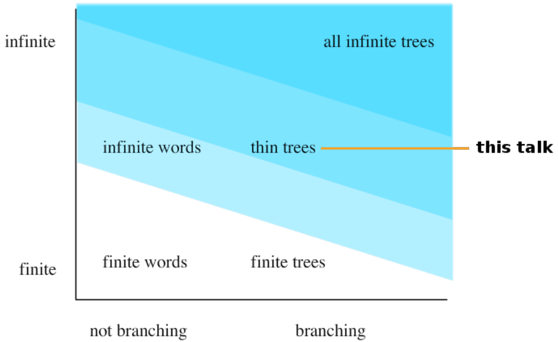
THIN



THICK

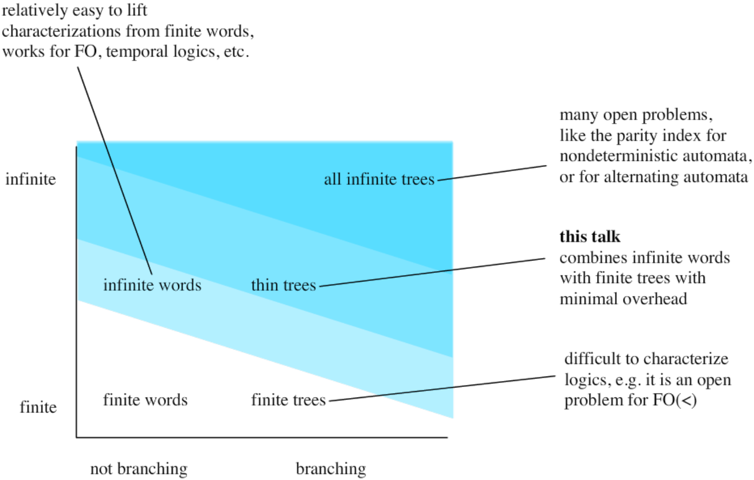


Diagram



infinite
all infinite trees
infinite words
thin trees
this talk
finite
finite words
finite trees
not branching
branching

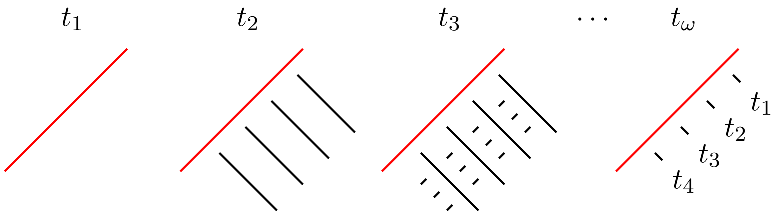
Diagram



relatively easy to lift characterizations from finite words, works for FO, temporal logics, etc.
many open problems, like the parity index for nondeterministic automata, or for alternating automata
infinite
all infinite trees
this talk
combines infinite words with finite trees with minimal overhead
infinite words
thin trees
difficult to characterize logics, e.g. it is an open problem for FO(<)
finite
finite words
finite trees
not branching
branching

Structural induction

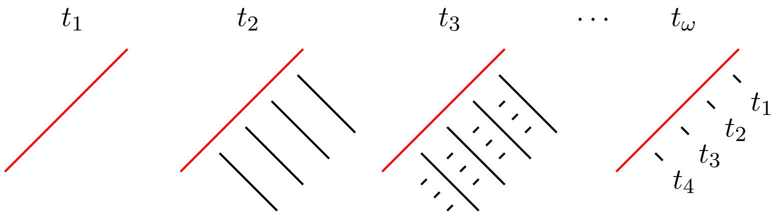
$\mathrm{rank}(t)$ — a measure of the complexity of t.



$\mathrm{rank}(t_1) = 1$ $\mathrm{rank}(t_2) = 2$ $\mathrm{rank}(t_3) = 3$ $\mathrm{rank}(t_\omega) = \omega$

Structural induction

$\mathrm{rank}(t)$ — a measure of the complexity of t.

