Extensions of ω-REG

Edon Kelmendi

Jointly with: Mikołaj Bojańczyk Rafał Stefański Georg Zetzsche

Monadic Second Order Theory of $(\omega, >)$

$$\exists X \ \forall y \ \exists x \qquad x \in X \ \wedge \ x > y$$

Monadic Second Order Theory of $(\omega, >)$

quantify over *sets* of positions

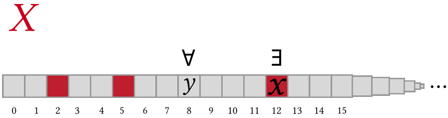
membership

order

$$\exists X \; \forall y \; \exists x \quad x \in X \;\wedge\; x > y$$

quantify over positions

Monadic Second Order Theory of $(\omega, >)$



quantify over *sets* of positions

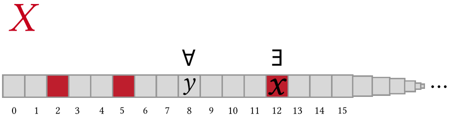
membership

order

$$\exists X \quad \forall y \ \ \exists x \quad x \in X \ \wedge \ x > y$$

quantify over positions

Monadic Second Order Theory of $(\omega, >)$



there is always a red position to the right

quantify over *sets* of positions

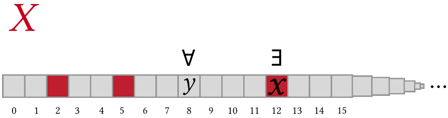
membership

order

$$\exists X \ \forall y \ \exists x \ \ x \in X \wedge x > y$$

quantify over positions

Monadic Second Order Theory of $(\omega, >)$



there is always a red position to the right

$$\exists X \ \forall y \ \exists x \quad \underbrace{x \in X \ \wedge \ x > y}$$

X has infinite cardinality

Büchi's theorem

Theorem (J. Richard Büchi, 1962)

MSO *theory of* $(\omega, >)$ *is decidable.*

Büchi's theorem

Theorem (J. Richard Büchi, 1962)

MSO *theory of* $(\omega, >)$ *is decidable.*

Are there more expressive logics that are decidable?

Are there decidable extensions?

- (R. M. Robinson, 1958) MSO extended with $f(n) = 2n$ is undecidable.

  Considered even before decidability of weak MSO by Büchi, Elgot, Trakhtenbrot, 1960, 1961.

Are there decidable extensions?

- (R. M. Robinson, 1958) MSO extended with $f(n) = 2n$ is undecidable.
  Considered even before decidability of weak MSO by Büchi, Elgot, Trakhtenbrot, 1960, 1961.

$\underbrace{\text{(C. Elgot, M. Rabin, 1966), (D. Siefkes, 1971), (W. Thomas, 1975), …}}$

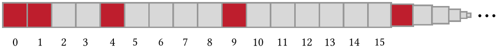
"for most natural examples of functions or binary relations
the corresponding monadic second order theory is undecidable"

Extending MSO

1. Add a function $f : \mathbb{N} \to \mathbb{N}$,
2. Add a single unary predicate (*i.e.* set) $W \subseteq \mathbb{N}$
3. Add a quantifier $Q(X) \,.\, \Phi(X)$
4. Add a language $L \subset \Sigma^\omega$

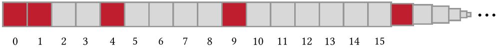
Adding unary predicates (sets)

$$W := \{n^2 \;:\; n \in \mathbb{N}\}$$



Adding unary predicates (sets)

$$W := \{n^2 \ : \ n \in \mathbb{N}\}$$



Problem

INPUT: *Non-deterministic Büchi automaton* $\mathcal{A}$

OUTPUT: *Does* $\mathcal{A}$ *accept* W*?*

Adding unary predicates (sets)

Problem

INPUT: *Non-deterministic Büchi automaton* $\mathcal{A}$
OUTPUT: *Does* $\mathcal{A}$ *accept* W*?*

decidable (C. Elgot, M. Rabin, 1966)

Adding unary predicates (sets)

