Commutative closures of regular languages

Jul, 28. 2009

Antoine Delignat-Lavaud
Computer Science Department,
École Normale Supérieure de Cachan

Outline

1. **Trace monoids and recognizability**
   - The free partially commutative monoid
   - Recognizability in non-free monoids
2. **Varieties**
   - Relationship with recognizable languages
3. **Commutative closures**
   - Closure under P_4
4. **References**

Commutation relation

- An *independence* or *commutation* relation is a symmetric and irreflexive relation I.

Commutation relation

- An *independence* or *commutation* relation is a symmetric and irreflexive relation I.
- (Σ, I) is the *independence alphabet*

Commutation relation

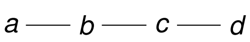
- An *independence* or *commutation* relation is a symmetric and irreflexive relation I.
- (Σ, I) is the *independence alphabet*
- Can be represented as undirected *commutation graph*.

Commutation relation

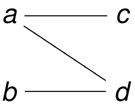
- An *independence* or *commutation* relation is a symmetric and irreflexive relation I.
- (Σ, I) is the *independence alphabet*
- Can be represented as undirected *commutation graph*.
- Complement of I is the *dependence* relation D.

Commutation relation

- An *independence* or *commutation* relation is a symmetric and irreflexive relation I.
- (Σ, I) is the *independence alphabet*
- Can be represented as undirected *commutation graph*.
- Complement of I is the *dependence* relation D.



I



D

Trace monoid

- The set of finite sequences of letters from an alphabet Σ is the free monoid Σ^*.

Trace monoid

- The set of finite sequences of letters from an alphabet Σ is the free monoid Σ^*.
- Given a commutation relation I, the commutation equivalence $\sim_I$ over Σ^* is the least congruence such that $ab \sim_I ba$ for all $(a, b) \in I$.

Trace monoid

- The set of finite sequences of letters from an alphabet Σ is the free monoid Σ^*.
- Given a commutation relation I, the commutation equivalence $\sim_I$ over Σ^* is the least congruence such that $ab \sim_I ba$ for all $(a, b) \in I$.
- The quotient $\mathbb{M}(\Sigma, I) = \Sigma^* / \sim_I$ is the *free partially I-commutative monoid* or *trace monoid*.

Trace monoid

- The set of finite sequences of letters from an alphabet Σ is the free monoid Σ^*.
- Given a commutation relation I, the commutation equivalence $\sim_I$ over Σ^* is the least congruence such that $ab \sim_I ba$ for all $(a, b) \in I$.
- The quotient $\mathbb{M}(\Sigma, I) = \Sigma^*/\sim_I$ is the *free partially I-commutative monoid* or *trace monoid*.
- Elements from $\mathbb{M}$ are called traces.

Trace monoid

- The set of finite sequences of letters from an alphabet Σ is the free monoid Σ^*.
- Given a commutation relation I, the commutation equivalence $\sim_I$ over Σ^* is the least congruence such that $ab \sim_I ba$ for all $(a, b) \in I$.
- The quotient $\mathbb{M}(\Sigma, I) = \Sigma^* / \sim_I$ is the *free partially I-commutative monoid* or *trace monoid*.
- Elements from $\mathbb{M}$ are called traces.
- ϕ denotes the canonical quotient homomorphism.

Trace monoid

- The set of finite sequences of letters from an alphabet Σ is the free monoid Σ^*.
- Given a commutation relation I, the commutation equivalence $\sim_I$ over Σ^* is the least congruence such that $ab \sim_I ba$ for all $(a, b) \in I$.
- The quotient $\mathbb{M}(\Sigma, I) = \Sigma^*/{\sim_I}$ is the *free partially I-commutative monoid* or *trace monoid*.
- Elements from $\mathbb{M}$ are called traces.
- ϕ denotes the canonical quotient homomorphism.
- $[\cdot]_I$ denotes the closure operator $\phi^{-1} \circ \phi$.

Trace monoid

- The set of finite sequences of letters from an alphabet Σ is the free monoid Σ^*.
- Given a commutation relation I, the commutation equivalence $\sim_I$ over Σ^* is the least congruence such that $ab \sim_I ba$ for all $(a, b) \in I$.
- The quotient $\mathbb{M}(\Sigma, I) = \Sigma^*/\sim_I$ is the *free partially I-commutative monoid* or *trace monoid*.
- Elements from $\mathbb{M}$ are called traces.
- ϕ denotes the canonical quotient homomorphism.
- $[\cdot]_I$ denotes the closure operator $\phi^{-1} \circ \phi$.
- If $t \in \mathbb{M}$, a word w such that $t = [w]_I$ is a linearization of t.

