Concatenation Hierarchies of Star-Free Languages

Jana Bartoňová

Department of Mathematics and Statistics, Faculty of Science,
Masaryk University

Outline

- Regular languages
- Star-free languages
- Characterizations of classes of regular languages
- Concatenation hierarchies of star-free languages
    - in terms of languages
    - in terms of finite monoids
    - in terms of logic on words
- Straubing–Thérien hierarchy

- A – a non-empty finite set – *alphabet*

Basic Definitions

- A – a non-empty finite set – *alphabet*
- *monoid* $(M, \cdot, 1)$ – semigroup $(M, \cdot)$ with a neutral element 1

Basic Definitions

- A – a non-empty finite set – *alphabet*
- *monoid* $(M, \cdot, 1)$ – semigroup $(M, \cdot)$ with a neutral element 1
- $(A^*, \cdot, \epsilon)$ – free monoid generated by A

Basic Definitions

- A – a non-empty finite set – *alphabet*
- *monoid* $(M, \cdot, 1)$ – semigroup $(M, \cdot)$ with a neutral element 1
- $(A^*, \cdot, \epsilon)$ – free monoid generated by A
  - A^* – a set of *words* over A

Basic Definitions

- A – a non-empty finite set – *alphabet*
- *monoid* $(M, \cdot, 1)$ – semigroup $(M, \cdot)$ with a neutral element 1
- $(A^*, \cdot, \epsilon)$ – free monoid generated by A
  - A^* – a set of *words* over A
  - $\cdot$ – an operation of *concatenation*

Basic Definitions

- A – a non-empty finite set – *alphabet*
- *monoid* $(M, \cdot, 1)$ – semigroup $(M, \cdot)$ with a neutral element 1
- $(A^*, \cdot, \epsilon)$ – free monoid generated by A
  - A^* – a set of *words* over A
  - $\cdot$ – an operation of *concatenation*
  - ϵ – *empty word*

Basic Definitions

- A – a non-empty finite set – *alphabet*

- *monoid* $(M, \cdot, 1)$ – semigroup $(M, \cdot)$ with a neutral element 1
- $(A^*, \cdot, \epsilon)$ – free monoid generated by A
  - A^* – a set of *words* over A
  - $\cdot$ – an operation of *concatenation*
  - ϵ – *empty word*

- $L \subseteq A^*$ – *language* over A

Basic Definitions

- A – a non-empty finite set – *alphabet*

- *monoid* $(M, \cdot, 1)$ – semigroup $(M, \cdot)$ with a neutral element 1
- $(A^*, \cdot, \epsilon)$ – free monoid generated by A
  - A^* – a set of *words* over A
  - $\cdot$ – an operation of *concatenation*
  - ϵ – *empty word*

- $L \subseteq A^*$ – *language* over A

- *concatenation* of languages: $K \cdot L = \{u \cdot v \mid u \in K, v \in L\}$

Basic Definitions

- A – a non-empty finite set – *alphabet*
- *monoid* $(M, \cdot, 1)$ – semigroup $(M, \cdot)$ with a neutral element 1
- $(A^*, \cdot, \epsilon)$ – free monoid generated by A
  - A^* – a set of *words* over A
  - $\cdot$ – an operation of *concatenation*
  - ϵ – *empty word*
- $L \subseteq A^*$ – *language* over A
- *concatenation* of languages: $K \cdot L = \{u \cdot v \mid u \in K, v \in L\}$
- *iteration* of a language: $L^* = \cup_{n \in \mathbb{N}} L^n$ – monoid generated by L

Regular Languages

Definition

A *regular language* over an alphabet A is a language which can be constructed from languages of the form $\emptyset$ or $\{a\}$, $a \in A$, by means of finitely many applications of the operations of *union*, *concatenation* and *iteration*.

Regular Languages

Definition

A *regular language* over an alphabet A is a language which can be constructed from languages of the form $\emptyset$ or $\{a\}$, $a \in A$, by means of finitely many applications of the operations of *union*, *concatenation* and *iteration*.

regular language = *recognizable* language (recognized by a finite monoid = recognized by a finite automaton)

Definition

A *regular language* over an alphabet A is a language which can be constructed from languages of the form $\emptyset$ or $\{a\}$, $a \in A$, by means of finitely many applications of the operations of *union*, *concatenation* and *iteration*.

regular language = *recognizable* language (recognized by a finite monoid = recognized by a finite automaton)

Definition

A language $L \subseteq A^*$ is said to be *recognized* by a finite monoid M if there exists a homomorphism of monoids $\alpha \colon A^* \to M$ and a set $F \subseteq M$ such that $L = \alpha^{-1}(F)$.