Problem

INPUT: *Non-deterministic Büchi automaton $\mathcal{A}$*
OUTPUT: *Does $\mathcal{A}$ accept W?*

decidable (C. Elgot, M. Rabin, 1966)

There is a computable $C \in \mathbb{N}$ such that:

$$a^{n_1} b a^{n_2} b a^{n_3} b a^{n_4} b \cdots \text{ is accepted by } \mathcal{A}$$

iff

$$a^{n_1 \bmod C} b a^{n_2 \bmod C} b a^{n_3 \bmod C} b a^{n_4 \bmod C} b \cdots \text{ is accepted by } \mathcal{A}$$

Adding unary predicates (sets)

Problem

INPUT: *Non-deterministic Büchi automaton $\mathcal{A}$*
OUTPUT: *Does $\mathcal{A}$ accept W?*

decidable (C. Elgot, M. Rabin, 1966)

There is a computable $C \in \mathbb{N}$ such that:

$$a^{n_1} b a^{n_2} b a^{n_3} b a^{n_4} b \cdots \text{ is accepted by } \mathcal{A}$$
$$\text{iff}$$
$$a^{n_1 \bmod C} b a^{n_2 \bmod C} b a^{n_3 \bmod C} b a^{n_4 \bmod C} b \cdots \text{ is accepted by } \mathcal{A}$$

(in case of W) the latter is of the form

$$v u^{\omega}$$

Adding unary predicates (sets)

squares, cubes, etc., powers of two, powers of three, etc., factorial

Thue-Morse word,all almost-periodic words (Muchnik, Semenov, Ushakov, 2003)

(A. Semenov 1984) and (Rabinovich, Thomas, 2006)
Characterisations of W with decidable MSO theory.

Cannot always be easily applied

Adding unary predicates (sets)

$$W := \{n \; : \; n \text{ is prime}\}$$

Consider the MSO formula

$$\exists \text{ infinite } V \subset W \;\; \forall x \;\; x \in V \Rightarrow (x+2) \in W$$

Adding unary predicates (sets)

$$W := \{n \;:\; n \text{ is prime}\}$$

Consider the MSO formula

$$\exists \text{ infinite } V \subset W \;\; \forall x \;\; x \in V \Rightarrow (x+2) \in W$$

twin prime conjecture

Adding a quantifier

Express asymptotic properties (more than just "*a* infinitely often")

- (Michalewski, Mio, 2015)
  A quantifier saying:
  "the formula holds for sets with full measure" undecidable
- (Mio, Skrzypczak, Michalewski, 2017)
  A quantifier related to Baire category $\subseteq$ MSO

MSO+U

(M. Bojańczyk 2004)

$$U\ X \qquad \Phi(X)$$

formula Φ holds for arbitrary large sets X

$\forall n \in \mathbb{N} \ \ \exists X \ \ \Phi(X)$ and $|X| \geq n$.

MSO+U

(M. Bojańczyk 2004)

$$U\,X \qquad \Phi(X)$$

formula Φ holds for arbitrary large sets X

$\forall n \in \mathbb{N} \;\; \exists X \;\; \Phi(X)$ and $|X| \geq n$.

- weak MSO+U is decidable (M. Bojańczyk, 2011)
- but the full logic is not (M. Bojańczyk, P. Parys, S. Toruńczyk, 2016)

Adding a language

$L \subseteq \{a, b, c\}^\omega$

$w = abcacbabcbacabc \cdots$

Adding a language

$L \subseteq \{a, b, c\}^\omega$

$w = abcacbabcbacabc \cdots$

X_a X_b X_c

MSO+L adds a second order predicate L

$$L(X_a, X_b, X_c) \quad \Leftrightarrow \quad w \in L$$

Adding a language (Examples)

- $\{(a^n b^n c)^\omega \ : \ n \in \mathbb{N}\}$,
- $\{uv^\omega \ : \ u, v \in \Sigma^*\}$,
- $\{w \ :$ distance between consec. b's is unbounded$\}$ $\equiv_{\text{MSO+U}}$

$$\cdots\, b \overbrace{aaa \cdots aaa}^{\text{unbounded}} b \cdots$$

- Main Theorem
- Corollaries
- Proof

Main theorem

Theorem

For any non-regular L with a neutral letter, the theory of MSO$+L$ *is undecidable.*

