Quantitative algebraic characterisations on *truly* infinite words.

Rémi Morvan
LaBRI, U. Bdx
https://morvan.xyz

EJCIM '22
Nice

based on joint works with

Thomas Colcombet & Sam van Gool
IRIF, U. Paris-Cité

Finite words: monoids and monadic second-order logic

Theorem [Büchi - Elgot - Trakhtenbrot ≈ '58]:

For $L \subseteq \Sigma^*$, the following are equivalent:

- L is regular;
- L is recognised by a finite monoid;
- L is described by a formula of MSO[<].

Def MSO[<] on finite words:

Signature: $\langle (a)_{a \in \Sigma}, < \rangle$

(a): unary predicate; $<$: binary predicate

Models: words $u \in \Sigma^*$

variables ⇝ positions

$a(x)$ ⇝ $u_x = a$

u_x: x-th letter of u

$x < y$ ⇝ natural order

Def Language defined by $\Phi \in \mathrm{MSO}[<]$:

$\{ u \in \Sigma^* \mid u \models \Phi \}$.

Ex: $u \models \exists x. \exists y.\ x < y \wedge a(x) \wedge b(y)$

$\iff u = |\sim\sim|\underset{x}{a}|\sim\sim\sim|\underset{y}{b}|\sim\sim|$

$\iff u \in \Sigma^* a \Sigma^* b \Sigma^*$

Def monoid $(M, \cdot, 1)$

set — neutral element

associative law

Ex:
- any group
- Σ^*
- $(\{0,1\}, \max, 0)$

Ex $\Sigma = \{a, b\}$
$L = (aa)^*$

$$f : \Sigma^* \longrightarrow M$$
$$u \longmapsto \begin{cases} \text{even} & \text{if } u \in (aa)^* \\ \text{odd} & \text{if } u \in (aa)^* a \\ \perp & \text{if } u \text{ contains a 'b'} \end{cases}$$

elements of the monoid

$\cdot$	even	odd	$\perp$
even	even	odd	$\perp$
odd	odd	even	$\perp$
$\perp$	$\perp$	$\perp$	$\perp$

Def A monoid M recognises $L \subseteq \Sigma^*$ iff there exists $f : \Sigma^* \to M$ morphism and $T \subseteq M$ st $f^{-1}[T] = L$.

$f(uv) = f(u) \cdot f(v)$

Theorem [Büchi-Elgot-Trakhtenbrot]:

For $L \subseteq \Sigma^*$, the following are equivalent:

- L is regular;
- L is recognised by a finite monoid;
- L is described by a formula of MSO[<].

Ex

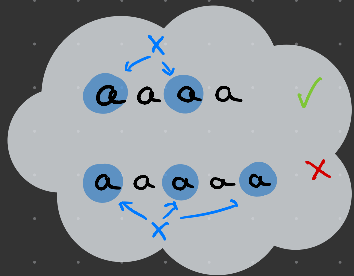
$L = (aa)^*$ on $\Sigma = \{a, b\}$.

Goal: Find $\Phi \in \mathrm{MSO}[<]$ that defines L.

1) Check that the word does not contain a 'b'

$$\forall x.\ \neg b(x)$$

2) Guess a set X of posi° $x \equiv 0 \mod 2$; is the last position odd?



$\exists X.$ the first posi° belongs to X $\wedge$ X contains every other position $\wedge$ the last posi° is not in X

First-order logic FO[<]

« la logique qu'il vous faut ! »
— Thomas C.

FO[<] $\approx$ MSO[<] with no set quantifiers.

Question: Which languages $L \subseteq \Sigma^*$ can be defined in FO[<] ?

Ex $\Sigma^* a \Sigma^* b \Sigma^*$ can be defined in FO[<]

Co-ex $(aa)^*$ cannot be defined in FO[<]

Theorem [Schützenberger '65 & McNaughton-Papert '71]
For any $L \subseteq \Sigma^*$, the following are equivalent:
- L is definable in FO[<]
- L is star-free ← not the topic of this talk
- L is recognised by a finite aperiodic monoid

Aperiodic monoids

Def A finite monoid M is aperiodic when every group $G \subseteq M$ is trivial.