The free monoid case

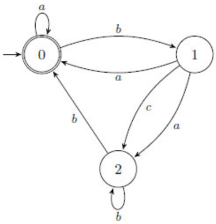
- A recognizable language $L \subseteq \Sigma^*$ is a language accepted by a (deterministic or nondeterministic) finite state machine $\mathcal{A} = (\Sigma, Q, I, \Delta, F)$.

The free monoid case

- A recognizable language $L \subseteq \Sigma^*$ is a language accepted by a (deterministic or nondeterministic) finite state machine $\mathcal{A} = (\Sigma, Q, I, \Delta, F)$.
- Kleene's theorem: the class of recognizable word languages is the closure of the class of finite languages under product, union and iteration. (the rational languages).

The free monoid case

- A recognizable language $L \subseteq \Sigma^*$ is a language accepted by a (deterministic or nondeterministic) finite state machine $\mathcal{A} = (\Sigma, Q, I, \Delta, F)$.
- Kleene's theorem: the class of recognizable word languages is the closure of the class of finite languages under product, union and iteration. (the rational languages).



$$(a^*(ba)^*(b(a+c)b^*b)^*)^*$$

Recognizability in non-free monoids

- An M-automaton is a tuple $\mathcal{A} = (Q, \delta, F)$.

Recognizability in non-free monoids

- An M-automaton is a tuple $\mathcal{A} = (Q, \delta, F)$.
- Q is a finite monoid, $F \subseteq Q$.

Recognizability in non-free monoids

- An M-automaton is a tuple $\mathcal{A} = (Q, \delta, F)$.
- Q is a finite monoid, $F \subseteq Q$.
- $\delta : M \to Q$ is an homomorphism of monoids.

Recognizability in non-free monoids

- An M-automaton is a tuple $\mathcal{A} = (Q, \delta, F)$.
- Q is a finite monoid, $F \subseteq Q$.
- $\delta : M \rightarrow Q$ is an homomorphism of monoids.
- The subset of M recognized by $\mathcal{A}$ is $\delta^{-1}(F)$.

Recognizability in non-free monoids

- An M-automaton is a tuple $\mathcal{A} = (Q, \delta, F)$.
- Q is a finite monoid, $F \subseteq Q$.
- $\delta : M \to Q$ is an homomorphism of monoids.
- The subset of M recognized by $\mathcal{A}$ is $\delta^{-1}(F)$.

Recognizable $\neq$ rational

$(\Sigma, I) = a$—b, $L = (ab)^*$.
$\phi^{-1}(L) = \{w \in \{a, b\}^* \mid |w|_a = |w|_b\}$

Syntactic monoid

Let M be a monoid and $T \subseteq M$. The syntactic preorder over T, $\leq_T$ is defined by:

$$x \leq_T y \text{ if } \forall u, v \in M, uyv \in T \Rightarrow uxv \in T$$

The syntactic equivalence $\equiv_T$ is defined by

$$x \equiv_T y \Leftrightarrow x \leq_T y \wedge y \leq_T x$$

The syntactic monoid of T is the quotient $M_T = M/\equiv_T$.

Syntactic monoid

Let M be a monoid and $T \subseteq M$. The syntactic preorder over T, $\leq_T$ is defined by:

$$x \leq_T y \text{ if } \forall u, v \in M, uyv \in T \Rightarrow uxv \in T$$

The syntactic equivalence $\equiv_T$ is defined by

$$x \equiv_T y \Leftrightarrow x \leq_T y \wedge y \leq_T x$$

The syntactic monoid of T is the quotient $M_T = M/\equiv_T$.

Characterization of recognizable languages

A subset $T \subseteq M$ is recognizable if and only if, its syntactic monoid M_T is finite.

Varieties of monoids and recognizable languages

Variety of monoids

A class of monoids closed under taking submonoids, left and right quotients and direct products.
A *pseudovariety of monoids* is a variety of finite monoids.