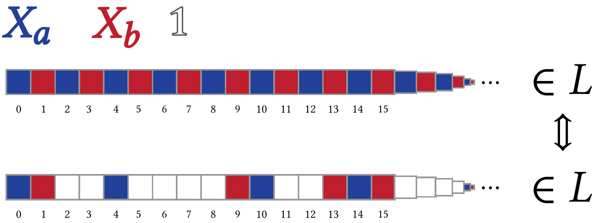
Theorem

For any non-regular L with a neutral letter, the theory of MSO+L *is undecidable.*

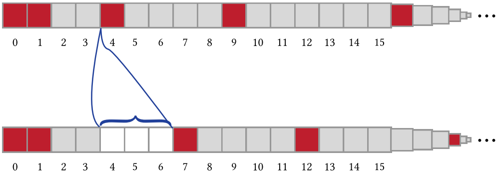
The letter $\mathbf{1} \in \Sigma$ is *neutral* if

$$w_1 \mathbf{1} w_2 \mathbf{1} \cdots \in L \quad \Leftrightarrow \quad w_1 w_2 \cdots \in L$$

for any $w_1, w_2, \ldots \in \Sigma^*$.



Squares



results in an undecidable logic

Corollaries

A class of languages $\mathcal{L}$ is a cone (or full-trio) if it is closed under:

- images under homomorphisms,
- inverse images under homomorphisms, and
- intersections with regular languages.

Examples: regular, context-free, recursively enumerable languages

Examples of *faithful cones*: context-sensitive, recursive languages

Corollaries

A class of languages $\mathcal{L}$ is a cone (or full-trio) if it is closed under:

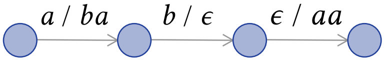
- images under homomorphisms,
- inverse images under homomorphisms, and
- intersections with regular languages.

Examples: regular, context-free, recursively enumerable languages
Examples of *faithful cones*: context-sensitive, recursive languages

equivalently (Nivat's theorem)

$\mathcal{L}$ is a cone if it is closed under:

- *transductions* (non-deterministic Büchi automaton with output)



Corollaries

$\mathcal{L}$ is a cone if it is closed under:

- *transductions* (Büchi automaton with output)

Corollary

Any Boolean-closed *cone $\mathcal{L}$, that contains a non-regular language, also contains the whole arithmetic hierarchy.*
I.e. *for any $L \subseteq \Sigma^*$ in the arithmetic hierarchy*

$$\{uv^\omega \;:\; u \in \Sigma^*, v \in L\} \in \mathcal{L}.$$

For languages over finite words: (Zetzsche, Kuske, Lohrey, 2017).

For any $\underbrace{\text{Boolean-closed}}_{\text{logic}}$ $\underbrace{\textit{cone}}_{\text{robust}}$ $\mathcal{L}$ either

- $\mathcal{L} = \omega$-REG, or
- $\mathcal{L}$ contains the whole $\underbrace{\text{arithmetic hierarchy}}_{\text{complicated}}$

For any $\underbrace{\text{Boolean-closed}}_{\text{logic}}\ \underbrace{\textit{cone}}_{\text{robust}}\ \mathcal{L}$ either

- $\mathcal{L} = \overbrace{\omega\text{-REG}}^{\textit{special}}$, or
- $\mathcal{L}$ contains the whole $\underbrace{\text{arithmetic hierarchy}}_{\text{complicated}}$

Proof

Fix L a non-regular language.

Recall:

$$U = \{w \in \{a,b\}^\omega \ : \ \text{distance between consecutive } b\text{'s is unbounded}\}$$

It suffices to show that:

$$U \qquad \in \qquad \textsc{mso} + L$$

Proof

Fix L a non-regular language.

Recall:

$$U = \{w \in \{a,b\}^\omega \;:\; \text{distance between consecutive } b\text{'s is unbounded}\}$$

It suffices to show that:

$$U \qquad \in \qquad \textsc{mso} + L$$

How to express unboundedness of distances between b's from the non-regularity of L?