Definition

A *regular language* over an alphabet A is a language which can be constructed from languages of the form $\emptyset$ or $\{a\}$, $a \in A$, by means of finitely many applications of the operations of *union*, *concatenation* and *iteration*.

regular language = *recognizable* language (recognized by a finite monoid = recognized by a finite automaton)

Definition

A language $L \subseteq A^*$ is said to be *recognized* by a finite monoid M if there exists a homomorphism of monoids $\alpha\colon A^* \to M$ and a set $F \subseteq M$ such that $L = \alpha^{-1}(F)$.

Example (important)

The set of all regular languages over A is closed under complementation: $A^* \setminus L = \alpha^{-1}(M \setminus F)$.

Variety of Regular Languages

Definition

A *class of regular languages* is a function $\mathcal{C}$ which assigns to every alphabet A a set of regular languages $\mathcal{C}(A)$ over A.

Definition

A *class of regular languages* is a function $\mathcal{C}$ which assigns to every alphabet A a set of regular languages $\mathcal{C}(A)$ over A.

Definition

A *variety of regular languages* is a class of regular languages $\mathcal{V}$ which satisfies the following conditions:

1. $L \in \mathcal{V}(A), a \in A \implies a^{-1}L, La^{-1} \in \mathcal{V}(A)$, where $a^{-1}L = \{u \in A^* \mid au \in L\}$, $La^{-1} = \{u \in A^* \mid ua \in L\}$,

Definition

A *class of regular languages* is a function $\mathcal{C}$ which assigns to every alphabet A a set of regular languages $\mathcal{C}(A)$ over A.

Definition

A *variety of regular languages* is a class of regular languages $\mathcal{V}$ which satisfies the following conditions:

1. $L \in \mathcal{V}(A), a \in A \implies a^{-1}L, La^{-1} \in \mathcal{V}(A)$, where $a^{-1}L = \{u \in A^* \mid au \in L\}$, $La^{-1} = \{u \in A^* \mid ua \in L\}$,
2. $K, L \in \mathcal{V}(A) \implies K \cup L, K \cap L, A^* \setminus L \in \mathcal{V}(A)$,

Variety of Regular Languages

Definition

A *class of regular languages* is a function $\mathcal{C}$ which assigns to every alphabet A a set of regular languages $\mathcal{C}(A)$ over A.

Definition

A *variety of regular languages* is a class of regular languages $\mathcal{V}$ which satisfies the following conditions:

1. $L \in \mathcal{V}(A), a \in A \implies a^{-1}L, La^{-1} \in \mathcal{V}(A)$, where $a^{-1}L = \{u \in A^* \mid au \in L\}$, $La^{-1} = \{u \in A^* \mid ua \in L\}$,
2. $K, L \in \mathcal{V}(A) \implies K \cup L, K \cap L, A^* \setminus L \in \mathcal{V}(A)$,
3. $L \in \mathcal{V}(A)$, $\alpha \colon B^* \to A^*$ an arbitrary homomorphism $\implies \alpha^{-1}(L) \in \mathcal{V}(B)$.

Star-Free Languages

star-free language = regular language constructed without any use of iteration and with a possible use of complementation instead

star-free language = regular language constructed without any use of iteration and with a possible use of complementation instead

Definition

A language over an alphabet A is called *star-free* if it can be constructed from languages of the form $\emptyset$ or $\{a\}$, $a \in A$, by means of finitely many applications of the operations of *union*, *concatenation* and *complementation*.

star-free language = regular language constructed without any use of iteration and with a possible use of complementation instead

Definition

A language over an alphabet A is called *star-free* if it can be constructed from languages of the form $\emptyset$ or $\{a\}$, $a \in A$, by means of finitely many applications of the operations of *union*, *concatenation* and *complementation*.

Star-free languages form a variety of regular languages.

Star-Free Languages

star-free language = regular language constructed without any use of iteration and with a possible use of complementation instead

Definition

A language over an alphabet A is called *star-free* if it can be constructed from languages of the form $\emptyset$ or $\{a\}$, $a \in A$, by means of finitely many applications of the operations of *union*, *concatenation* and *complementation*.

Star-free languages form a variety of regular languages.

How can we describe it?

Syntactic Monoid of a Regular Language and Pseudovariety of monoids

A *syntactic monoid* M_L of regular language L is the smallest monoid recognizing L.

Syntactic Monoid of a Regular Language and Pseudovariety of monoids

A *syntactic monoid* M_L of regular language L is the smallest monoid recognizing L.

- syntactic monoid can be compute algorithmically

Syntactic Monoid of a Regular Language and Pseudovariety of monoids

A *syntactic monoid* M_L of regular language L is the smallest monoid recognizing L.

- syntactic monoid can be compute algorithmically

We are looking for a class of *finite* monoids which corresponds to a given variety of regular languages.

Syntactic Monoid of a Regular Language and Pseudovariety of monoids

A *syntactic monoid* M_L of regular language L is the smallest monoid recognizing L.

