Regularity Conditions for Iterated Shuffle on Commutative Regular Languages

Stefan Hoffmann, Universität Trier, Germany

International Conference on Implementation and Application of Automata (CIAA) 2021,
July 19 – July 22, 2021 (virtuell)
Organizers: Sebastian Maneth, Peter Leupold, Kathryn Lorenz, Martin Vu

July, 2021

Notation

Let Σ be some finite alphabet.
Languages: subsets of Σ^* (the free monoid).
ε: the empty word.

Definition (Commutative language)

A language $L \subseteq \Sigma^*$ is called commutative if it is closed under arbitrary permutation of words.

Example 0.1

$U = \{aab, aba, baa\}$ is commutative, $V = \{ab\}$ is not, as $ba \notin V$.

The Parikh map $\psi : \Sigma^* \to \mathbb{N}_0^{|\Sigma|}$ is

$$\psi(w) = (|w|_{a_1}, \ldots, |w|_{a_k}),$$

where $|w|_{a_i}$ denotes the number of occurrences of the letter a_i in w.

Notation

Let Σ be some finite alphabet.
Languages: subsets of Σ^* (the free monoid).
ε: the empty word.

Definition (Commutative language)

A language $L \subseteq \Sigma^*$ is called commutative if it is closed under arbitrary permutation of words.

Example 0.1

$U = \{aab, aba, baa\}$ is commutative, $V = \{ab\}$ is not, as $ba \notin V$.

The Parikh map $\psi : \Sigma^* \to \mathbb{N}_0^{|\Sigma|}$ is

$$\psi(w) = (|w|_{a_1}, \ldots, |w|_{a_k}),$$

where $|w|_{a_i}$ denotes the number of occurrences of the letter a_i in w.

Notation

Let Σ be some finite alphabet.
Languages: subsets of Σ^* (the free monoid).
ε: the empty word.

Definition (Commutative language)

A language $L \subseteq \Sigma^*$ is called commutative if it is closed under arbitrary permutation of words.

Example 0.1

$U = \{aab, aba, baa\}$ is commutative, $V = \{ab\}$ is not, as $ba \notin V$.

The Parikh map $\psi : \Sigma^* \to \mathbb{N}_0^{|\Sigma|}$ is

$$\psi(w) = (|w|_{a_1}, \ldots, |w|_{a_k}),$$

where $|w|_{a_i}$ denotes the number of occurrences of the letter a_i in w.

Shuffle Operation

Definition (Shuffle operation)

The shuffle operation, denoted by $\sqcup\!\sqcup$, is defined by

$$u \sqcup\!\sqcup v := \left\{ x_1 y_1 x_2 y_2 \cdots x_n y_n \;\middle|\; \begin{array}{l} u = x_1 x_2 \cdots x_n,\, v = y_1 y_2 \cdots y_n, \\ x_i, y_i \in \Sigma^*, 1 \leq i \leq n, n \geq 1 \end{array} \right\},$$

for $u, v \in \Sigma^*$ and $L_1 \sqcup\!\sqcup L_2 := \bigcup_{x \in L_1, y \in L_2} (x \sqcup\!\sqcup y)$ for $L_1, L_2 \subseteq \Sigma^*$.

$\{ab\} \sqcup\!\sqcup \{cd\} = \{abcd, acbd, acdb, cadb, cdab, cabd\}$

Shuffle Operation

Definition (Shuffle operation)

The shuffle operation, denoted by $\sqcup\!\sqcup$, is defined by

$$u \sqcup\!\sqcup v := \left\{ x_1 y_1 x_2 y_2 \cdots x_n y_n \;\middle|\; \begin{array}{l} u = x_1 x_2 \cdots x_n, v = y_1 y_2 \cdots y_n, \\ x_i, y_i \in \Sigma^*, 1 \le i \le n, n \ge 1 \end{array} \right\},$$

for $u, v \in \Sigma^*$ and $L_1 \sqcup\!\sqcup L_2 := \bigcup_{x \in L_1, y \in L_2} (x \sqcup\!\sqcup y)$ for $L_1, L_2 \subseteq \Sigma^*$.

$\{ab\} \sqcup\!\sqcup \{cd\} = \{abcd, acbd, acdb, cadb, cdab, cabd\}$

Iterated Shuffle

1. The Iterated Shuffle is $L^{\sqcup\!\sqcup,*} = \bigcup_{i \geq 0} L^{i,\sqcup\!\sqcup}$ where $L^{0,\sqcup\!\sqcup} = \{\varepsilon\}$ and $L^{i+1,\sqcup\!\sqcup} = L^{i,\sqcup\!\sqcup} \sqcup\!\sqcup L$.
2. The regular languages are not closed under iterated shuffle in general.
3. Example: $\{ab, ba\}^{\sqcup\!\sqcup,*} = \{$words with equal number of a's and b's$\}$.

For what languages is the iterated shuffle regularlity-preserving?

In my previous talk (Monday) I addressed this issue and mentioned that for group languages, this holds true (actually, more general, for polynomials of group languages).

