Regular languages of thin trees

Mikołaj Bojańczyk Tomasz Idziaszek Michał Skrzypczak

University of Warsaw

GAMES 2012, Napoli

Motto

"Trees are harder than words"

Motto

"Trees are harder than words"

Outline

- Thin trees: structures in-between words and trees.
- Positive results: several equational characterisations.
- Negative results: thin trees are much poorer then all trees.
- Tool: thin forest algebra.

Motto

"Trees are harder than words"

Outline

- Thin trees: structures in-between words and trees.
- Positive results: several equational characterisations.
- Negative results: thin trees are much poorer then all trees.
- Tool: thin forest algebra.

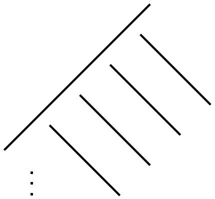
Setting

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

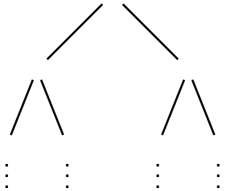
Thin trees

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

THIN



THICK



Lemma

A tree is either:

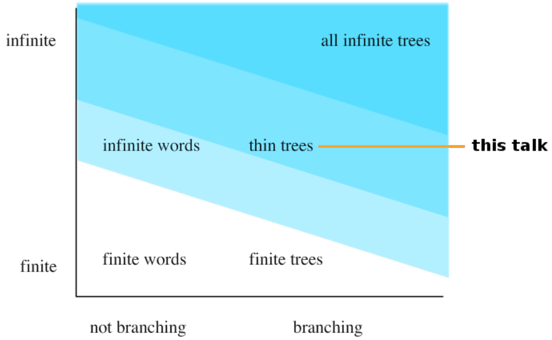
- **thin** — has countably many infinite branches,
- **thick** — contains a full binary tree as a minor.

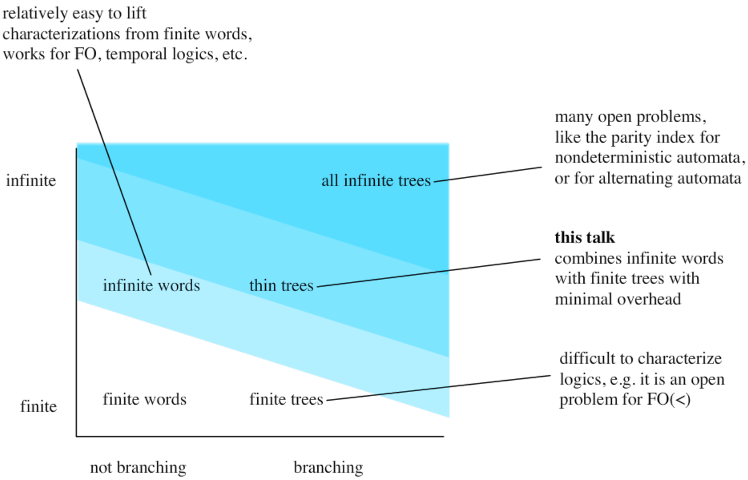
Lemma

A tree is either:

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

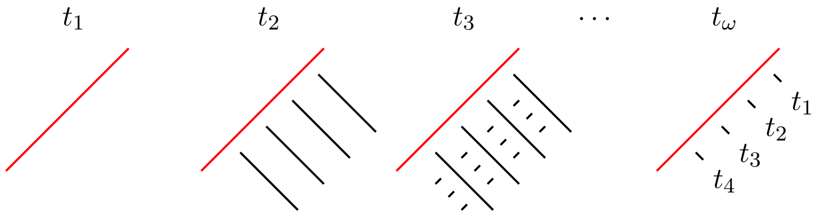
- Being a thick tree is MSO definable by an existential formula.
- Thin trees are coanalytic ($\mathbf{\Pi}^1_1$)-complete among all (thin and thick) trees.
- Being a thin tree is **not** weak-MSO definable.


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


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

Structural induction

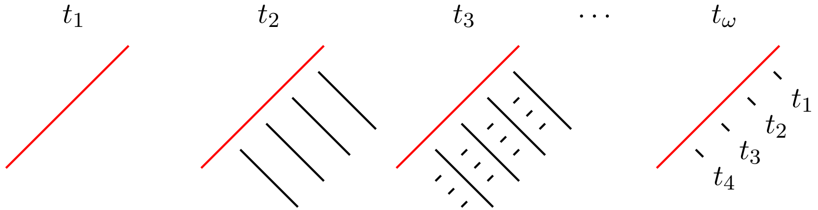
$\mathrm{rank}(t)$ — a measure of the complexity of t.