- syntactic monoid can be compute algorithmically

We are looking for a class of *finite* monoids which corresponds to a given variety of regular languages.

Definition

A *pseudovariety* of monoids is a class of *finite* monoids closed under submonoids, homomorphic images and *finite* direct products.

- an analogy to a variety of monoids (in the sense of universal algebra) . . . a class of monoids closed under submonoids, homomorphic images and direct products

- pseudovariety of monoids $\mathbf{V} \mapsto \lambda(\mathbf{V})$, where $\lambda(\mathbf{V})$ is the class of regular languages such that for every alphabet A
  $\lambda(\mathbf{V})(A) = \{L \subseteq A^* \mid M_L \in \mathbf{V}\}$

- pseudovariety of monoids $\mathbf{V} \mapsto \lambda(\mathbf{V})$, where $\lambda(\mathbf{V})$ is the class of regular languages such that for every alphabet A
  $\lambda(\mathbf{V})(A) = \{L \subseteq A^* \mid M_L \in \mathbf{V}\}$
  $\ldots \lambda(\mathbf{V})$ is always a variety of regular languages

- pseudovariety of monoids $\mathbf{V} \;\mapsto\; \lambda(\mathbf{V})$, where $\lambda(\mathbf{V})$ is the class of regular languages such that for every alphabet A
  $\lambda(\mathbf{V})(A) = \{L \subseteq A^* \mid M_L \in \mathbf{V}\}$
  $\ldots \lambda(\mathbf{V})$ is always a variety of regular languages
- variety of regular languages $\mathcal{V} \;\mapsto \mu(\mathbf{V})$, where $\mu(\mathbf{V})$ is the pseudovariety of monoids generated by the class of all syntactic monoids M_L such that $L \in \mathcal{V}(A)$ for some alphabet A

- pseudovariety of monoids $\mathbf{V} \mapsto \lambda(\mathbf{V})$, where $\lambda(\mathbf{V})$ is the class of regular languages such that for every alphabet A
  $\lambda(\mathbf{V})(A) = \{L \subseteq A^* \mid M_L \in \mathbf{V}\}$
  ... $\lambda(\mathbf{V})$ is always a variety of regular languages
- variety of regular languages $\mathcal{V} \mapsto \mu(\mathbf{V})$, where $\mu(\mathbf{V})$ is the pseudovariety of monoids generated by the class of all syntactic monoids M_L such that $L \in \mathcal{V}(A)$ for some alphabet A

Theorem (Eilenberg, 1974)

The mappings λ and μ are mutually inverse isomorphisms between the lattice of all pseudovarieties of monoids and the lattice of all varieties of regular languages.

M . . . a finite monoid

Every element $m \in M$ has an unique idempotent power, denoted by m^ω. . . . $(m^\omega)^2 = m^\omega$

M ... a finite monoid
Every element $m \in M$ has an unique idempotent power, denoted by m^{ω}. $\quad \ldots (m^{\omega})^2 = m^{\omega}$

Definition

A finite monoid M is *aperiodic* if for every $m \in M$ the equality $m^{\omega} = m^{\omega+1}$ holds.

$m^{\omega+1} = m^{\omega} \cdot m$

Characterization of Star-Free Languages

M ... a finite monoid

Every element $m \in M$ has an unique idempotent power, denoted by m^ω. $\quad \dots (m^\omega)^2 = m^\omega$

Definition

A finite monoid M is *aperiodic* if for every $m \in M$ the equality $m^\omega = m^{\omega+1}$ holds.

$m^{\omega+1} = m^\omega \cdot m$

$\mathcal{W}$ ... the variety of all star-free languages

$\mathbf{W}$ ... the corresponding pseudovariety of monoids

M ... a finite monoid

Every element $m \in M$ has an unique idempotent power, denoted by m^ω. $\quad \ldots (m^\omega)^2 = m^\omega$

Definition

A finite monoid M is *aperiodic* if for every $m \in M$ the equality $m^\omega = m^{\omega+1}$ holds.

$m^{\omega+1} = m^\omega \cdot m$

$\mathcal{W}$ ... the variety of all star-free languages

$\mathbf{W}$ ... the corresponding pseudovariety of monoids

Theorem (Schützenberger, 1965)

The pseudovariety $\mathbf{W}$ *is the class of all aperiodic monoids.*

Hierarchies of Star-Free Languages

- level 0 – a variety of regular languages

Hierarchies of Star-Free Languages

- level 0 – a variety of regular languages
- level $n + \frac{1}{2}$ – *polynomial closure* of level n ($\sim$ closure under concatenation and union)

Hierarchies of Star-Free Languages

- level 0 – a variety of regular languages
- level $n + \frac{1}{2}$ – *polynomial closure* of level n
  ($\sim$ closure under concatenation and union)
- level $n + 1$ – *Boolean closure* of level $n + \frac{1}{2}$
  ($\sim$ closure under union and complementation)