Here, we investigate this question for the subclass of commutative regular languages.

Iterated Shuffle

1. The Iterated Shuffle is $L^{\sqcup\!\sqcup,*} = \bigcup_{i\geq 0} L^{i,\sqcup\!\sqcup}$ where $L^{0,\sqcup\!\sqcup} = \{\varepsilon\}$ and $L^{i+1,\sqcup\!\sqcup} = L^{i,\sqcup\!\sqcup} \sqcup\!\sqcup L$.
2. The regular languages are not closed under iterated shuffle in general.
3. Example: $\{ab, ba\}^{\sqcup\!\sqcup,*} = \{$words with equal number of a's and b's$\}$.

For what languages is the iterated shuffle regularlity-preserving?

In my previous talk (Monday) I addressed this issue and mentioned that for group languages, this holds true (actually, more general, for polynomials of group languages).

Here, we investigate this question for the subclass of commutative regular languages.

Iterated Shuffle

1. The Iterated Shuffle is $L^{\sqcup\!\sqcup,*} = \bigcup_{i \geq 0} L^{i,\sqcup\!\sqcup}$ where $L^{0,\sqcup\!\sqcup} = \{\varepsilon\}$ and $L^{i+1,\sqcup\!\sqcup} = L^{i,\sqcup\!\sqcup} \sqcup\!\sqcup L$.
2. The regular languages are not closed under iterated shuffle in general.
3. Example: $\{ab, ba\}^{\sqcup\!\sqcup,*} = \{$words with equal number of a's and b's$\}$.

For what languages is the iterated shuffle regularlity-preserving?

In my previous talk (Monday) I addressed this issue and mentioned that for group languages, this holds true (actually, more general, for polynomials of group languages).

Here, we investigate this question for the subclass of commutative regular languages.

Iterated Shuffle

1. The Iterated Shuffle is $L^{\sqcup\!\sqcup,*} = \bigcup_{i \geq 0} L^{i,\sqcup\!\sqcup}$ where $L^{0,\sqcup\!\sqcup} = \{\varepsilon\}$ and $L^{i+1,\sqcup\!\sqcup} = L^{i,\sqcup\!\sqcup} \sqcup\!\sqcup L$.
2. The regular languages are not closed under iterated shuffle in general.
3. Example: $\{ab, ba\}^{\sqcup\!\sqcup,*} = \{$words with equal number of a's and b's$\}$.

For what languages is the iterated shuffle regularlity-preserving?

In my previous talk (Monday) I addressed this issue and mentioned that for group languages, this holds true (actually, more general, for polynomials of group languages).

Here, we investigate this question for the subclass of commutative regular languages.

Commutative Regular Languages

We will use a characterization of commutative regular languages by boolean operations.
Set

1. $F(a, r, n) = \{u \in \Sigma^* : |u|_a \equiv r \pmod{n}\}$,
2. $F(a, t) = \{u \in \Sigma^+ : |u|_a \geq t\}$.

Let Σ be a non-empty alphabet, $a \in \Sigma$ and $\Gamma \subseteq \Sigma$.

1. $F(a, 0, 1) = \Sigma^*$.
2. $F(a, 0, 2) \cap F(a, 3, 4) = \emptyset$.
3. $F(a, 1) = \Sigma^* a \Sigma^*$.
4. $\Gamma^* = \Sigma^* \setminus \left(\bigcup_{b \in \Sigma \setminus \Gamma} F(b, 1)\right)$.

Commutative Regular Languages

We will use a characterization of commutative regular languages by boolean operations.
Set

1. $F(a, r, n) = \{u \in \Sigma^* : |u|_a \equiv r \pmod{n}\}$,
2. $F(a, t) = \{u \in \Sigma^+ : |u|_a \geq t\}$.

Let Σ be a non-empty alphabet, $a \in \Sigma$ and $\Gamma \subseteq \Sigma$.

1. $F(a, 0, 1) = \Sigma^*$.
2. $F(a, 0, 2) \cap F(a, 3, 4) = \emptyset$.
3. $F(a, 1) = \Sigma^* a \Sigma^*$.
4. $\Gamma^* = \Sigma^* \setminus \left(\bigcup_{b \in \Sigma \setminus \Gamma} F(b, 1)\right)$.

Commutative Regular Languages

Theorem (Jean-Eric Pin, 1997)

The class of commutative regular languages is the boolean algebra generated by languages of the from $F(a, t), F(a, r, n)$, where $t \geq 0$, $0 \leq r < n$ and $a \in \Sigma$.

Definition

A diagonal periodic language over $\Gamma \subseteq \Sigma$ is a language of the form

$$\bigsqcup_{a \in \Gamma} a^{k_a}(a^{p_a})^*,$$

where $k_a \geq 0$ and $p_a > 0$ for $a \in \Gamma$ when $\Gamma \neq \emptyset$, or the language $\{\varepsilon\}$.

Proposition

The diagonal periodic languages are regular and commutative.