Ex (Syntactic) monoid of $(aa)^*$ on $\Sigma = \{a, b\}$

$\cdot$	even	odd	$\perp$
even	even	odd	$\perp$
odd	odd	even	$\perp$
$\perp$	$\perp$	$\perp$	$\perp$

Groups: $\{\perp\}$, $\{\text{even}, \text{odd}\}$, $\{\text{even}\}$, $\{\text{odd}\}$

Deciding first-order definability

Def A morphism $f : \Sigma^* \to M$ (finite) is $FO[<]$-definable when it can be written as

$$f : \begin{array}{ccl} \Sigma^* & \longrightarrow & M \\ u & \longmapsto & \begin{cases} m_1 & \text{if } u \in L_1 \\ \vdots & \\ m_n & \text{if } u \in L_n \end{cases} \end{array}$$

where $L_1, \dots, L_n \in FO[<]$.

Ex

$$f : \begin{array}{ccl} \Sigma^* & \longrightarrow & M \\ u & \longmapsto & \begin{cases} \text{even} & \text{if } u \in (aa)^* \\ \text{odd} & \text{if } u \in (aa)^* a \\ \bot & \text{if } u \text{ contains a 'b'} \end{cases} \end{array}$$

$\cdot$	even	odd	$\bot$
even	even	odd	$\bot$
odd	odd	even	$\bot$
$\bot$	$\bot$	$\bot$	$\bot$

is not $FO[<]$-definable.

(Reformulation of) Schützenberger-McNaughton-Papert theorem:

A morphism $f : \Sigma^* \to M$ (finite) is $FO[<]$-definable iff M is aperiodic.

Qualitative vs quantitative results

In this case, could we extract all FO[<]-definable information from f?

Ex

$$f : \begin{array}{ccl} \Sigma^* & \longrightarrow & M \\ u & \longmapsto & \begin{cases} \text{even} & \text{if } u \in (aa)^* \\ \text{odd} & \text{if } u \in (aa)^* a \\ \bot & \text{if } u \text{ contains a 'b'} \end{cases} \end{array}$$

$$g : \begin{array}{ccl} \Sigma^* & \longrightarrow & \\ u & \longmapsto & \begin{cases} \{\text{even}, \text{odd}\} & \text{if } u \in (aa)^* \cup (aa)^* a = a^* \\ \{\bot\} & \text{if } u \text{ contains a 'b'} \end{cases} \end{array}$$

Def An FO[<]-approximant of $f: \Sigma^* \to M$ is a function $g: \Sigma^* \to \mathcal{P}(M)$ such that:
- g is FO[<]-definable
- $\forall u \in \Sigma^*$, $f(u) \in g(u)$

Rk $g: \begin{array}{ccc} \Sigma^* & \longrightarrow & \mathcal{P}(M) \\ u & \longmapsto & \end{array}$ is always an FO[<]-approximant

$\to$ notion of "<u>minimal</u>" approximants (not detailled here, it's technical)

<u>Thm</u> [Henckell '88] The following specification is computable.

$$\left(f: \Sigma^* \to M \right) \mapsto \left(\text{minimal } g: \Sigma^* \to \mathcal{P}(M) \text{ FO[<]-approximation} \right)$$

<u>Idea</u> g is obtained from f by "merging groups."

Finite words....

- Qualitative characterisa° of FO[<]   [Schützenberger '65 & McNaughton-Papert '71]

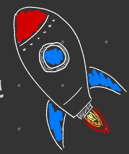
  "FO[<] $\approx$ no (non-trivial) group"

- Quantitative characterisa° of FO[<]   [Henckell '88]

  "to obtain something in FO[<], get rid of groups"

[Perrin '84] & [Place-Zeitoun '16] : extension to ω-words

have you ever heard about Ariane V ?!



Beyond finite / ω-words

Goal: understand logics (FO[<], MSO[<], ...) on complex structures.

Countable ordinals:

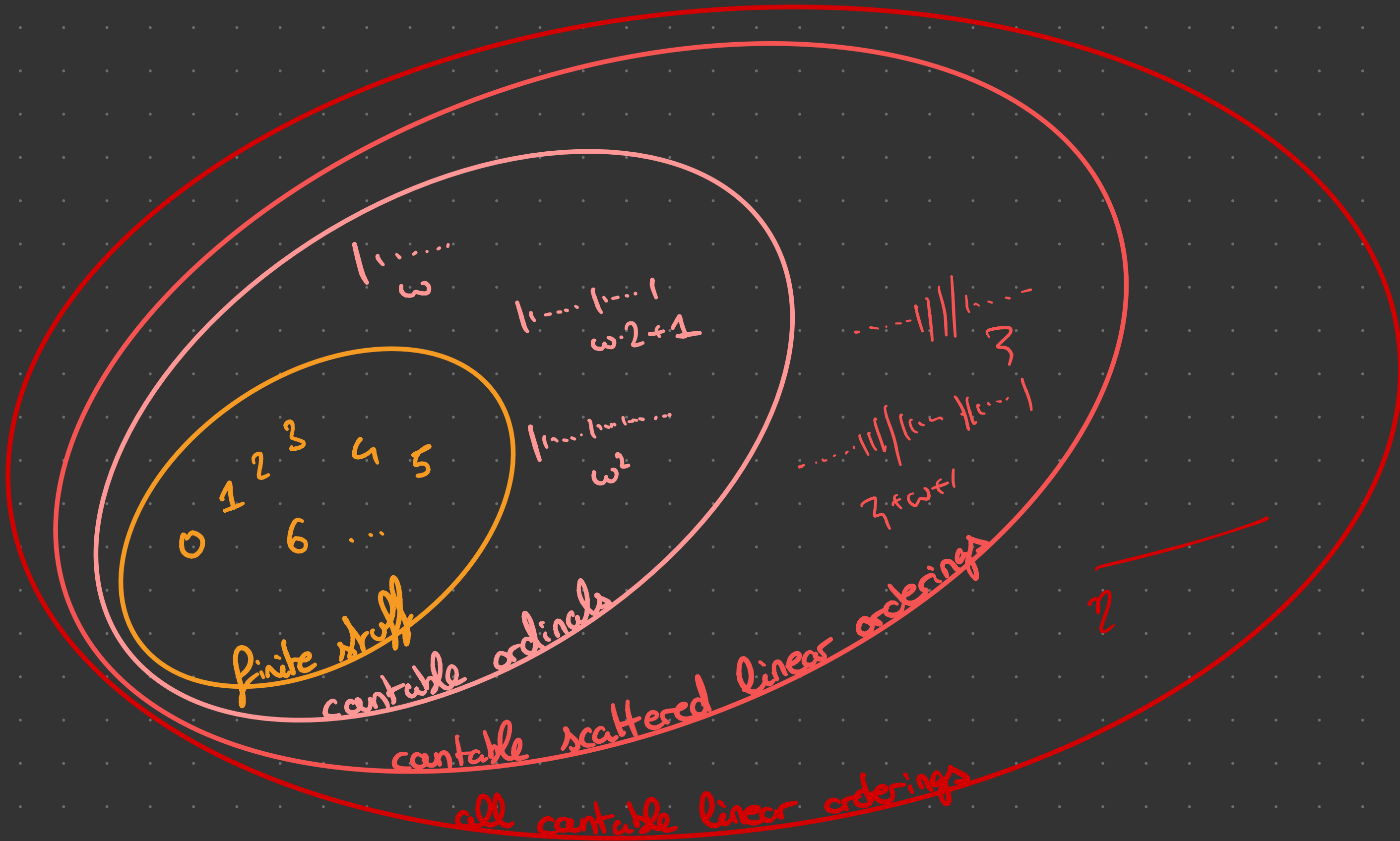
3 $\qquad$ ω $\qquad$ a baaab ab: $\omega + 2$ $\qquad$ $\omega \cdot 2$

MSO[<] / FO[<] on countable ordinal words:

"words with no last position": $\forall x.\ \exists y,\ x < y$

"every a is followed by infinitely many b's":

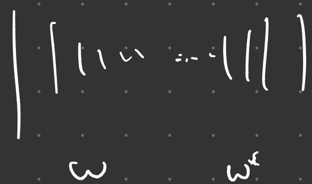
Countable linear orderings



FO[<] is not very expressive...

Can we find a formula $\Phi \in FO[<]$ defining all finite words?

i.e. $\forall u$ word over countable linear order, $u \models \Phi$ iff $u \in \Sigma^*$.



Conclusion: characterisations of FO[<]

Recall: given a morphism $f: \Sigma^* \to M$ (or f: all words over some domain $\longrightarrow$ nice algebras)
we are interested in:
- is f definable in FO[<]?
- can we compute an "optimal" FO[<]-approximant of f?

This is always decidable!

Domain	Characterisation: forbidden patterns	Qualitative	Quantitative
Finite	no groups	[Schützenberger '65 & McNaughton-Papert '71]	[Henckell '88]
Omega	no groups	[Perrin '84]	[Place-Zeitoun '16]
Cnt. ordinals	no groups	[Bedon '01]	[Colcombet-van Gool-Morvan '22]
Cnt scattered words	no groups, no gaps	[Bès-Carton '11]	[Colcombet-Morvan (unpublished)]
Cnt. linear orderings	no groups, no gaps, no shuffle	[Colcombet-Sreejith '15]	ongoing work...