- level 0 – a variety of regular languages
- level $n + \frac{1}{2}$ – *polynomial closure* of level n ($\sim$ closure under concatenation and union)
- level $n + 1$ – *Boolean closure* of level $n + \frac{1}{2}$ ($\sim$ closure under union and complementation)

Definition

A *polynomial closure* of a class of regular languages $\mathcal{C}$ is a class of regular languages $\mathrm{Pol}\,\mathcal{C}$ such that for every alphabet A
$\mathrm{Pol}\,\mathcal{C}(A)$ consists of all finite unions of languages of the form $L_0 a_1 L_1 a_2 \ldots a_n L_n$ where $L_0, \ldots L_n \in \mathcal{C}(A)$ and $a_1, \ldots, a_n \in A$.

Hierarchies of Star-Free Languages

- level 0 – a variety of regular languages
- level $n + \frac{1}{2}$ – *polynomial closure* of level n ($\sim$ closure under concatenation and union)
- level $n + 1$ – *Boolean closure* of level $n + \frac{1}{2}$ ($\sim$ closure under union and complementation)

Definition

A *polynomial closure* of a class of regular languages $\mathcal{C}$ is a class of regular languages $\operatorname{Pol}\mathcal{C}$ such that for every alphabet A
$\operatorname{Pol}\mathcal{C}(A)$ consists of all finite unions of languages of the form $L_0 a_1 L_1 a_2 \dots a_n L_n$ where $L_0, \dots L_n \in \mathcal{C}(A)$ and $a_1, \dots, a_n \in A$.

Theorem

Every integer level in the hierarchy is a variety of regular languages.

- level 0 – a variety of regular languages
- level $n + \frac{1}{2}$ – *polynomial closure* of level n ($\sim$ closure under concatenation and union)
- level $n + 1$ – *Boolean closure* of level $n + \frac{1}{2}$ ($\sim$ closure under union and complementation)

Definition

A *polynomial closure* of a class of regular languages $\mathcal{C}$ is a class of regular languages $\mathrm{Pol}\,\mathcal{C}$ such that for every alphabet A
$\mathrm{Pol}\,\mathcal{C}(A)$ consists of all finite unions of languages of the form $L_0 a_1 L_1 a_2 \ldots a_n L_n$ where $L_0, \ldots L_n \in \mathcal{C}(A)$ and $a_1, \ldots, a_n \in A$.

Theorem

Every integer level in the hierarchy is a variety of regular languages.

Half levels are not closed under complementation.

Positive Variety of Regular Languages

Definition

A *variety of regular languages* is a class of regular languages $\mathcal{V}$ which satisfies the following conditions:

1. $L \in \mathcal{V}(A), a \in A \implies a^{-1}L, La^{-1} \in \mathcal{V}(A)$, where $a^{-1}L = \{u \in A^* \mid au \in L\}$, $La^{-1} = \{u \in A^* \mid ua \in L\}$,
2. $K, L \in \mathcal{V}(A) \implies K \cup L, K \cap L, A^* \setminus L \in \mathcal{V}(A)$,
3. $L \in \mathcal{V}(A)$, $\alpha \colon B^* \to A^*$ an arbitrary homomorphism $\implies \alpha^{-1}(L) \in \mathcal{V}(B)$.

Definition

A *positive variety* of regular languages is a class of regular languages satisfying all the conditions from previous definition except the closure under complementation (in item 2).

Positive Variety of Regular Languages

Definition

A *variety of regular languages* is a class of regular languages $\mathcal{V}$ which satisfies the following conditions:

1. $L \in \mathcal{V}(A), a \in A \implies a^{-1}L, La^{-1} \in \mathcal{V}(A)$, where $a^{-1}L = \{u \in A^* \mid au \in L\}$, $La^{-1} = \{u \in A^* \mid ua \in L\}$,
2. $K, L \in \mathcal{V}(A) \implies K \cup L, K \cap L, A^* \setminus L \in \mathcal{V}(A)$,
3. $L \in \mathcal{V}(A)$, $\alpha \colon B^* \to A^*$ an arbitrary homomorphism $\implies \alpha^{-1}(L) \in \mathcal{V}(B)$.

Definition

A *positive variety* of regular languages is a class of regular languages satisfying all the conditions from previous definition except the closure under complementation (in item 2).

Positive varieties of regular languages correspond to so-called *pseudovarieties of* **ordered monoids**.