Commutative Regular Languages

Theorem (Jean-Eric Pin, 1997)

The class of commutative regular languages is the boolean algebra generated by languages of the from $F(a, t), F(a, r, n)$, where $t \geq 0$, $0 \leq r < n$ and $a \in \Sigma$.

Definition

A diagonal periodic language over $\Gamma \subseteq \Sigma$ is a language of the form

$$\bigsqcup_{a \in \Gamma} a^{k_a} (a^{p_a})^*,$$

where $k_a \geq 0$ and $p_a > 0$ for $a \in \Gamma$ when $\Gamma \neq \emptyset$, or the language $\{\varepsilon\}$.

Proposition

The diagonal periodic languages are regular and commutative.

Commutative Regular Languages

Proposition

The iterated shuffle of a diagonal periodic language $L \subseteq \Sigma^*$ over $\Gamma \subseteq \Sigma^*$ is a finite union of diagonal periodic languages. In particular, it is regular.

Proof.

Let $L = \coprod_{a\in\Gamma} a^{k_a}(a^{p_a})^*$ with $k_a \geq 0, p_a > 0$.

$\Rightarrow$ $L^{ш,*} = \{\varepsilon\} \cup \bigcup_{m>0} ш_{a\in\Gamma} a^{m\cdot k_a}(a^{p_a})^*$.

Set $N = \mathrm{lcm}((p_a)_{a\in\Gamma})$.

We have $m \cdot k_a + r_a \cdot p_a = (m - tN)\cdot k_a + (r_a + t\frac{N}{p_a}k_a)\cdot p_a$. Choose $t \geq 0$ s.t. $1 \leq m - tN \leq N$.

$\Rightarrow$ $\bigcup_{m>0} ш_{a\in\Gamma} a^{m\cdot k_a}(a^{p_a})^* = \bigcup_{i=1}^{N} ш_{a\in\Gamma} a^{i\cdot k_a}(a^{p_a})^*$.

So,

$$L^{ш,*} = \{\varepsilon\} \cup \bigcup_{i=1}^{N} ш_{a\in\Gamma} a^{i\cdot k_a}(a^{p_a})^*.$$

□

Commutative Regular Languages

Proposition

The iterated shuffle of a diagonal periodic language $L \subseteq \Sigma^*$ over $\Gamma \subseteq \Sigma^*$ is a finite union of diagonal periodic languages. In particular, it is regular.

Proof.

Let $L = \bigsqcup_{a \in \Gamma} a^{k_a}(a^{p_a})^*$ with $k_a \geq 0, p_a > 0$.

$\Rightarrow$ $L^{\sqcup\sqcup,*} = \{\varepsilon\} \cup \bigcup_{m>0} \sqcup\sqcup_{a\in\Gamma} a^{m\cdot k_a}(a^{p_a})^*$.

Set $N = \mathrm{lcm}((p_a)_{a\in\Gamma})$.

We have $m \cdot k_a + r_a \cdot p_a = (m - tN) \cdot k_a + (r_a + t\frac{N}{p_a}k_a) \cdot p_a$. Choose $t \geq 0$ s.t. $1 \leq m - tN \leq N$.

$\Rightarrow$ $\bigcup_{m>0} \sqcup\sqcup_{a\in\Gamma} a^{m\cdot k_a}(a^{p_a})^* = \bigcup_{i=1}^{N} \sqcup\sqcup_{a\in\Gamma} a^{i\cdot k_a}(a^{p_a})^*$.

So,

$$L^{\sqcup\sqcup,*} = \{\varepsilon\} \cup \bigcup_{i=1}^{N} \sqcup\sqcup_{a\in\Gamma} a^{i\cdot k_a}(a^{p_a})^*.$$

□

Commutative Regular Languages

Proposition

The iterated shuffle of a diagonal periodic language $L \subseteq \Sigma^*$ over $\Gamma \subseteq \Sigma^*$ is a finite union of diagonal periodic languages. In particular, it is regular.

Proof.

Let $L = \coprod_{a \in \Gamma} a^{k_a}(a^{p_a})^*$ with $k_a \geq 0, p_a > 0$.

$\Rightarrow$ $L^{ш,*} = \{\varepsilon\} \cup \bigcup_{m>0} ш_{a\in\Gamma} a^{m \cdot k_a}(a^{p_a})^*$.

Set $N = \text{lcm}((p_a)_{a\in\Gamma})$.

We have $m \cdot k_a + r_a \cdot p_a = (m - tN) \cdot k_a + (r_a + t\frac{N}{p_a}k_a) \cdot p_a$. Choose $t \geq 0$ s.t. $1 \leq m - tN \leq N$.

$\Rightarrow$ $\bigcup_{m>0} ш_{a\in\Gamma} a^{m \cdot k_a}(a^{p_a})^* = \bigcup_{i=1}^{N} ш_{a\in\Gamma} a^{i \cdot k_a}(a^{p_a})^*$.

