Positive first-order logic on words

Thomas Colcombet\textsuperscript{1}, Amina Doumane\textsuperscript{2},
\textbf{Denis Kuperberg}\textsuperscript{2}, Sam van Gool\textsuperscript{1}

\textsuperscript{1}IRIF, U. Paris
\textsuperscript{2}LIP, ENS Lyon

Highlights 2020
Upward-closed languages

**Alphabet:** \( A = 2^\Sigma \)

Letter from \( A \) = Set of atoms from \( \Sigma \).
Upward-closed languages

**Alphabet:** \( A = 2^\Sigma \)
Letter from \( A \) = Set of atoms from \( \Sigma \).

**Definition (Upward-closed languages)**

\( L \subseteq A^* \) is upward-closed if \( \forall \) words \( u, v \) and letters \( a, b \),

\[ uav \in L \text{ and } a \subseteq b \implies ubv \in L. \]

*Semantic* notion.
Upward-closed languages

Alphabet: $A = 2^\Sigma$
Letter from $A = \text{Set of atoms from } \Sigma$.

Definition (Upward-closed languages)
$L \subseteq A^*$ is upward-closed if $\forall$ words $u, v$ and letters $a, b$,

$$uav \in L \text{ and } a \subseteq b \implies ubv \in L.$$  

Semantic notion.

Example
On $\Sigma = \{1, 2\}$:
- $L_0 = A^*\{1, 2\}A^*$ is upward-closed.
- $L_1 = A^*\{1\}A^*$ is not upward-closed.
Positive first-order logic

How to syntactically define upward-closed languages?
Positive first-order logic

How to syntactically define upward-closed languages?

Definition (FO$^+$)

Positive first-order logic (FO$^+$) is FO on words with:

- no negation: all predicates appear positively.
- atomic predicate $a^\uparrow(x)$ with $a \in \Sigma$: label of $x$ contains $a$.

A language defined in FO$^+$ is upward-closed by construction.
Positive first-order logic

How to syntactically define upward-closed languages?

Definition (FO⁺)

Positive first-order logic (FO⁺) is FO on words with:

▶ no negation: all predicates appear positively.
▶ atomic predicate $a^\uparrow(x)$ with $a \in \Sigma$: label of $x$ contains $a$.

A language defined in FO⁺ is upward-closed by construction.

Does this characterize upward-closed FO-definable languages? (Syntax vs Semantics)
Our result

**Theorem (Syntax ⊊ Semantics)**

There exists an upward-closed FO language on $\Sigma = \{1, 2, 3\}$ that is not $\text{FO}^+$-definable.
Our result

Theorem (Syntax $\subsetneq$ Semantics)

There exists an upward-closed FO language on $\Sigma = \{1, 2, 3\}$ that is not $FO^+$-definable.
Background: Lyndon’s theorem

Zoom out: FO with arbitrary relational signature, on all structures.

**Theorem (Lyndon 1959)**

\(\text{FO-definable and upward-closed} \iff \text{FO}^+\text{-definable.}\)

\(\varphi\) preserved by surjective morphisms \(\iff\) equivalent to a positive formula.
Background: Lyndon’s theorem

Zoom out: FO with arbitrary relational signature, on all structures.

**Theorem (Lyndon 1959)**

$\text{FO-definable and upward-closed } \iff \text{FO}^+\text{-definable.}$

$\varphi$ preserved by surjective morphisms $\iff$ equivalent to a positive formula.

**Theorem**

Lyndon’s theorem fails on finite structures:

- on signature $(4, 3, 3, 3, 2, 1, 1)$ [Ajtai Gurevich 1987] (lattices, probabilities, number theory, topology)
- on signature $(2, 2)$ [Stolboushkin 1995] (Ehrenfeucht-Fraïssé games on grids, involved)
Background: Lyndon’s theorem

Zoom out: FO with arbitrary relational signature, on all structures.

**Theorem (Lyndon 1959)**

*FO-definable and upward-closed ⇔ FO⁺-definable.*

ϕ preserved by surjective morphisms ⇔ equivalent to a positive formula.

**Theorem**

*Lyndon’s theorem fails on finite structures:*  
- on signature \(\langle 4, 3, 3, 3, 3, 2, 1, 1 \rangle\) [Ajtai Gurevich 1987]  
  (lattices, probabilities, number theory, topology)  
- on signature \(\langle 2, 2 \rangle\) [Stolboushkin 1995]  
  (Ehrenfeucht-Fraïssé games on grids, involved)  
- on signature \(\langle 2, 1, 1, 1 \rangle\) [This work]  
  (E-F games on words, easier)
Ongoing work

Open problem

Can we decide whether a regular language is $\text{FO}^+$-definable?

Is there an algebraic characterization?
Open problem

Can we decide whether a regular language is $\text{FO}^+$-definable ?

Is there an algebraic characterization ?

Thanks for your attention !