- *ordered monoid* $(M, \leq)$ – monoid M with a compatible partial order $\leq$ on M:
  $\forall m_1, m_2, n_1, n_2 \in M : \; m_1 \leq m_2 \wedge n_1 \leq n_2 \implies m_1 n_1 \leq m_2 n_2$

Pseudovariety of Ordered Monoids

- *ordered monoid* $(M, \leq)$ – monoid M with a compatible partial order $\leq$ on M:
  $\forall m_1, m_2, n_1, n_2 \in M : \; m_1 \leq m_2 \wedge n_1 \leq n_2 \implies m_1 n_1 \leq m_2 n_2$
- submonoid N of an ordered monoid $(M, \leq)$ – restriction of $\leq$ to N

Pseudovariety of Ordered Monoids

- *ordered monoid* $(M, \leq)$ – monoid M with a compatible partial order $\leq$ on M:
  $\forall m_1, m_2, n_1, n_2 \in M: \; m_1 \leq m_2 \wedge n_1 \leq n_2 \implies m_1 n_1 \leq m_2 n_2$
- submonoid N of an ordered monoid $(M, \leq)$ – restriction of $\leq$ to N
- *homomorphism of ordered monoids* $\alpha\colon M \to N$
  … a homomorphism of monoids which preserves orderings:
  $\forall m, n \in M: \; m \leq n \implies \alpha(m) \leq \alpha(n)$

Pseudovariety of Ordered Monoids

- *ordered monoid* $(M, \leq)$ – monoid M with a compatible partial order $\leq$ on M:
  $\forall m_1, m_2, n_1, n_2 \in M: \; m_1 \leq m_2 \wedge n_1 \leq n_2 \implies m_1 n_1 \leq m_2 n_2$
- submonoid N of an ordered monoid $(M, \leq)$ – restriction of $\leq$ to N
- *homomorphism of ordered monoids* $\alpha \colon M \to N$
  ... a homomorphism of monoids which preserves orderings:
  $\forall m, n \in M: \; m \leq n \implies \alpha(m) \leq \alpha(n)$
- *product of ordered monoids* $M \times N$:
  $(m_1, m_2) \leq (n_1, n_2) \iff m_1 \leq n_1 \;\wedge\; m_2 \leq n_2$

Pseudovariety of Ordered Monoids

- *ordered monoid* $(M, \leq)$ – monoid M with a compatible partial order $\leq$ on M:
  $\forall m_1, m_2, n_1, n_2 \in M: \; m_1 \leq m_2 \wedge n_1 \leq n_2 \implies m_1 n_1 \leq m_2 n_2$
- submonoid N of an ordered monoid $(M, \leq)$ – restriction of $\leq$ to N
- *homomorphism of ordered monoids* $\alpha\colon M \to N$
  ... a homomorphism of monoids which preserves orderings:
  $\forall m, n \in M: \; m \leq n \implies \alpha(m) \leq \alpha(n)$
- *product of ordered monoids* $M \times N$:
  $(m_1, m_2) \leq (n_1, n_2) \iff m_1 \leq n_1 \; \wedge \; m_2 \leq n_2$

Definition

A *pseudovariety of ordered monoids* is a class of *finite* monoids closed under submonoids, homomorphic images and finite direct products.

Syntactic Ordered Monoid

Definition

A language $L \subseteq A^*$ is said to be *recognized* by a finite *ordered* monoid M if there exists a homomorphism of monoids $\alpha\colon A^* \to M$ and a *downward closed* set $F \subseteq M$ such that $L = \alpha^{-1}(F)$.

$F \subseteq M$ is *downward closed*:
$\forall m, n \in M : m \in F \wedge n \leq m \implies n \in F$

Syntactic Ordered Monoid

Definition

A language $L \subseteq A^*$ is said to be *recognized* by a finite *ordered* monoid M if there exists a homomorphism of monoids $\alpha \colon A^* \to M$ and a *downward closed* set $F \subseteq M$ such that $L = \alpha^{-1}(F)$.

$F \subseteq M$ is *downward closed*:
$\forall m, n \in M : m \in F \wedge n \leq m \implies n \in F$

Syntactic ordered monoid of regular language L:
$(M_L, \leq_L)$, where $\leq_L$ is the largest possible partial order on M_L such that $(M_L, \leq_L)$ recognizes L.

A Correspondence with Positive Varieties

- pseudovariety of *ordered* monoids $\mathbf{V} \mapsto \lambda(\mathbf{V})$, where $\lambda(\mathbf{V})$ is the class of regular languages such that for every alphabet A $\lambda(\mathbf{V})(A) = \{L \subseteq A^* \mid (M_L, \leq_L) \in \mathbf{V}\}$

A Correspondence with Positive Varieties

- pseudovariety of *ordered* monoids $\mathbf{V} \mapsto \lambda(\mathbf{V})$, where $\lambda(\mathbf{V})$ is the class of regular languages such that for every alphabet A
  $\lambda(\mathbf{V})(A) = \{L \subseteq A^* \mid (M_L, \leq_L) \in \mathbf{V}\}$
  $\ldots \lambda(\mathbf{V})$ is always a *positive* variety of regular languages