So,

$$L^{ш,*} = \{\varepsilon\} \cup \bigcup_{i=1}^{N} ш_{a\in\Gamma} a^{i \cdot k_a}(a^{p_a})^*.$$

□

Commutative Regular Languages

Proposition

The iterated shuffle of a diagonal periodic language $L \subseteq \Sigma^*$ over $\Gamma \subseteq \Sigma^*$ is a finite union of diagonal periodic languages. In particular, it is regular.

Proof.

Let $L = \coprod_{a \in \Gamma} a^{k_a}(a^{p_a})^*$ with $k_a \geq 0, p_a > 0$.

$\Rightarrow$ $L^{ш,*} = \{\varepsilon\} \cup \bigcup_{m>0} ш_{a \in \Gamma} a^{m \cdot k_a}(a^{p_a})^*$.

Set $N = \mathrm{lcm}((p_a)_{a \in \Gamma})$.

We have $m \cdot k_a + r_a \cdot p_a = (m - tN) \cdot k_a + (r_a + t\frac{N}{p_a}k_a) \cdot p_a$. Choose $t \geq 0$ s.t. $1 \leq m - tN \leq N$.

$\Rightarrow$ $\bigcup_{m>0} ш_{a \in \Gamma} a^{m \cdot k_a}(a^{p_a})^* = \bigcup_{i=1}^{N} ш_{a \in \Gamma} a^{i \cdot k_a}(a^{p_a})^*$.

So,

$$L^{ш,*} = \{\varepsilon\} \cup \bigcup_{i=1}^{N} ш_{a \in \Gamma} a^{i \cdot k_a}(a^{p_a})^*.$$

□

Commutative Regular Languages

Proposition

The iterated shuffle of a diagonal periodic language $L \subseteq \Sigma^*$ over $\Gamma \subseteq \Sigma^*$ is a finite union of diagonal periodic languages. In particular, it is regular.

Proof.

Let $L = \coprod_{a \in \Gamma} a^{k_a}(a^{p_a})^*$ with $k_a \geq 0, p_a > 0$.

$\Rightarrow$ $L^{\sqcup\!\sqcup,*} = \{\varepsilon\} \cup \bigcup_{m>0} \sqcup\!\sqcup_{a\in\Gamma} a^{m\cdot k_a}(a^{p_a})^*$.

Set $N = \mathrm{lcm}((p_a)_{a\in\Gamma})$.

We have $m \cdot k_a + r_a \cdot p_a = (m - tN) \cdot k_a + (r_a + t\frac{N}{p_a}k_a) \cdot p_a$. Choose $t \geq 0$ s.t. $1 \leq m - tN \leq N$.

$\Rightarrow$ $\bigcup_{m>0} \sqcup\!\sqcup_{a\in\Gamma} a^{m\cdot k_a}(a^{p_a})^* = \bigcup_{i=1}^{N} \sqcup\!\sqcup_{a\in\Gamma} a^{i\cdot k_a}(a^{p_a})^*$.

So,

$$L^{\sqcup\!\sqcup,*} = \{\varepsilon\} \cup \bigcup_{i=1}^{N} \sqcup\!\sqcup_{a\in\Gamma} a^{i\cdot k_a}(a^{p_a})^*.$$

□

Commutative Regular Languages

Proposition

The iterated shuffle of a diagonal periodic language $L \subseteq \Sigma^*$ over $\Gamma \subseteq \Sigma^*$ is a finite union of diagonal periodic languages. In particular, it is regular.

Proof.

Let $L = \coprod_{a \in \Gamma} a^{k_a}(a^{p_a})^*$ with $k_a \geq 0, p_a > 0$.

$\Rightarrow$ $L^{\sqcup\!\sqcup,*} = \{\varepsilon\} \cup \bigcup_{m>0} \sqcup\!\sqcup_{a \in \Gamma} a^{m \cdot k_a}(a^{p_a})^*$.

Set $N = \text{lcm}((p_a)_{a \in \Gamma})$.

We have $m \cdot k_a + r_a \cdot p_a = (m - tN) \cdot k_a + (r_a + t\frac{N}{p_a}k_a) \cdot p_a$. Choose $t \geq 0$ s.t. $1 \leq m - tN \leq N$.

$\Rightarrow$ $\bigcup_{m>0} \sqcup\!\sqcup_{a \in \Gamma} a^{m \cdot k_a}(a^{p_a})^* = \bigcup_{i=1}^{N} \sqcup\!\sqcup_{a \in \Gamma} a^{i \cdot k_a}(a^{p_a})^*$.

So,

$$L^{\sqcup\!\sqcup,*} = \{\varepsilon\} \cup \bigcup_{i=1}^{N} \sqcup\!\sqcup_{a \in \Gamma} a^{i \cdot k_a}(a^{p_a})^*. \qquad \square$$

Commutative Regular Languages

Proposition

The positive boolean algebra generated by languages of the form $F(a,k)$, $F(a,k,n)$, $0 \leq k < n$, Γ^+ and Γ^*, $\Gamma \subseteq \Sigma$, is precisely the language class of finite unions of the diagonal periodic languages.