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

Structural induction

rank(t) — a measure of the complexity of t.

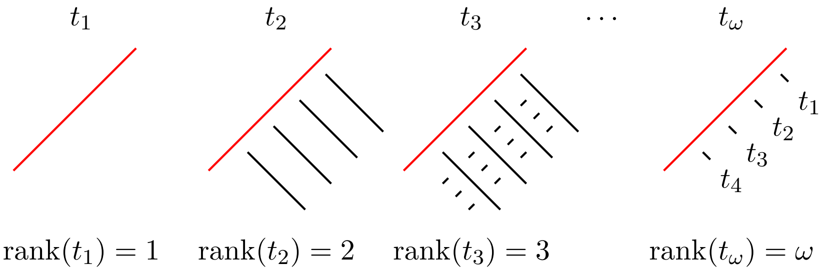


$\mathrm{rank}(t_1) = 1$ $\mathrm{rank}(t_2) = 2$ $\mathrm{rank}(t_3) = 3$ $\mathrm{rank}(t_\omega) = \omega$

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 can be arbitrarily arranged in the tree!

Thin forest algebra

Technical manoeuvre

Instead of trees we work with unranked forests.

Thin forest algebra

Technical manoeuvre

Instead of trees we work with unranked forests.

Thin forest algebra

Two-sorted algebra (H, V) where

- H contains types of forests
- V contains types of contexts
- standard operations $+$ and $\cdot$
- infinite power $V \ni v \mapsto v^{\infty} \in H$

Images to appear!

Thin forest algebra

Technical manoeuvre

Instead of trees we work with unranked forests.

Thin forest algebra

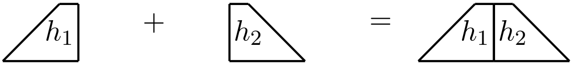
Two-sorted algebra (H, V) where

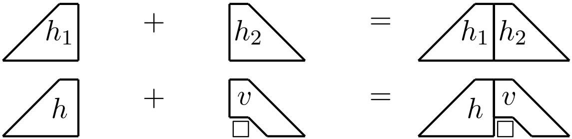
- H contains types of forests
- V contains types of contexts
- standard operations $+$ and $\cdot$
- infinite power $V \ni v \mapsto v^{\infty} \in H$

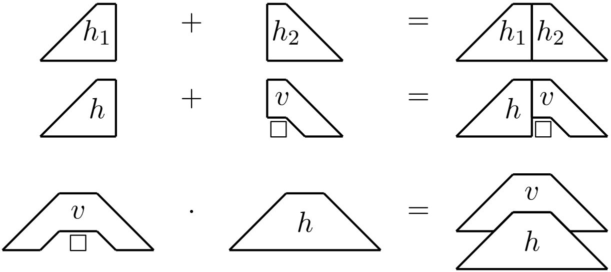
Images to appear!

Motivations

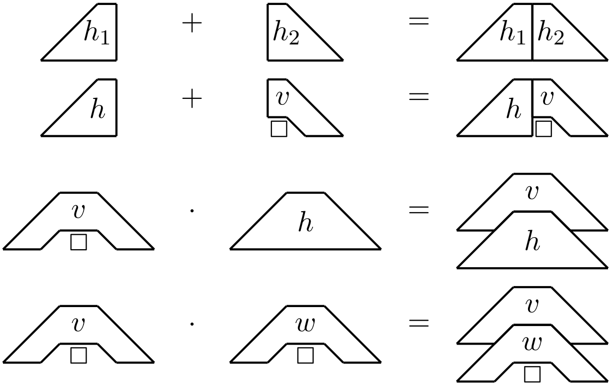
- Composition method [Shelah]
- Forest algebra [Bojańczyk, Walukiewicz]
- Wilke algebras, ω-semigroups