A Correspondence with Positive Varieties

- pseudovariety of *ordered* monoids $\mathbf{V} \mapsto \lambda(\mathbf{V})$, where $\lambda(\mathbf{V})$ is the class of regular languages such that for every alphabet A
  $\lambda(\mathbf{V})(A) = \{L \subseteq A^* \mid (M_L, \leq_L) \in \mathbf{V}\}$
  $\ldots \lambda(\mathbf{V})$ is always a *positive* variety of regular languages
- *positive* variety of regular languages $\mathcal{V} \mapsto \mu(\mathbf{V})$, where $\mu(\mathbf{V})$ is the pseudovariety of *ordered* monoids generated by by the class of all syntactic *ordered* monoids $(M_L, \leq_L)$ such that $L \in \mathcal{V}(A)$ for some alphabet A

A Correspondence with Positive Varieties

- pseudovariety of *ordered* monoids $\mathbf{V} \mapsto \lambda(\mathbf{V})$, where $\lambda(\mathbf{V})$ is the class of regular languages such that for every alphabet A $\lambda(\mathbf{V})(A) = \{L \subseteq A^* \mid (M_L, \leq_L) \in \mathbf{V}\}$
  ... $\lambda(\mathbf{V})$ is always a *positive* variety of regular languages
- *positive* variety of regular languages $\mathcal{V} \mapsto \mu(\mathbf{V})$, where $\mu(\mathbf{V})$ is the pseudovariety of *ordered* monoids generated by by the class of all syntactic *ordered* monoids $(M_L, \leq_L)$ such that $L \in \mathcal{V}(A)$ for some alphabet A

Theorem (Pin, 1995)

The mappings λ and μ are mutually inverse isomorphisms between the lattice of all pseudovarieties of ordeder monoids and the lattice of all positive varieties of regular languages.

Back to Concatenation Hierarchies

- integer level of a hierarchy . . . variety of regular languages
  $\longleftrightarrow$ pseudovariety of monoids

Back to Concatenation Hierarchies

- integer level of a hierarchy ... variety of regular languages $\longleftrightarrow$ pseudovariety of monoids
- half level of a hierarchy ... positive variety of regular languages

Back to Concatenation Hierarchies

- integer level of a hierarchy ... variety of regular languages $\longleftrightarrow$ pseudovariety of monoids
- half level of a hierarchy ... positive variety of regular languages $\longleftrightarrow$ pseudovariety of ordered monoids

Back to Concatenation Hierarchies

- integer level of a hierarchy . . . variety of regular languages $\longleftrightarrow$ pseudovariety of monoids
- half level of a hierarchy . . . positive variety of regular languages $\longleftrightarrow$ pseudovariety of ordered monoids

THE MAIN PROBLEM – for a given level of a hierarchy to find an algorithm which enables to decide whether a given star-free language belongs to this level

Back to Concatenation Hierarchies

- integer level of a hierarchy . . . variety of regular languages $\longleftrightarrow$ pseudovariety of monoids
- half level of a hierarchy . . . positive variety of regular languages $\longleftrightarrow$ pseudovariety of ordered monoids

THE MAIN PROBLEM – for a given level of a hierarchy to find an algorithm which enables to decide whether a given star-free language belongs to this level . . . *decidability* of the level

- integer level of a hierarchy ... variety of regular languages $\longleftrightarrow$ pseudovariety of monoids
- half level of a hierarchy ... positive variety of regular languages $\longleftrightarrow$ pseudovariety of ordered monoids

THE MAIN PROBLEM – for a given level of a hierarchy to find an algorithm which enables to decide whether a given star-free language belongs to this level ... *decidability* of the level

- To solve a decidability of a level it suffices to prove a decidability of the corresponding pseudovariety of (ordered) monoids:

A language L belongs to a (positive) variety of regular languages if and only if its (ordered) syntactic monoid belongs to the corresponding pseudovariety of (ordered) monoids.

Another Approach: First-Order Logic on Words

FO[$<$] ... a fragment of the first-order logic

- formulas are interpreted on words – finite sets of positions $N = \{1, \ldots, n\}$ where to every position there's assigned a letter from a fixed alphabet A

Another Approach: First-Order Logic on Words

FO[$<$] ... a fragment of the first-order logic

- formulas are interpreted on words – finite sets of positions $N = \{1, \ldots, n\}$ where to every position there's assigned a letter from a fixed alphabet A
- a binary predicate symbol $<$
  – interpreted on a word as a usual ordering on natural numbers

Another Approach: First-Order Logic on Words

FO[$<$] . . . a fragment of the first-order logic

- formulas are interpreted on words – finite sets of positions $N = \{1, \ldots, n\}$ where to every position there's assigned a letter from a fixed alphabet A
- a binary predicate symbol $<$
  – interpreted on a word as a usual ordering on natural numbers
- for every letter $a \in A$ an unary predicate symbol P_a
  – interpreted on a word as a set of positions labeled by a