Theorem

Let $L \subseteq \Sigma^*$ be in the positive boolean algebra generated by languages of the form $F(a,k)$, $F(a,k,n)$, Γ^+ and Γ^* for $\Gamma \subseteq \Sigma$. Then, the iterated shuffle of L is contained in this positive boolean algebra. In particular, the iterated shuffle is regular.

Proof.

By the above, $(U \cup V)^{\sqcup\!\sqcup,*} = U^{\sqcup\!\sqcup,*} \sqcup\!\sqcup V^{\sqcup\!\sqcup,*}$ and the fact that shuffle distributes over union. □

Commutative Regular Languages

Proposition

The positive boolean algebra generated by languages of the form $F(a,k)$, $F(a,k,n)$, $0 \leq k < n$, Γ^+ and Γ^*, $\Gamma \subseteq \Sigma$, is precisely the language class of finite unions of the diagonal periodic languages.

Theorem

Let $L \subseteq \Sigma^*$ be in the positive boolean algebra generated by languages of the form $F(a,k)$, $F(a,k,n)$, Γ^+ and Γ^* for $\Gamma \subseteq \Sigma$. Then, the iterated shuffle of L is contained in this positive boolean algebra. In particular, the iterated shuffle is regular.

Proof.

By the above, $(U \cup V)^{⧢,*} = U^{⧢,*} ⧢ V^{⧢,*}$ and the fact that shuffle distributes over union. □

Commutative Regular Languages

Proposition

The positive boolean algebra generated by languages of the form $F(a, k)$, $F(a, k, n)$, $0 \leq k < n$, Γ^+ and Γ^*, $\Gamma \subseteq \Sigma$, is precisely the language class of finite unions of the diagonal periodic languages.

Theorem

Let $L \subseteq \Sigma^*$ be in the positive boolean algebra generated by languages of the form $F(a, k)$, $F(a, k, n)$, Γ^+ and Γ^* for $\Gamma \subseteq \Sigma$. Then, the iterated shuffle of L is contained in this positive boolean algebra. In particular, the iterated shuffle is regular.

Proof.

By the above, $(U \cup V)^{\sqcup\!\sqcup,*} = U^{\sqcup\!\sqcup,*} \sqcup\!\sqcup V^{\sqcup\!\sqcup,*}$ and the fact that shuffle distributes over union. $\square$

Finite Commutative Languages

Theorem

Let $L \subseteq \Sigma^*$ be a finite language. Then, $\mathrm{perm}(L)^{\sqcup\!\sqcup,*}$ is regular if and only if for any $a \in \Sigma$ with $\Sigma^* a \Sigma^* \cap L \neq \emptyset$ we have $a^+ \cap L \neq \emptyset$.

Corollary

Let $u \in \Sigma$ and $L \subseteq \Sigma^*$ be a finite language. Then, $\mathrm{perm}(u) \sqcup\!\sqcup \mathrm{perm}(L)^{\sqcup\!\sqcup,*}$ is regular if and only if for any $a \in \Sigma$ with $\Sigma^* a \Sigma^* \cap L \neq \emptyset$, we have $a^+ \cap L \neq \emptyset$.

1. The iterated shuffle of $\{ab, ba\} \cup \{c\} \sqcup\!\sqcup \{a, b\}^*$ is not regular.
2. The iterated shuffle of $\{ab, ba\} \sqcup\!\sqcup \{c\}^* \cup \{ac\} \sqcup\!\sqcup \{a, b\}^*$ is not regular.
3. The iterated shuffle of $\{ab, ba\} \cup \{c\} \sqcup\!\sqcup \{a, b\}^* \cup \mathrm{perm}(abb) \sqcup\!\sqcup \{a, b\}^*$ is regular.
4. The iterated shuffle of $\{ab, ba\} \cup \{c\} \sqcup\!\sqcup \{a, b\}^* \cup \mathrm{perm}(abb) \sqcup\!\sqcup \{a\}^* \cup \{bb\}$ is regular.

Finite Commutative Languages

Theorem

Let $L \subseteq \Sigma^*$ be a finite language. Then, $\mathrm{perm}(L)^{\sqcup\!\sqcup,*}$ is regular if and only if for any $a \in \Sigma$ with $\Sigma^* a \Sigma^* \cap L \neq \emptyset$ we have $a^+ \cap L \neq \emptyset$.

Corollary

Let $u \in \Sigma$ and $L \subseteq \Sigma^*$ be a finite language. Then, $\mathrm{perm}(u) \sqcup\!\sqcup \mathrm{perm}(L)^{\sqcup\!\sqcup,*}$ is regular if and only if for any $a \in \Sigma$ with $\Sigma^* a \Sigma^* \cap L \neq \emptyset$, we have $a^+ \cap L \neq \emptyset$.

