

CM0081 Formal Languages and Automata

§ 4.2 Closure Properties of Regular Languages

Andrés Sicard-Ramírez

Universidad EAFIT

Semester 2024-1

Preliminaries

Conventions

- ▶ The number and page numbers assigned to chapters, examples, exercises, figures, quotes, sections and theorems on these slides correspond to the numbers assigned in the textbook [Hopcroft, Motwani and Ullman (1979) 2007].
- ▶ The natural numbers include the zero, that is, $\mathbb{N} = \{0, 1, 2, \dots\}$.
- ▶ The power set of a set A , that is, the set of its subsets, is denoted by $\mathcal{P}A$.

Introduction

Let L and L' be regular languages. The following languages are regular:

$L \cup L'$	(union)
$L \cap L'$	(intersection)
\overline{L}	(complement)
$L - L'$	(difference)
L^R	(reversal)
L^*	(closure)
$L \cdot L'$	(concatenation)
$h(L)$	(homomorphism)
$h^{-1}(L)$	(inverse homomorphism)

Closure Under Union

Theorem 4.4

If L and L' are regular languages, then so is $L \cup L'$.

Closure Under Union

Theorem 4.4

If L and L' are regular languages, then so is $L \cup L'$.

Proof

(Using regular expressions)

Closure Under Complementation

Definition

Let L be a language over alphabet Σ . The **complement** of L is defined by

$$\overline{L} := \Sigma^* - L.$$

Closure Under Complementation

Theorem 4.5

If L is a regular language, then so is \overline{L} .

Closure Under Complementation

Theorem 4.5

If L is a regular language, then so is \overline{L} .

Proof

Let $A = (Q, \Sigma, \delta, q_0, F)$ be a DFA that accepts L . Then $B = (Q, \Sigma, \delta, q_0, Q - F)$ is a DFA that accepts \overline{L} . ■

Closure Under Complementation

Question

'Do you see how to take a regular expression and change it into one that defines the complement language?' [Hopcroft, Motwani and Ullman [1979] 2007, p. 136]

Closure Under Complementation

Observation

Using the closure properties we can prove that a language is not regular.

Closure Under Complementation

Observation

Using the closure properties we can prove that a language is not regular.

Example

Given that

$$L_{=} = \{ w \in \{0,1\}^* \mid w \text{ has an equal numbers of 0's and 1's} \}$$

is a language not regular. Prove that

$$L_{\neq} = \{ w \in \{0,1\}^* \mid w \text{ has an unequal numbers of 0's and 1's} \}$$

is a language not regular.

Closure Under Complementation

Observation

Using the closure properties we can prove that a language is not regular.

Example

Given that

$$L_{=} = \{ w \in \{0,1\}^* \mid w \text{ has an equal numbers of 0's and 1's} \}$$

is a language not regular. Prove that

$$L_{\neq} = \{ w \in \{0,1\}^* \mid w \text{ has an unequal numbers of 0's and 1's} \}$$

is a language not regular.

Proof

Whiteboard.

Product Construction

Construction

Let A_L , A_M and A be DFAs given by

$$A_L = (Q_L, \Sigma, \delta_L, q_L, F_L),$$

$$A_M = (Q_M, \Sigma, \delta_M, q_M, F_M),$$

$$A = (Q_L \times Q_M, \Sigma, \delta, (q_L, q_M), F_L \times F_M),$$

where

$$\delta : (Q_L \times Q_M) \times \Sigma \rightarrow Q_L \times Q_M$$

$$\delta((p, q), a) = (\delta_L(p, a), \delta_M(q, a)).$$

Product Construction

Theorem (Exercise 4.2.15)

For all $w \in \Sigma^*$,

$$\hat{\delta}((q_L, q_M), w) = (\hat{\delta}_L(q_L, w), \hat{\delta}_M(q_M, w)).$$

(continued on next slide)

Product Construction

Proof by induction on w

1. Basis step

$$\begin{aligned}\hat{\delta}((q_L, q_M), \varepsilon) &= (q_L, q_M) && \text{(def. of } \hat{\delta}) \\ &= (\hat{\delta}_L(q_L, \varepsilon), \hat{\delta}_M(q_M, \varepsilon)) && \text{(def. of } \hat{\delta}_L \text{ and } \hat{\delta}_M)\end{aligned}$$

(continued on next slide)