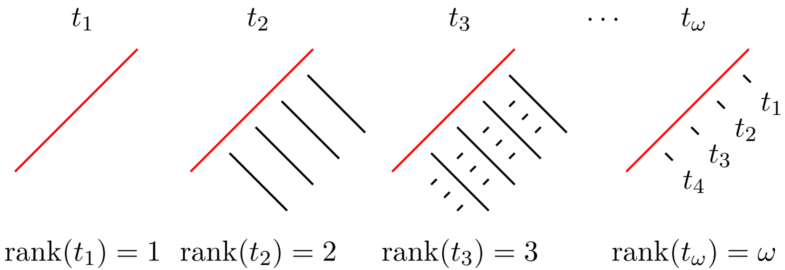


$\mathrm{rank}(t_1) = 1$ $\mathrm{rank}(t_2) = 2$ $\mathrm{rank}(t_3) = 3$ $\mathrm{rank}(t_\omega) = \omega$

A thin tree consists of a *spine* and subtrees of smaller rank along it.

Structural induction

$\mathrm{rank}(t)$ — a measure of the complexity of t.



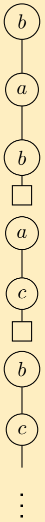
A thin tree consists of a *spine* and subtrees of smaller rank along it.

Cannot assign rank to a thick tree.
Every thin tree t has $\mathrm{rank}(t) < \omega_1$.
The spine need not be the leftmost branch!

Thin forest algebra

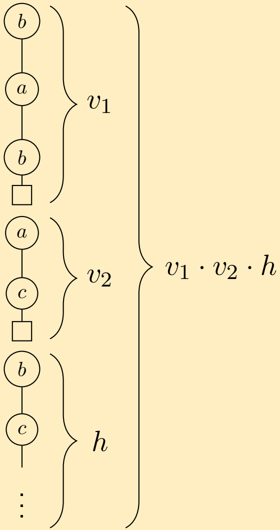
Thin forest algebra

Wilke algebra



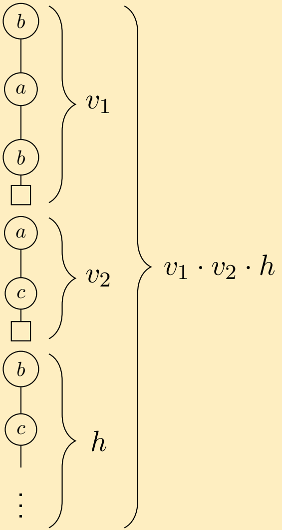
Thin forest algebra

Wilke algebra



Thin forest algebra

Wilke algebra



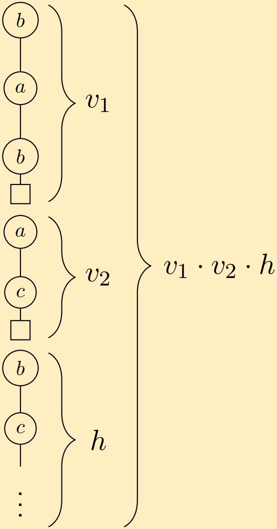
⊕

For the sake of algebra

Use forests instead of trees!

Thin forest algebra

Wilke algebra

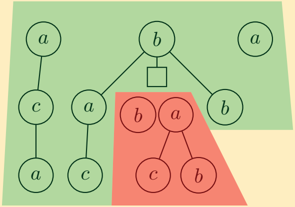


$\oplus$

For the sake of algebra

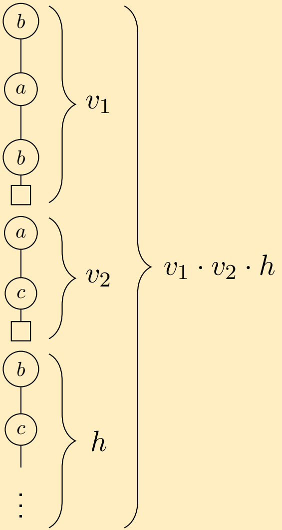
Use forests instead of trees!