1. The iterated shuffle of $\{ab, ba\} \cup \{c\} \sqcup\!\sqcup \{a, b\}^*$ is not regular.
2. The iterated shuffle of $\{ab, ba\} \sqcup\!\sqcup \{c\}^* \cup \{ac\} \sqcup\!\sqcup \{a, b\}^*$ is not regular.
3. The iterated shuffle of $\{ab, ba\} \cup \{c\} \sqcup\!\sqcup \{a, b\}^* \cup \mathrm{perm}(abb) \sqcup\!\sqcup \{a, b\}^*$ is regular.
4. The iterated shuffle of $\{ab, ba\} \cup \{c\} \sqcup\!\sqcup \{a, b\}^* \cup \mathrm{perm}(abb) \sqcup\!\sqcup \{a\}^* \cup \{bb\}$ is regular.

Finite Commutative Languages

Theorem
Let $L \subseteq \Sigma^*$ be a finite language. Then, $\mathrm{perm}(L)^{\sqcup\!\sqcup,*}$ is regular if and only if for any $a \in \Sigma$ with $\Sigma^* a \Sigma^* \cap L \neq \emptyset$ we have $a^+ \cap L \neq \emptyset$.

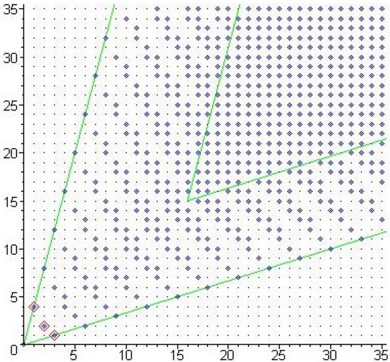
Corollary
Let $u \in \Sigma$ and $L \subseteq \Sigma^*$ be a finite language. Then, $\mathrm{perm}(u) \sqcup\!\sqcup \mathrm{perm}(L)^{\sqcup\!\sqcup,*}$ is regular if and only if for any $a \in \Sigma$ with $\Sigma^* a \Sigma^* \cap L \neq \emptyset$, we have $a^+ \cap L \neq \emptyset$.

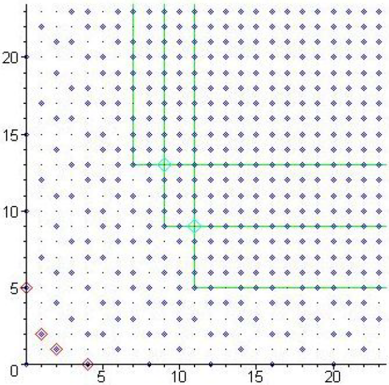
1. The iterated shuffle of $\{ab, ba\} \cup \{c\} \sqcup\!\sqcup \{a, b\}^*$ is not regular.
2. The iterated shuffle of $\{ab, ba\} \sqcup\!\sqcup \{c\}^* \cup \{ac\} \sqcup\!\sqcup \{a, b\}^*$ is not regular.
3. The iterated shuffle of $\{ab, ba\} \cup \{c\} \sqcup\!\sqcup \{a, b\}^* \cup \mathrm{perm}(abb) \sqcup\!\sqcup \{a, b\}^*$ is regular.
4. The iterated shuffle of $\{ab, ba\} \cup \{c\} \sqcup\!\sqcup \{a, b\}^* \cup \mathrm{perm}(abb) \sqcup\!\sqcup \{a\}^* \cup \{bb\}$ is regular.

Finite Commutative Languages

Theorem

Let $L \subseteq \Sigma^*$ be a finite language. Then, $\mathrm{perm}(L)^{\sqcup\!\sqcup,*}$ is regular if and only if for any $a \in \Sigma$ with $\Sigma^* a \Sigma^* \cap L \neq \emptyset$ we have $a^+ \cap L \neq \emptyset$.





Images from Vadim Ponomarenko, *The Multi-Dimensional Frobenius Problem and Vector GCDs* (Talk), `https://vadim.sdsu.edu//frob-gcd.pdf`

$\Rightarrow$

1. W.l.o.g there exists $a \in \Sigma$ such that $|u|_a > 0$ implies $u \notin a^+$ and $|v|_a > 0$ for at least one $v \in L$.
2. (Ehrenfeucht, Haussler, Rozenberg 1983) perm(L) is regular iff $\psi(L)$ is a finite union of sets of the form

$$u + \{c_1 u_1 + \ldots + c_m u_m \mid c_1, \ldots, c_m \in \mathbb{N}_0, u_1, \ldots, u_m \in A\}$$

   for $u \in \mathbb{N}_0$ and finite $A \subseteq \{c \cdot \psi(a) \mid a \in \Sigma \text{ and } c \in \mathbb{N}_0\}$ ($\psi(a)$ are *standard basis vectors*).
3. With the assumption, we can show that an affine one-dimensional set of the form $u + \mathbb{N}_0 \cdot ce_i$ for some stretched standard basis vector ce_i intersects with finitely many (slanted and shifted) rays infinitely often. However, by assumption these rays are not parallel to any e_i, hence (linear algebra) each ray has precisely one intersection with the ray $\mathbb{N}_0 \cdot e_i$.