An Observation

Theorem

A language $K \subseteq \Sigma^\omega$ *is* ω*-regular if and only if there exists* $\sim \subseteq \Sigma^* \times \Sigma^*$ *that is*

- *an equivalence relation with finite index,*

such that for all sequences of finite words u_i, u'_i*:*

$$\left(\bigwedge_{i \in \{1,2\}} u_i \sim u'_i \right) \quad \Rightarrow \quad u_1 u_2 \sim u'_1 u'_2$$

$$\left(\bigwedge_{i \in \mathbb{N}} u_i \sim u'_i \right) \quad \Rightarrow \quad (u_1 u_2 \cdots \in K \Leftrightarrow u'_1 u'_2 \cdots \in K)$$

An Observation

Theorem

A language $K \subseteq \Sigma^\omega$ *is* ω*-regular if and only if there exists* $\sim \subseteq \Sigma^* \times \Sigma^*$ *that is*

- *an equivalence relation with finite index,*

such that for all sequences of finite words u_i, u_i'*:*

$$\left(\bigwedge_{i \in \{1,2\}} u_i \sim u_i'\right) \quad \Rightarrow \quad u_1 u_2 \sim u_1' u_2'$$

$$\left(\bigwedge_{i \in \mathbb{N}} u_i \sim u_i'\right) \quad \Rightarrow \quad \left(u_1 u_2 \cdots \in K \;\Leftrightarrow\; u_1' u_2' \cdots \in K \right)$$

$$\overbrace{u_0}^{\in\Sigma^*} \quad \overbrace{u_1}^{\in\Sigma^*} \quad \overbrace{u_2}^{\in\Sigma^*} \quad \overbrace{u_3}^{\in\Sigma^*} \quad \cdots$$

An Observation

Theorem

A language $K \subseteq \Sigma^\omega$ *is* ω*-regular if and only if there exists* $\sim \subseteq \Sigma^* \times \Sigma^*$ *that is*

- *an equivalence relation with finite index,*

such that for all sequences of finite words u_i, u_i'*:*

$$\left(\bigwedge_{i \in \{1,2\}} u_i \sim u_i'\right) \quad \Rightarrow \quad u_1 u_2 \sim u_1' u_2'$$

$$\left(\bigwedge_{i \in \mathbb{N}} u_i \sim u_i'\right) \quad \Rightarrow \quad \left(u_1 u_2 \cdots \in K \;\Leftrightarrow\; u_1' u_2' \cdots \in K \right)$$

$$\begin{array}{cccccc}
\overbrace{u_0}^{\in\Sigma^*} & \overbrace{u_1}^{\in\Sigma^*} & \overbrace{u_2}^{\in\Sigma^*} & \overbrace{u_3}^{\in\Sigma^*} & \cdots & \in K \\
\wr & \wr & \wr & \wr & & \Updownarrow \\
\underbrace{u_0'}_{\in\Sigma^{\leq 5}} & \underbrace{u_1'}_{\in\Sigma^{\leq 5}} & \underbrace{u_2'}_{\in\Sigma^{\leq 5}} & \underbrace{u_3'}_{\in\Sigma^{\leq 5}} & \cdots & \in K
\end{array}$$

An Observation

Theorem

If K is not *ω-regular, then there is* no *equivalence relation $\sim$ with finite index such that for all sequences of finite words u_i, u_i':*

$$\left(\bigwedge_{i \in \mathbb{N}} u_i \sim u_i' \right) \quad \Rightarrow \quad \left(u_1 u_2 \cdots \in K \;\Leftrightarrow\; u_1' u_2' \cdots \in K \right)$$

$$\begin{array}{cccccc}
\overbrace{u_0}^{\in\Sigma^*} & \overbrace{u_1}^{\in\Sigma^*} & \overbrace{u_2}^{\in\Sigma^*} & \overbrace{u_3}^{\in\Sigma^*} & \cdots & \in K \\
\wr & \wr & \wr & \wr & & \Updownarrow \\
\underbrace{u_0'}_{\in\Sigma^{n_0}} & \underbrace{u_1'}_{\in\Sigma^{n_1}} & \underbrace{u_2'}_{\in\Sigma^{n_2}} & \underbrace{u_3'}_{\in\Sigma^{n_3}} & \cdots & \in K
\end{array}$$