Forest algebra



Thin forest algebra

Wilke algebra

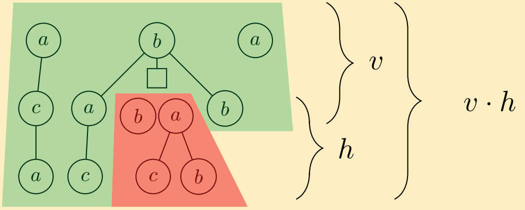


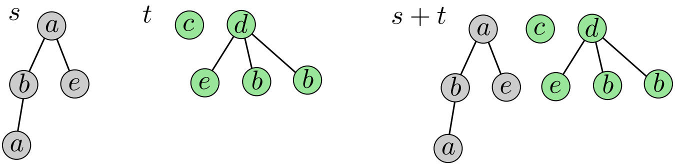
$\oplus$

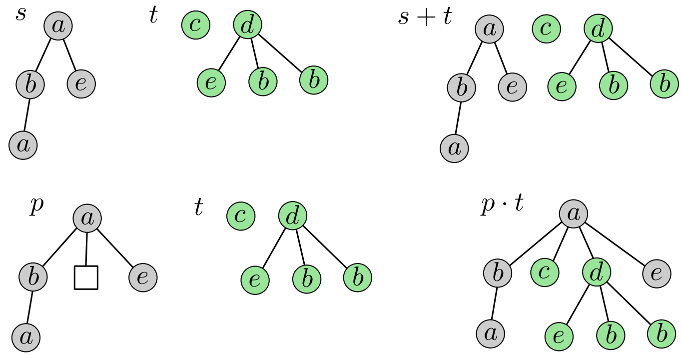
For the sake of algebra

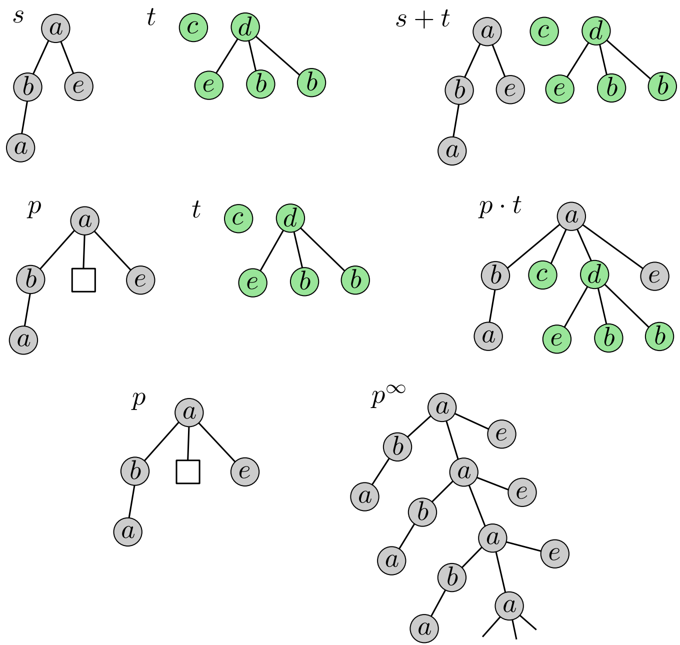
Use forests instead of trees!

Forest algebra




s
t
s + t


s
t
s + t
p
t
p · t


s
t
s + t
p
t
p · t
p
p^∞

Effective characterisations

$L = \mathrm{L}(\mathcal{A})$	$\longrightarrow$	S_L
a representation of L		the *canonical* finite algebra

Effective characterisations

$L = \mathrm{L}(\mathcal{A})$ a representation of L	$\longrightarrow$	S_L the *canonical* finite algebra
Check if L is:	$\longrightarrow$	Verify equations in S_L:

Effective characterisations

$L = \mathrm{L}(\mathcal{A})$ a representation of L	$\longrightarrow$	S_L the *canonical* finite algebra
Check if L is:	$\longrightarrow$	Verify equations in S_L:
closed under arbitrary permutations of siblings		$h + v = v + h$

Effective characterisations

$L = \mathrm{L}(\mathcal{A})$ a representation of L	$\longrightarrow$	S_L the *canonical* finite algebra
Check if L is:	$\longrightarrow$	Verify equations in S_L:
closed under arbitrary permutations of siblings		$h + v = v + h$
closed under well-founded[1] permutations of siblings		$g + h = h + g$

[1] Only finitely many changes on every path.