Example

The class of all finite monoids is a pseudovariety.
The class of finite groups is a pseudovariety denoted **G**.

Variety of languages

A *variety of languages* $\mathcal{V}(\Sigma)$ over some alphabet Σ is a family of recognizable languages such that for all Σ,

- $\mathcal{V}(\Sigma)$ is a boolean algebra

Variety of languages

A *variety of languages* $\mathcal{V}(\Sigma)$ over some alphabet Σ is a family of recognizable languages such that for all Σ,

- $\mathcal{V}(\Sigma)$ is a boolean algebra
- $\mathcal{V}(\Sigma)$ is closed under inverse of morphisms.

Variety of languages

A *variety of languages* $\mathcal{V}(\Sigma)$ over some alphabet Σ is a family of recognizable languages such that for all Σ,

- $\mathcal{V}(\Sigma)$ is a boolean algebra
- $\mathcal{V}(\Sigma)$ is closed under inverse of morphisms.
- $\mathcal{V}(\Sigma)$ is closed under residuals $a^{-1}L, La^{-1}, a \in \Sigma, L \in \mathcal{V}(\Sigma)$.

Variety of languages

A *variety of languages* $\mathcal{V}(\Sigma)$ over some alphabet Σ is a family of recognizable languages such that for all Σ,

- $\mathcal{V}(\Sigma)$ is a boolean algebra
- $\mathcal{V}(\Sigma)$ is closed under inverse of morphisms.
- $\mathcal{V}(\Sigma)$ is closed under residuals $a^{-1}L, La^{-1}, a \in \Sigma, L \in \mathcal{V}(\Sigma)$.

Variety theorem

The correspondance $\mathbf{V} \to \mathcal{V}$ that maps to a variety of monoids $\mathbf{V}$ the variety of languages whose syntactic monoid is in $\mathbf{V}$ is one to one.

Polynomial closure

The polynomial closure $Pol(L)$ of a class of language $\mathcal{L}$ over Σ is the set finite unions of $L_0 a_1 L_1 a_2 \cdots a_n L_n$ whith $a_i \in \Sigma$ and $L_i \in \mathcal{L}$.

Polynomial closure

The polynomial closure $Pol(L)$ of a class of language $\mathcal{L}$ over Σ is the set finite unions of $L_0 a_1 L_1 a_2 \cdots a_n L_n$ whith $a_i \in \Sigma$ and $L_i \in \mathcal{L}$.

Polynomials of group languages

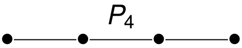
- $Pol(\mathcal{G})$ (the polynomial closure of the variety associated to **G**) is closed unter total commutation.
- It is closed under I-commutation if (Σ, I) does not have one of the two following induced subgraphs:

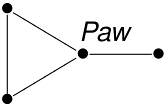
Polynomial closure

The polynomial closure $Pol(L)$ of a class of language $\mathcal{L}$ over Σ is the set finite unions of $L_0 a_1 L_1 a_2 \cdots a_n L_n$ whith $a_i \in \Sigma$ and $L_i \in \mathcal{L}$.

Polynomials of group languages

- $Pol(\mathcal{G})$ (the polynomial closure of the variety associated to **G**) is closed unter total commutation.
- It is closed under I-commutation if (Σ, I) does not have one of the two following induced subgraphs:





P_4-closure of a group language

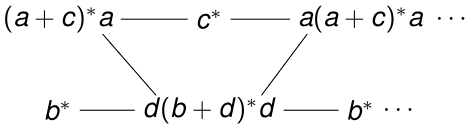
During the internship, we proved the following result:

The P_4-closure of a group language is recognizable.

P_4-closure of a group language

During the internship, we proved the following result:

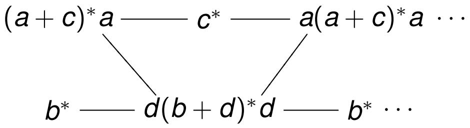
The P_4-closure of a group language is recognizable.



P_4-closure of a group language

During the internship, we proved the following result:

The P_4-closure of a group language is recognizable.



Proof

- Proof relies on structure of traces.
- Recognizability is shown using MSO logic
- See the internship report for more details.

References

Gómez, Antonio Cano and Guaiana, Giovanna and Pin, Jean-Éric
When Does Partial Commutative Closure Preserve Regularity?
ICALP'2008