$n_0, n_1, n_2, n_3 \cdots$ unbounded

Congruence Game

A *congruence game* on $u \in \{a, b\}^\omega$ is played between Spoiler and Duplicator

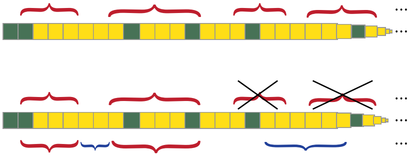
(1) Spoiler chooses an infinite family $\mathcal{W}$ of pairwise disjoint intervals

(1) Spoiler chooses an infinite family $\mathcal{W}$ of pairwise disjoint intervals

(2) Duplicator chooses intervals

$$W_1 < V_1 < W_2 < V_2 < \cdots$$

such that $W_1, W_2, \ldots$ are from $\mathcal{W}$ and $V_1, V_2, \ldots$ contain only positions with label a in the word u

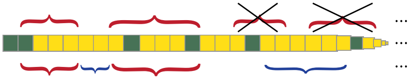


(1) Spoiler chooses an infinite family $\mathcal{W}$ of pairwise disjoint intervals

(2) Duplicator chooses intervals

$$W_1 < V_1 < W_2 < V_2 < \cdots$$

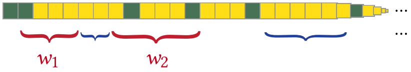
such that $W_1, W_2, \ldots$ are from $\mathcal{W}$ and $V_1, V_2, \ldots$ contain only positions with label a in the word u



(3) Spoiler chooses words

$$w_1, w_2, \ldots \in \Sigma^*$$

such that $|w_i| < |W_i|$



(1) Spoiler chooses an infinite family $\mathcal{W}$
of pairwise disjoint intervals

(2) Duplicator chooses intervals

$$W_1 < V_1 < W_2 < V_2 < \cdots$$

such that $W_1, W_2, \ldots$ are from $\mathcal{W}$ and $V_1, V_2, \ldots$
contain only positions with label a in the word u

(3) Spoiler chooses words

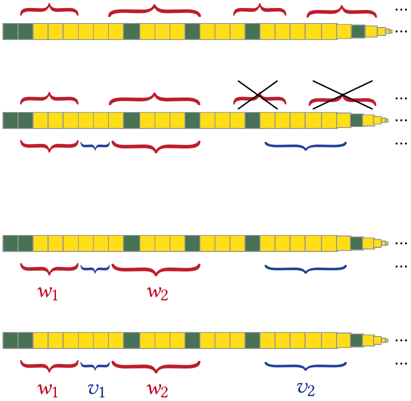
$w_1, w_2, \ldots \in \Sigma^*$

such that $|w_i| < |W_i|$

(4) Duplicator chooses words

$v_1, v_2, \ldots \in \Sigma^*$

such that $|v_i| < |V_i|$



(1) Spoiler chooses an infinite family $\mathcal{W}$ of pairwise disjoint intervals

(2) Duplicator chooses intervals

$$W_1 < V_1 < W_2 < V_2 < \cdots$$

such that $W_1, W_2, \ldots$ are from $\mathcal{W}$ and $V_1, V_2, \ldots$ contain only positions with label a in the word u

(3) Spoiler chooses words

$$w_1, w_2, \ldots \in \Sigma^*$$

such that $|w_i| < |W_i|$

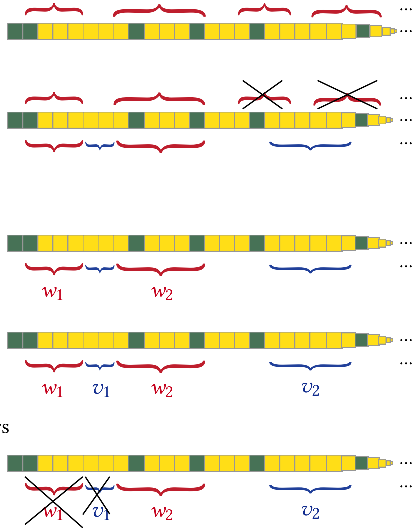
(4) Duplicator chooses words

$$v_1, v_2, \ldots \in \Sigma^*$$

such that $|v_i| < |V_i|$

(5) Spoiler chooses a sequence of natural numbers

$$i_1 < i_2 < \cdots$$



(1) Spoiler chooses an infinite family $\mathcal{W}$ of pairwise disjoint intervals