2. Inductive step

$$\begin{aligned}\hat{\delta}((q_L, q_M), xa) \\&= \delta(\hat{\delta}((q_L, q_M), x), a) && \text{(def. of } \hat{\delta}) \\&= \delta((\hat{\delta}_L(q_L, x), \hat{\delta}_M(q_M, x)), a) && \text{(by IH)} \\&= (\delta_L(\hat{\delta}_L(q_L, x), a), \delta_M(\hat{\delta}_M(q_M, x), a)) && \text{(def. of } \delta) \\&= (\hat{\delta}_L(q_L, xa), \hat{\delta}_M(q_M, xa)) && \text{(def. of } \hat{\delta}_L \text{ and } \hat{\delta}_M)\end{aligned}$$



Closure Under Intersection

Theorem 4.8

If L and L' are regular languages, then so is $L \cap L'$.

Closure Under Intersection

Theorem 4.8

If L and L' are regular languages, then so is $L \cap L'$.

Proof

Let A_L and $A_{L'}$ be DFAs accepting L and L' . The product construction of A_L and $A_{L'}$ accepts $L \cap L'$. ■

Closure Under Intersection

Theorem 4.8

If L and L' are regular languages, then so is $L \cap L'$.

Proof

Let A_L and $A_{L'}$ be DFAs accepting L and L' . The product construction of A_L and $A_{L'}$ accepts $L \cap L'$. ■

Different proof

The regular languages are closure under union and complement, and

$$L \cap L' = \overline{\overline{L} \cup \overline{L'}}.$$

■

Closure Under Reversal

Definition

Let $w = a_1 a_2 \cdots a_n$ be a word. The **reversal** of w is defined by

$$w^R := a_n a_{n-1} \cdots a_1.$$

Closure Under Reversal

Definition

Let $w = a_1 a_2 \cdots a_n$ be a word. The **reversal** of w is defined by

$$w^R := a_n a_{n-1} \cdots a_1.$$

Definition

Let L be a language on alphabet Σ . The **reversal** of L is defined by

$$L^R := \{ w^R \in \Sigma^* \mid w \in L \}.$$

Closure Under Reversal

Definition

Let $w = a_1 a_2 \cdots a_n$ be a word. The **reversal** of w is defined by

$$w^R := a_n a_{n-1} \cdots a_1.$$

Definition

Let L be a language on alphabet Σ . The **reversal** of L is defined by

$$L^R := \{ w^R \in \Sigma^* \mid w \in L \}.$$

Theorem 4.11

If L is regular language, then so is L^R (proof using automata or regular expressions)

Closure Under Reversal

Proof using automata

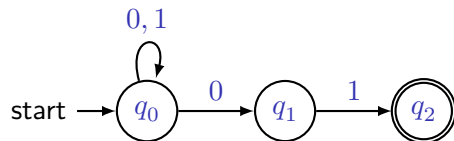
Let L be recognized by a finite automaton A . From the automaton A we get a finite automaton for L^R , by

1. Reversing all arcs.
2. Make the start state of A be the only accepting state.
3. Create a new start state p_0 with transitions $\delta(p_0, \varepsilon) = f$, where $f \in F$ are the accepting states of A . ■

Closure Under Reversal

Example

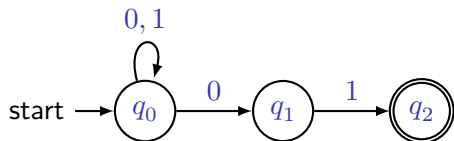
A NFA accepting all the binary strings that end in 01.



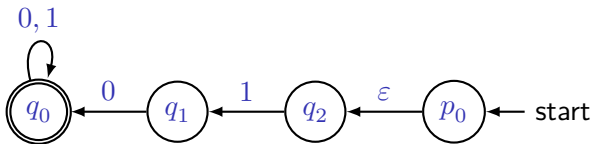
Closure Under Reversal

Example

A NFA accepting all the binary strings that end in 01.



A NFA accepting all the binary strings that start with 10.



Homomorphisms

Definition

An **algebraic structure** on a set $A \neq \emptyset$ is essentially a collection of n -ary operations on A [Birkhoff 1946, 1987].

Homomorphisms

Definition

An **algebraic structure** on a set $A \neq \emptyset$ is essentially a collection of n -ary operations on A [Birkhoff 1946, 1987].

Example (Semigroup)

A **semigroup** $(S, *)$ is a set S with an associative binary operation $* : S \times S \rightarrow S$.