Another Approach: First-Order Logic on Words

FO[$<$] ... a fragment of the first-order logic

- formulas are interpreted on words – finite sets of positions $N = \{1, \dots, n\}$ where to every position there's assigned a letter from a fixed alphabet A
- a binary predicate symbol $<$
  – interpreted on a word as a usual ordering on natural numbers
- for every letter $a \in A$ an unary predicate symbol P_a
  – interpreted on a word as a set of positions labeled by a

Definition

Let φ be a closed formula of FO[$<$]. A language *defined* by φ is the language $L(\varphi)$ of all words on which φ is satisfied:
$L(\varphi) = \{u \in A^* \mid u \models \varphi\}$.

FO[$<$] ... a fragment of the first-order logic

- formulas are interpreted on words – finite sets of positions $N = \{1, \dots, n\}$ where to every position there's assigned a letter from a fixed alphabet A
- a binary predicate symbol $<$
  – interpreted on a word as a usual ordering on natural numbers
- for every letter $a \in A$ an unary predicate symbol P_a
  – interpreted on a word as a set of positions labeled by a

Definition

Let φ be a closed formula of FO[$<$]. A language *defined* by φ is the language $L(\varphi)$ of all words on which φ is satisfied:
$L(\varphi) = \{u \in A^* \mid u \models \varphi\}$.

Example

$\varphi = (\exists x\ \exists y)(\mathrm{P}_a(x)\ \wedge\ \mathrm{P}_b(y)\ \wedge\ x < y)$

FO[$<$] ... a fragment of the first-order logic

- formulas are interpreted on words – finite sets of positions $N = \{1, \ldots, n\}$ where to every position there's assigned a letter from a fixed alphabet A
- a binary predicate symbol $<$
  – interpreted on a word as a usual ordering on natural numbers
- for every letter $a \in A$ an unary predicate symbol P_a
  – interpreted on a word as a set of positions labeled by a

Definition

Let φ be a closed formula of FO[$<$]. A language *defined* by φ is the language $L(\varphi)$ of all words on which φ is satisfied:
$L(\varphi) = \{u \in A^* \mid u \models \varphi\}$.

Example

$\varphi = (\exists x \; \exists y)(\mathrm{P}_a(x) \;\wedge\; \mathrm{P}_b(y) \;\wedge\; x < y)$
$L(\varphi) = A^* a A^* b A^*$

A Connection with Star-Free Languages

Theorem (Mc Naughton, Pappert, 1971)

A language is definable in FO[$<$] *if and only if it is star-free.*

A Connection with Star-Free Languages

Theorem (Mc Naughton, Pappert, 1971)

A language is definable in FO$[<]$ *if and only if it is star-free.*

- Thomas (1982); Perrin, Pin (1986): a correspondence with levels of concatenation hierarchies

A Connection with Star-Free Languages

Theorem (Mc Naughton, Pappert, 1971)

A language is definable in FO[$<$] *if and only if it is star-free.*

- Thomas (1982); Perrin, Pin (1986): a correspondence with levels of concatenation hierarchies

How does this correspondence look like?

Step $n \longrightarrow n + \frac{1}{2}$

- Let Φ be a set of all formulas φ such that $L(\varphi)$ is a language of level n of a concatenation hierarchy (for a fixed alphabet A).

Step $n \longrightarrow n + \frac{1}{2}$

- Let Φ be a set of all formulas φ such that $L(\varphi)$ is a language of level n of a concatenation hierarchy (for a fixed alphabet A).
- Ψ ... a set of formulas consisting of all formulas created from formulas from Φ by a possible removing some quantifiers (more precisely strings of the form "$\exists x$"or "$\forall x$")

Step $n \longrightarrow n + \frac{1}{2}$

- Let Φ be a set of all formulas φ such that $L(\varphi)$ is a language of level n of a concatenation hierarchy (for a fixed alphabet A).
- Ψ . . . a set of formulas consisting of all formulas created from formulas from Φ by a possible removing some quantifiers (more precisely strings of the form "$\exists x$"or "$\forall x$")
- $\bar{\Psi}$ . . . the smallest set of formulas containing Ψ and closed under conjunction, disjunction and existential quantifier:
  - $\varphi \in \Psi \implies \varphi \in \bar{\Psi}$
  - $\varphi, \psi \in \bar{\Psi} \implies \varphi \wedge \psi, \varphi \vee \psi \in \bar{\Psi}$
  - $\varphi \in \bar{\Psi} \implies \exists x\, \varphi \in \bar{\Psi}$ for every variable x