(2) Duplicator chooses intervals

$$W_1 < V_1 < W_2 < V_2 < \cdots$$

such that $W_1, W_2, \ldots$ are from $\mathcal{W}$ and $V_1, V_2, \ldots$ contain only positions with label a in the word u

(3) Spoiler chooses words

$$w_1, w_2, \ldots \in \Sigma^*$$

such that $|w_i| < |W_i|$

(4) Duplicator chooses words

$$v_1, v_2, \ldots \in \Sigma^*$$

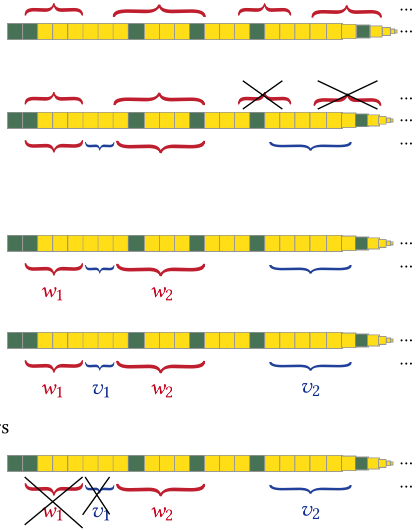
such that $|v_i| < |V_i|$

(5) Spoiler chooses a sequence of natural numbers

$$i_1 < i_2 < \cdots$$

(6) Duplicator wins the game if and only if

$$w_{i_1} w_{i_2} \cdots \in L \qquad \Longleftrightarrow \qquad v_{i_1} v_{i_2} \cdots \in L$$



Congruence Game

Theorem

Duplicator wins the congruence game for $u \;\Leftrightarrow\; u \in U$.

$u \in U \Rightarrow$ Duplicator wins the congruence game for u

(1) Spoiler chooses an infinite family $\mathcal{W}$ of pairwise disjoint intervals

(2) Duplicator chooses intervals

$$W_1 < V_1 < W_2 < V_2 < \cdots$$

such that $W_1, W_2, \ldots$ are from $\mathcal{W}$ and $V_1, V_2, \ldots$ contain only positions with label a in the word u

Since $u \in U$ we can choose the intervals such that $|V_i| > |W_i|$ for all i

(3) Spoiler chooses words

$$w_1, w_2, \ldots \in \Sigma^*$$

such that $|w_i| < |W_i|$

(4) Duplicator chooses words

$$v_1, v_2, \ldots \in \Sigma^*$$

such that $|v_i| < |V_i|$

We can choose $v_i = w_i$ for all i

(5) Spoiler chooses a sequence of natural numbers

$$i_1 < i_2 < \cdots$$

(6) Duplicator wins the game if and only if

$$w_{i_1} w_{i_2} \cdots \in L \quad \Longleftrightarrow \quad v_{i_1} v_{i_2} \cdots \in L$$

Duplicator wins because $w_{i_1} w_{i_2} \cdots = v_{i_1} v_{i_2} \cdots$

$u \notin U \Rightarrow$ Spoiler wins the congruence game for u

(1) Spoiler chooses an infinite family $\mathcal{W}$ of pairwise disjoint intervals

(2) Duplicator chooses intervals

$$W_1 < V_1 < W_2 < V_2 < \cdots$$

such that $W_1, W_2, \ldots$ are from $\mathcal{W}$ and $V_1, V_2, \ldots$ contain only positions with label a in the word u