Effective characterisations

$L = \mathrm{L}(\mathcal{A})$ a representation of L	$\longrightarrow$	S_L the *canonical* finite algebra
Check if L is:	$\longrightarrow$	Verify equations in S_L:
closed under arbitrary permutations of siblings		$h + v = v + h$
closed under well-founded[1] permutations of siblings		$g + h = h + g$

Intermediate step

For every forests s and t check:

$$s + t \ \sim_L \ t + s \quad \text{(Myhill-Nerode style equivalence)}$$

[1] Only finitely many changes on every path.

Theorem

A regular language of thin trees is closed under bisimulational equivalence iff its syntactic algebra satisfies identities

$$h + v = v + h$$
$$h + h = h$$
$$(v^\infty + v)^\infty = v^\infty$$

Theorem

A regular language of thin trees is closed under bisimulational equivalence iff its syntactic algebra satisfies identities

$$h + v = v + h$$
$$h + h = h$$
$$(v^\infty + v)^\infty = v^\infty$$

Remark

For thick trees no such equational characterisation is known!

Theorem

The following conditions are equivalent for a regular language of thin trees L:

Theorem

The following conditions are equivalent for a regular language of thin trees L:

1. L *is weak MSO-definable among all trees,*

Theorem

The following conditions are equivalent for a regular language of thin trees L:

1. L *is weak MSO-definable among all trees,*
2. *exists* $M \in \mathbb{N}$ *such that every tree* $t \in L$ *has rank at most* M*,*

Theorem

The following conditions are equivalent for a regular language of thin trees L:

1. *L is weak MSO-definable among all trees,*
2. *exists $M \in \mathbb{N}$ such that every tree $t \in L$ has rank at most M,*
3. *L is **not** co-analytic ($\mathbf{\Pi}^1_1$)-hard among all trees,*

Theorem

The following conditions are equivalent for a regular language of thin trees L:

1. L *is weak MSO-definable among all trees,*
2. *exists* $M \in \mathbb{N}$ *such that every tree* $t \in L$ *has rank at most* M*,*
3. L *is* ***not*** *co-analytic* $(\mathbf{\Pi}^1_1)$*-hard among all trees,*
4. *the syntactic morphism for* L *satisfies condition*

$$\text{if} \quad h = v(w+h)^\infty \quad \text{or} \quad h = v(h+w)^\infty \quad \text{then} \quad h = \perp.$$

Descriptive complexity

Theorem

Every regular language of thin trees is co-analytic $(\mathbf{\Pi}^1_1)$ *among all trees.*

Descriptive complexity

Theorem

Every regular language of thin trees is co-analytic $(\mathbf{\Pi}^1_1)$ *among all trees.*

Conjecture in all trees (*a gap property*):

$$\text{Borel and regular} \Longrightarrow \text{weak MSO-definable}$$

Descriptive complexity

Theorem

Every regular language of thin trees is co-analytic $(\mathbf{\Pi}^1_1)$ *among all trees.*

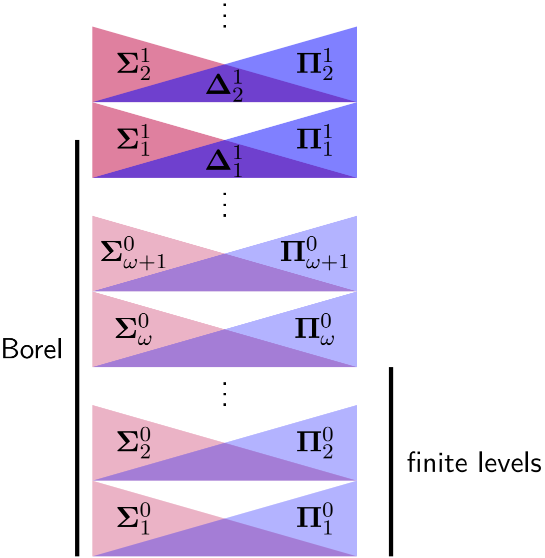
Conjecture in all trees (*a gap property*):