Step $n \longrightarrow n + \frac{1}{2}$

- Let Φ be a set of all formulas φ such that $L(\varphi)$ is a language of level n of a concatenation hierarchy (for a fixed alphabet A).
- Ψ . . . a set of formulas consisting of all formulas created from formulas from Φ by a possible removing some quantifiers (more precisely strings of the form "$\exists x$"or "$\forall x$")
- $\bar{\Psi}$ . . . the smallest set of formulas containing Ψ and closed under conjunction, disjunction and existential quantifier:
  - $\varphi \in \Psi \implies \varphi \in \bar{\Psi}$
  - $\varphi, \psi \in \bar{\Psi} \implies \varphi \wedge \psi, \varphi \vee \psi \in \bar{\Psi}$
  - $\varphi \in \bar{\Psi} \implies \exists x\, \varphi \in \bar{\Psi}$ for every variable x
- Then level $n + \frac{1}{2}$ is defined by the set of formulas $\bar{\Phi}$ consisting of all closed formulas from $\bar{\Psi}$.

- Let Φ be a set of all formulas φ such that $L(\varphi)$ is a language of level $n + \frac{1}{2}$ of a concatenation hierarchy (for a fixed alphabet A).

Step $n+\frac{1}{2} \longrightarrow n+1$

- Let Φ be a set of all formulas φ such that $L(\varphi)$ is a language of level $n+\frac{1}{2}$ of a concatenation hierarchy (for a fixed alphabet A).
- Then level $n+1$ is defined by the smallest set of formulas $\bar{\Phi}$ which contains Φ and is closed under conjunction, disjunction and negation:
  - $\varphi \in \Phi \implies \varphi \in \bar{\Phi}$
  - $\varphi, \psi \in \bar{\Phi} \implies \varphi \wedge \psi, \varphi \vee \psi, \neg\varphi \in \bar{\Phi}$

Straubing–Thérien Hierarchy

- level 0 … $\mathcal{V}_0(A) = \{\emptyset, A^*\}$ for every alphabet A

- level 0 . . . $\mathcal{V}_0(A) = \{\emptyset, A^*\}$ for every alphabet A

Theorem (Brzozowski, Knast, 1977)

The Straubing–Thérien hierarchy is infinite.

Straubing–Thérien Hierarchy

- level 0 . . . $\mathcal{V}_0(A) = \{\emptyset, A^*\}$ for every alphabet A

Theorem (Brzozowski, Knast, 1977)

The Straubing–Thérien hierarchy is infinite.

- $\mathbf{V}_0$ . . . the pseudovariety of trivial monoids ($M = \{1\}$)

- level 0 ... $\mathcal{V}_0(A) = \{\emptyset, A^*\}$ for every alphabet A

Theorem (Brzozowski, Knast, 1977)

The Straubing–Thérien hierarchy is infinite.

- $\mathbf{V}_0$ ... the pseudovariety of trivial monoids ($M = \{1\}$)
- $\mathbf{V}_{\frac{1}{2}}$ ... the pseudovariety of all ordered monoids satisfying the inequality $1 \leq m$ for every $m \in M$

- level 0 … $\mathcal{V}_0(A) = \{\emptyset, A^*\}$ for every alphabet A

Theorem (Brzozowski, Knast, 1977)

The Straubing–Thérien hierarchy is infinite.

- $\mathbf{V}_0$ … the pseudovariety of trivial monoids ($M = \{1\}$)
- $\mathbf{V}_{\frac{1}{2}}$ … the pseudovariety of all ordered monoids satisfying the inequality $1 \leq m$ for every $m \in M$
- level 1 … *piecewise testable* languages
  Simon (1975): $\mathbf{V}_1$ … the pseudovariety of $\mathcal{J}$-trivial monoids

- level 0 . . . $\mathcal{V}_0(A) = \{\emptyset, A^*\}$ for every alphabet A

Theorem (Brzozowski, Knast, 1977)

The Straubing–Thérien hierarchy is infinite.

- $\mathbf{V}_0$ . . . the pseudovariety of trivial monoids ($M = \{1\}$)
- $\mathbf{V}_{\frac{1}{2}}$ . . . the pseudovariety of all ordered monoids satisfying the inequality $1 \leq m$ for every $m \in M$
- level 1 . . . *piecewise testable* languages
  Simon (1975): $\mathbf{V}_1$ . . . the pseudovariety of $\mathcal{J}$-trivial monoids
- level $\frac{3}{2}$ . . . a proof of decidability by Hashigushi, Arfi (1987); Pin, Weil (1997)

- levels 2, $\frac{5}{2}$ – a proof of decidability by Place, Zeitoun (2014)

- levels 2, $\frac{5}{2}$ – a proof of decidability by Place, Zeitoun (2014)
- level $\frac{7}{2}$ – a proof of decidability by Place (2015)

All these new results in terms of the logic on words.

Thank you for your attention.