(3) Spoiler chooses words

$$w_1, w_2, \ldots \in \Sigma^*$$

such that $|w_i| < |W_i|$

(4) Duplicator chooses words

$$v_1, v_2, \ldots \in \Sigma^*$$

such that $|v_i| < |V_i|$

(5) Spoiler chooses a sequence of natural numbers

$$i_1 < i_2 < \cdots$$

(6) Duplicator wins the game if and only if

$$w_{i_1} w_{i_2} \cdots \in L \quad \Longleftrightarrow \quad v_{i_1} v_{i_2} \cdots \in L$$

such that the lengths of intervals tend to infinity

by choice in (1), $\liminf |W_i| = \infty$

since $u \notin U$, $\limsup |V_i| < \infty$

every finite word appears infinitely often in $w_1, w_2, \ldots$

Duplicator constructs an equivalence relation $\sim$ with finite index

Theorem

If K is not ω-regular, then there is no equivalence relation $\sim$ with finite index such that for all sequences of finite words u_i, u'_i:

$$\left(\bigwedge_{i \in \mathbb{N}} u_i \sim u'_i \right) \Rightarrow \left(u_1 u_2 \cdots \in K \Leftrightarrow u'_1 u'_2 \cdots \in K \right)$$

gives a choice for step (5) and Spoiler wins

We proved that

$$U = \{w \in \{a, b\}^\omega \;\;:\;\; \text{distance between consecutive } b\text{'s is unbounded}\}$$

is the set of arenas where Duplicator wins.

Theorem

If L is not ω-regular and has a neutral letter then MSO$+L$ *is undecidable.*

Proof.

- Suffices to show that:

$$\{u \;:\; \text{Duplicator wins the congruence game for } u\}$$

  is expressible in MSO $+L$.

Theorem

If L is not ω-regular and has a neutral letter then MSO+L is undecidable.

Proof.

- Suffices to show that:

$$\{u \;:\; \text{Duplicator wins the congruence game for } u\}$$

  is expressible in MSO $+L$.
- A family of intervals can be represented by two sets of positions:

X - the leftmost positions in intervals

Y - the rightmost positions in intervals

Theorem

If L is not ω-regular and has a neutral letter then $\textsc{mso}+L$ is undecidable.

Proof.

- Suffices to show that:

$$\{u \ : \ \text{Duplicator wins the congruence game for } u\}$$

  is expressible in $\textsc{mso} + L$.

- A family of intervals can be represented by two sets of positions:

$$X \text{ - the leftmost positions in intervals}$$
$$Y \text{ - the rightmost positions in intervals}$$

- For round (1) use $\forall$, for round (2) use $\exists$
- For rounds (3) and (4) color intervals by $\Sigma \setminus \{\mathbf{1}\}$ and everything else by $\mathbf{1}$

Theorem

If L is not ω-regular and has a neutral letter then MSO+L is undecidable.

Proof.

- Suffices to show that:

$$\{u \;:\; \text{Duplicator wins the congruence game for } u\}$$

  is expressible in MSO $+L$.
- A family of intervals can be represented by two sets of positions:

$$X \text{ - the leftmost positions in intervals}$$
$$Y \text{ - the rightmost positions in intervals}$$

- For round (1) use $\forall$, for round (2) use $\exists$
- For rounds (3) and (4) color intervals by $\Sigma \setminus \{\mathbf{1}\}$ and everything else by $\mathbf{1}$
- In round (5) quantify over subsets of intervals, and

Theorem

If L is not ω-regular and has a neutral letter then MSO$+L$ *is undecidable.*

Proof.

- Suffices to show that:

$$\{u \;:\; \text{Duplicator wins the congruence game for } u\}$$

  is expressible in MSO $+L$.
- A family of intervals can be represented by two sets of positions:

$$X \text{ - the leftmost positions in intervals}$$
$$Y \text{ - the rightmost positions in intervals}$$

- For round (1) use $\forall$, for round (2) use $\exists$
- For rounds (3) and (4) color intervals by $\Sigma \setminus \{\mathbf{1}\}$ and everything else by $\mathbf{1}$
- In round (5) quantify over subsets of intervals, and
- The winning condition in round (6) is checked by the predicate L.

□