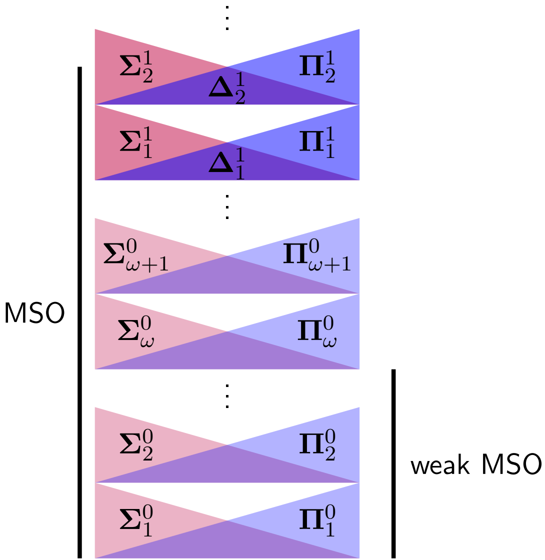
$$\text{Borel and regular} \implies \text{weak MSO-definable}$$

Corollary

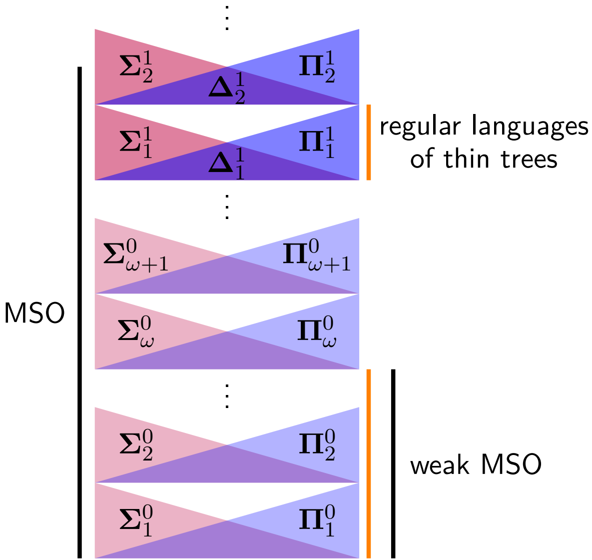
A regular language of thin trees is either:

- *definable in weak MSO among all trees,*
- $\mathbf{\Pi}^1_1$*-complete among all trees.*


Σ^1_2
Δ^1_2
Π^1_2
Σ^1_1
Δ^1_1
Π^1_1
$\Sigma^0_{\omega+1}$
$\Pi^0_{\omega+1}$
Σ^0_ω
Π^0_ω
Σ^0_2
Π^0_2
Σ^0_1
Π^0_1
Borel
finite levels



Σ^1_2
Π^1_2
Δ^1_2
Σ^1_1
Π^1_1
Δ^1_1
$\Sigma^0_{\omega+1}$
$\Pi^0_{\omega+1}$
Σ^0_ω
Π^0_ω
Σ^0_2
Π^0_2
Σ^0_1
Π^0_1
MSO
weak MSO


Σ^1_2
Π^1_2
Δ^1_2
Σ^1_1
Π^1_1
Δ^1_1
regular languages of thin trees
$\Sigma^0_{\omega+1}$
$\Pi^0_{\omega+1}$
MSO
Σ^0_ω
Π^0_ω
Σ^0_2
Π^0_2
weak MSO
Σ^0_1
Π^0_1

Summary

- Structures in-between words and trees.
- Nice (simple) algebras.
- Equational characterisations of various properties.
- Collapse of the complexity comparing to all trees.

Summary

- Structures in-between words and trees.
- Nice (simple) algebras.
- Equational characterisations of various properties.
- Collapse of the complexity comparing to all trees.

Open problems

- Decidability of the weak MSO-definability **among thin trees**?
- Is it possible to extend these techniques/results to all trees?

Effective characterisations

Effective (equational) characterisations of regular languages of thin trees that are:

- open in the standard topology,
- commutative (in two flavours),
- invariant under bisimulation (in two flavours),
- weak MSO-definable among all trees.

Descriptive complexity

Every regular language of thin trees is:

- co-analytic ($\mathbf{\Pi}^1_1$) among all trees,
- recognisable by a non-det. (1,3)-automaton among all trees,
- recognisable by an unambiguous automaton among thin trees,
- not harder than Borel sets (as a subset of thin trees).

Thank you for your attention!