h_1
+
h_2
=
h_1 h_2

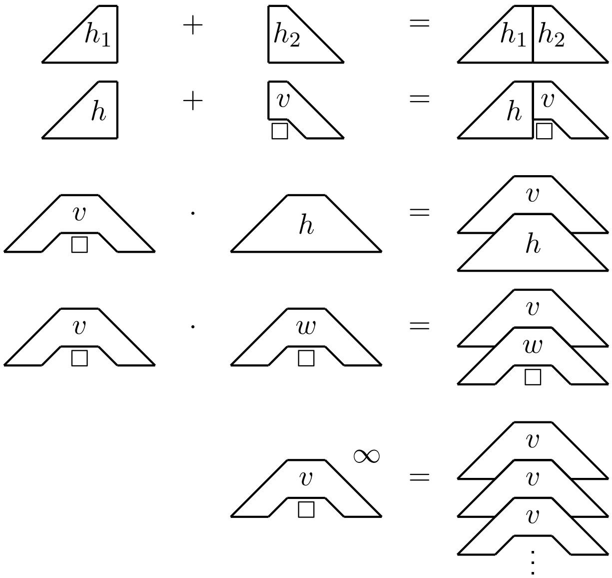

h1
+
h2
=
h1
h2
h
+
v
=
h
v



h1 + h2 = h1 h2
h + v = h v
v · h = v h



h_1 + h_2 = h_1 h_2
h + v = h v
v · h = v h
v · w = v w


h1
+
h2
=
h1
h2
h
+
v
=
h
v
v
·
h
=
v
h
v
·
w
=
v
w
v
∞
=
v
v
v

Results 1 ("positive")

Equational characterisations

Characterisation of the form:
A regular language L has property $\mathcal{P}$ if and only if the syntactic algebra $\mathcal{A}_L$ for L satisfies equations $E_{\mathcal{P}}$.

Results 1 ("positive")

Equational characterisations

Characterisation of the form:
A regular language L has property $\mathcal{P}$ if and only if the syntactic algebra $\mathcal{A}_L$ for L satisfies equations $E_{\mathcal{P}}$.

- decidability
- good understanding
- algebraic properties: varieties, quotients, ...

Theorem

A regular language of thin trees is closed under well-founded[a] *commutations iff its syntactic algebra satisfies identity*

$$h + g = g + h$$

[a] Only finitely many changes on every path.

Theorem

A regular language of thin trees is closed under well-founded[a] *commutations iff its syntactic algebra satisfies identity*

$$h + g = g + h$$

[a] Only finitely many changes on every path.

Theorem

A regular language of thin trees is closed under arbitrary commutations iff its syntactic algebra satisfies identity

$$h + v = v + h$$

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

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

Also, a similar characterisation for EF-bisimulational equivalence: equivalence induced by the bisimulation game where players can make more than one step down at once.

Theorem

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

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

Also, a similar characterisation for EF-bisimulational equivalence: equivalence induced by the bisimulation game where players can make more than one step down at once.

Remark

No such equational characterisation for all trees known!

Definition

A set of trees L is *open* if for every tree $t \in L$ there exists a depth $d \in \mathbb{N}$ such that all trees agreeing with t up to depth d belong to L.

Definition

A set of trees L is *open* if for every tree $t \in L$ there exists a depth $d \in \mathbb{N}$ such that all trees agreeing with t up to depth d belong to L.

Theorem

A regular language of thin trees is open iff its syntactic morphism α satisfies:

$$\textit{if} \quad v^{\infty} \in \alpha(L) \quad \textit{then} \quad v^{\omega} h \in \alpha(L)$$

Definition

A set of trees L is *open* if for every tree $t \in L$ there exists a depth $d \in \mathbb{N}$ such that all trees agreeing with t up to depth d belong to L.

Theorem

A regular language of thin trees is open iff its syntactic morphism α satisfies:

$$\textit{if} \quad v^\infty \in \alpha(L) \quad \textit{then} \quad v^\omega h \in \alpha(L)$$

The same condition characterises open languages of ω-words.

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** *coanalytic* ($\mathbf{\Pi}^1_1$)*-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** *coanalytic* ($\mathbf{\Pi}^1_1$)*-hard among all trees*
4. *the syntactic morphism for* L *satisfies condition*

$$\textit{if} \quad h = v(w+h)^\infty \quad \textit{or} \quad h = v(h+w)^\infty \quad \textit{then} \quad h = \bot$$

Results 2 ("negative")

Theorem

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

Results 2 ("negative")

Theorem

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

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

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

Results 2 ("negative")

Theorem

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

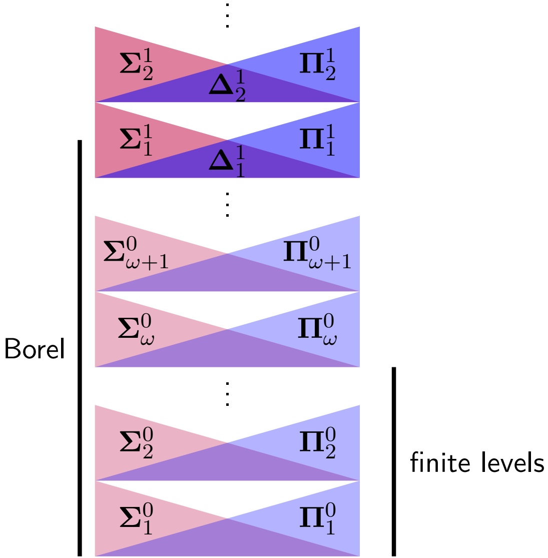
Conjecture in all trees (*a gap property*):