$\Rightarrow$

1. W.l.o.g there exists $a \in \Sigma$ such that $|u|_a > 0$ implies $u \notin a^+$ and $|v|_a > 0$ for at least one $v \in L$.
2. (Ehrenfeucht, Haussler, Rozenberg 1983) perm(L) is regular iff $\psi(L)$ is a finite union of sets of the form

$$u + \{c_1 u_1 + \ldots + c_m u_m \mid c_1, \ldots, c_m \in \mathbb{N}_0, u_1, \ldots, u_m \in A\}$$

   for $u \in \mathbb{N}_0$ and finite $A \subseteq \{c \cdot \psi(a) \mid a \in \Sigma \text{ and } c \in \mathbb{N}_0\}$ ($\psi(a)$ are *standard basis vectors*).
3. With the assumption, we can show that an affine one-dimensional set of the form $u + \mathbb{N}_0 \cdot ce_i$ for some stretched standard basis vector ce_i intersects with finitely many (slanted and shifted) rays infinitely often. However, by assumption these rays are not parallel to any e_i, hence (linear algebra) each ray has precisely one intersection with the ray $\mathbb{N}_0 \cdot e_i$.

$\Rightarrow$

1. W.l.o.g there exists $a \in \Sigma$ such that $|u|_a > 0$ implies $u \notin a^+$ and $|v|_a > 0$ for at least one $v \in L$.
2. (Ehrenfeucht, Haussler, Rozenberg 1983) perm(L) is regular iff $\psi(L)$ is a finite union of sets of the form

$$u + \{c_1 u_1 + \ldots + c_m u_m \mid c_1, \ldots, c_m \in \mathbb{N}_0, u_1, \ldots, u_m \in A\}$$

   for $u \in \mathbb{N}_0$ and finite $A \subseteq \{c \cdot \psi(a) \mid a \in \Sigma \text{ and } c \in \mathbb{N}_0\}$ ($\psi(a)$ are *standard basis vectors*).
3. With the assumption, we can show that an affine one-dimensional set of the form $u + \mathbb{N}_0 \cdot ce_i$ for some stretched standard basis vector ce_i intersects with finitely many (slanted and shifted) rays infinitely often. However, by assumption these rays are not parallel to any e_i, hence (linear algebra) each ray has precisely one intersection with the ray $\mathbb{N}_0 \cdot e_i$.

$\Rightarrow$

1. W.l.o.g there exists $a \in \Sigma$ such that $|u|_a > 0$ implies $u \notin a^+$ and $|v|_a > 0$ for at least one $v \in L$.
2. (Ehrenfeucht, Haussler, Rozenberg 1983) perm(L) is regular iff $\psi(L)$ is a finite union of sets of the form

$$u + \{c_1 u_1 + \ldots + c_m u_m \mid c_1, \ldots, c_m \in \mathbb{N}_0, u_1, \ldots, u_m \in A\}$$

   for $u \in \mathbb{N}_0$ and finite $A \subseteq \{c \cdot \psi(a) \mid a \in \Sigma \text{ and } c \in \mathbb{N}_0\}$ ($\psi(a)$ are *standard basis vectors*).
3. With the assumption, we can show that an affine one-dimensional set of the form $u + \mathbb{N}_0 \cdot ce_i$ for some stretched standard basis vector ce_i intersects with finitely many (slanted and shifted) rays infinitely often. However, by assumption these rays are not parallel to any e_i, hence (linear algebra) each ray has precisely one intersection with the ray $\mathbb{N}_0 \cdot e_i$.

$\Leftarrow$

1. Let $L = \{u_1, \ldots, u_n\}$. W.l.o.g. $\Sigma = \{a_1, \ldots, a_k\}$ and $u_i \in a_i^+$ for $\{u_1, \ldots, u_k\} \subseteq L$. Set $m_i = |u_i| = \psi(u_i)$ for $i \in \{1, \ldots, k\}$.
2. For $i > k$ write $\psi(u_i) = x_1\psi(a_1) + \ldots + x_k\psi(a_k)$.
   Then, $m_1 \cdots m_k \psi(u_i) = x_1 \frac{m_1 \cdots m_k}{m_1}\psi(u_1) + \ldots + x_k \frac{m_1 \cdots m_k}{m_k}\psi(u_k)$.
3. So $c_1\psi(u_1) + \ldots + c_n\psi(u_n)$ equals

$$\left(c_1 + x_{a_1}\frac{m_1 \cdots m_k}{m_1}\right)\psi(u_1) + \ldots + \left(c_k + x_{a_k}\frac{m_1 \cdots m_k}{m_k}\right)\psi(u_k) + \ldots +$$
$$c_{i-1}\psi(u_{i-1}) + (c_i - m_1 \cdots m_k)\psi(u_i) + c_{i+1}\psi(u_{i+1}) + \ldots + c_n\psi(u_n).$$