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

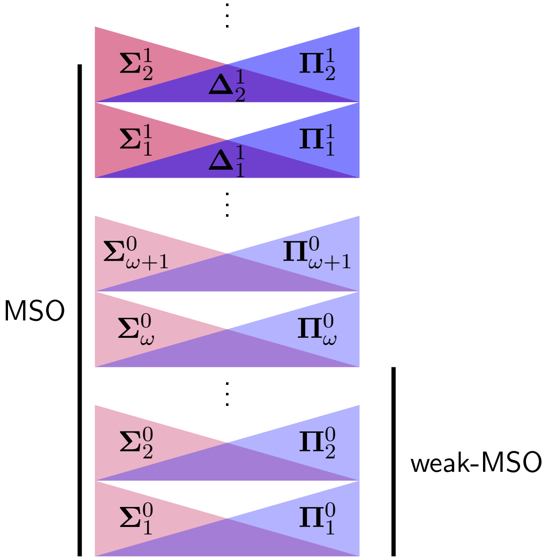
Corollary

A regular language of thin trees is either:

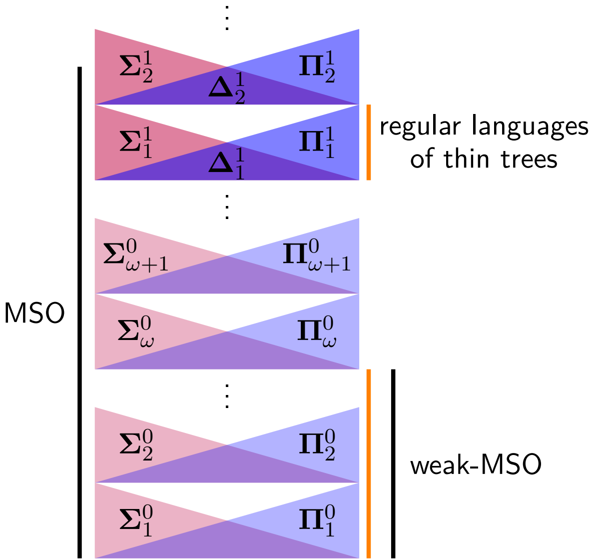
- *weak-MSO definable among all trees*
- $\mathbf{\Pi}^1_1$*-complete among all trees*



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
Borel
Σ^0_2
Π^0_2
Σ^0_1
Π^0_1
finite levels



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



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

Theorem

Every regular language of thin trees can be recognised by a nondeterministic $(1, 3)$ *parity automaton among all trees.*

Theorem

Every regular language of thin trees can be recognised by a nondeterministic $(1,3)$ parity automaton among all trees.

Theorem

Every regular language of thin trees L can be recognised by an unambiguous automaton $\mathcal{A}_L$ **among thin trees***:*
For every **thin tree** *t the automaton $\mathcal{A}_L$ has at most one accepting run on t and*

$$t \in \mathrm{L}(\mathcal{A}_L) \;\Leftrightarrow\; t \in L$$

When there are only thin trees...

Theorem

If f is a continuous function from a Polish topological space X **to thin trees** *and L is a regular language of thin trees then $f^{-1}(L)$ is Borel in X.*

Roughly speaking...
No regular language is topologically harder then Borel.

When there are only thin trees...

Theorem

If f is a continuous function from a Polish topological space X **to thin trees** *and L is a regular language of thin trees then $f^{-1}(L)$ is Borel in X.*

Roughly speaking...
No regular language is topologically harder then Borel.

Theorem

There exists a regular language of thin trees L such that every Borel subset B of a Polish topological space can be continuously reduced to L **in thin trees***: there exists a continuous function f mapping elements of X to* **thin trees** *such that $f^{-1}(L) = B$.*

Roughly speaking...
Language L is Borel-hard.

Conclusions

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

Conclusions

- 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?

Results 1 ("positive")

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

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

Results 2 ("negative")

Every regular language of thin trees is:

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

Thank you for your attention!