4. Hence, $L^{\sqcup\sqcup,*}$ equals:

$$\bigcup_{\substack{(c_{k+1},\ldots,c_n) \\ 0 \le c_i < m_1 \cdots m_k}} \mathrm{perm}(u_1^*) \sqcup\sqcup \ldots \sqcup\sqcup \mathrm{perm}(u_k^*) \sqcup\sqcup \mathrm{perm}(u_{k+1}^{c_{k+1}}) \sqcup\sqcup \ldots \sqcup\sqcup \mathrm{perm}(u_n^{c_n}).$$

For $i \in \{1, \ldots, k\}$, we have $u_i \in a_i^+$, so $\mathrm{perm}(u_i^*) = u_i^*$. $\square$

$\Leftarrow$

1. Let $L = \{u_1, \ldots, u_n\}$. W.l.o.g. $\Sigma = \{a_1, \ldots, a_k\}$ and $u_i \in a_i^+$ for $\{u_1, \ldots, u_k\} \subseteq L$. Set $m_i = |u_i| = \psi(u_i)$ for $i \in \{1, \ldots, k\}$.
2. For $i > k$ write $\psi(u_i) = x_1\psi(a_1) + \ldots + x_k\psi(a_k)$.
   Then, $m_1 \cdots m_k \psi(u_i) = x_1 \frac{m_1 \cdots m_k}{m_1} \psi(u_1) + \ldots + x_k \frac{m_1 \cdots m_k}{m_k} \psi(u_k)$.
3. So $c_1\psi(u_1) + \ldots + c_n\psi(u_n)$ equals

$$\left(c_1 + x_{a_1} \frac{m_1 \cdots m_k}{m_1}\right) \psi(u_1) + \ldots + \left(c_k + x_{a_k} \frac{m_1 \cdots m_k}{m_k}\right) \psi(u_k) + \ldots +$$
$$c_{i-1}\psi(u_{i-1}) + (c_i - m_1 \cdots m_k)\psi(u_i) + c_{i+1}\psi(u_{i+1}) + \ldots + c_n\psi(u_n).$$

4. Hence, $L^{\sqcup,*}$ equals:

$$\bigcup_{\substack{(c_{k+1},\ldots,c_n) \\ 0 \le c_i < m_1 \cdots m_k}} \mathrm{perm}(u_1^*) \sqcup \ldots \sqcup \mathrm{perm}(u_k^*) \sqcup \mathrm{perm}(u_{k+1}^{c_{k+1}}) \sqcup \ldots \sqcup \mathrm{perm}(u_n^{c_n}).$$

For $i \in \{1, \ldots, k\}$, we have $u_i \in a_i^+$, so $\mathrm{perm}(u_i^*) = u_i^*$. $\square$

$\Leftarrow$

1. Let $L = \{u_1, \ldots, u_n\}$. W.l.o.g. $\Sigma = \{a_1, \ldots, a_k\}$ and $u_i \in a_i^+$ for $\{u_1, \ldots, u_k\} \subseteq L$. Set $m_i = |u_i| = \psi(u_i)$ for $i \in \{1, \ldots, k\}$.
2. For $i > k$ write $\psi(u_i) = x_1\psi(a_1) + \ldots + x_k\psi(a_k)$.
   Then, $m_1 \cdots m_k \psi(u_i) = x_1 \frac{m_1 \cdots m_k}{m_1} \psi(u_1) + \ldots + x_k \frac{m_1 \cdots m_k}{m_k} \psi(u_k)$.
3. So $c_1\psi(u_1) + \ldots + c_n\psi(u_n)$ equals

$$\left(c_1 + x_{a_1} \frac{m_1 \cdots m_k}{m_1}\right) \psi(u_1) + \ldots + \left(c_k + x_{a_k} \frac{m_1 \cdots m_k}{m_k}\right) \psi(u_k) + \ldots +$$
$$c_{i-1}\psi(u_{i-1}) + (c_i - m_1 \cdots m_k)\psi(u_i) + c_{i+1}\psi(u_{i+1}) + \ldots + c_n\psi(u_n).$$

4. Hence, $L^{\sqcup\!\sqcup,*}$ equals:

$$\bigcup_{\substack{(c_{k+1}, \ldots, c_n) \\ 0 \leq c_i < m_1 \cdots m_k}} \mathrm{perm}(u_1^*) \sqcup\!\sqcup \ldots \sqcup\!\sqcup \mathrm{perm}(u_k^*) \sqcup\!\sqcup \mathrm{perm}(u_{k+1}^{c_{k+1}}) \sqcup\!\sqcup \ldots \sqcup\!\sqcup \mathrm{perm}(u_n^{c_n}).$$

For $i \in \{1, \ldots, k\}$, we have $u_i \in a_i^+$, so $\mathrm{perm}(u_i^*) = u_i^*$. $\square$

Open Problems

1. Characterize the class of commutative languages whose iterated shuffle is regular in easy terms (for example, as a (positive) boolean algebra).
2. Smaller Step: Do this for the star-free commutative languages (a sufficient condition is in the paper).
3. Computational Complexity to decide regularity of the iterated shuffle of commutative languages. (The given criterion for finite languages is easily seen to be in P).
4. Construct (optimal) automata in case the iterated shuffle is regular.