Homomorphisms

Definition

An **algebraic structure** on a set $A \neq \emptyset$ is essentially a collection of n -ary operations on A [Birkhoff 1946, 1987].

Example (Semigroup)

A **semigroup** $(S, *)$ is a set S with an associative binary operation $* : S \times S \rightarrow S$.

Example (Monoid)

A **monoid** $(M, *, \varepsilon)$ is a semigroup $(M, *)$ with an element $\varepsilon \in M$ which is an unit for $*$, i.e. $(\forall x)(x * \varepsilon = \varepsilon * x = x)$.

Homomorphisms

Definition

A **homomorphism** is a structure-preserving map between two algebraic structures.

Homomorphisms

Definition

A **homomorphism** is a structure-preserving map between two algebraic structures.

Example

A homomorphism between two semigroups $(S, *)$ and $(S', *')$ is a function $\varphi : S \rightarrow S'$ such that:

$$(\forall x)(\forall y)[\varphi(x * y) = \varphi(x) *' \varphi(y)].$$

Homomorphisms

Definition

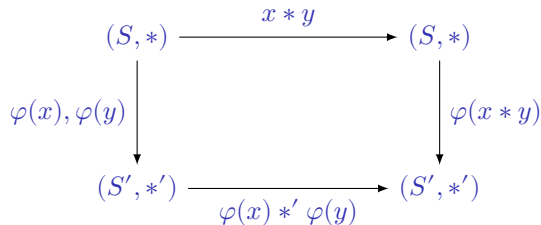
A **homomorphism** is a structure-preserving map between two algebraic structures.

Example

A homomorphism between two semigroups $(S, *)$ and $(S', *')$ is a function $\varphi : S \rightarrow S'$ such that:

$$(\forall x)(\forall y)[\varphi(x * y) = \varphi(x) *' \varphi(y)].$$

Graphically,



Homomorphisms

Example

A homomorphism between two monoids $(M, *, \varepsilon)$ and $(M', *', \varepsilon')$ is a function $\varphi : M \rightarrow M'$ such that:

$$\begin{aligned} &(\forall x)(\forall y)[\varphi(x * y) = \varphi(x) *' \varphi(y)], \\ &\varphi(\varepsilon) = \varepsilon'. \end{aligned}$$

Homomorphisms

Definition

A homomorphism φ between two algebraic structures is [Cohn (1965) 1981]:

- ▶ a **monomorphism** if φ is an injection,
- ▶ an **epimorphism** if φ is a surjection,
- ▶ an **endomorphism** if φ is from an algebraic structure to itself,
- ▶ an **isomorphism** if φ is a bijection,
- ▶ an **automorphism** if φ is a bijective endomorphism.

Closure Under Homomorphism

Definition

Let Σ and Γ be two alphabets. A **homomorphism** between (the monoids) $(\Sigma^*, \cdot, \varepsilon)$ and $(\Gamma^*, \cdot, \varepsilon)$ is a function

$$\begin{aligned}h &: \Sigma^* \rightarrow \Gamma^* \\a_1 a_2 \cdots a_n &\mapsto h(a_1) h(a_2) \cdots h(a_n) \\ \varepsilon &\mapsto \varepsilon\end{aligned}$$

Note: For this reason the textbook talks about a homomorphism $h : \Sigma \rightarrow \Gamma^*$.

Closure Under Homomorphism

Example

Let $h : \{0, 1\}^* \rightarrow \{a, b\}^*$ be a homomorphism defined by

$$h(0) = ab, \quad h(1) = \varepsilon.$$

Then

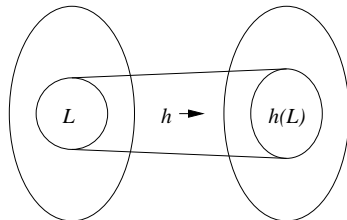
$$\begin{aligned} h(0011) &= h(0)h(0)h(1)h(1) \\ &= abab. \end{aligned}$$

Closure Under Homomorphism

Definition

Let L be a language over an alphabet Σ and let h be a homomorphism on Σ . The **application** of h to L , denoted $h(L)$, is defined by[†]

$$h(L) := \{ h(w) \mid w \in L \}.$$



[†]Figure from Hopcroft, Motwani and Ullman [(1979) 2007, Fig. 4.5a].

Closure Under Homomorphism

Example

Let $h : \{0, 1\}^* \rightarrow \{a, b\}^*$ be a homomorphism defined by

$$h(0) = ab, \quad h(1) = \varepsilon.$$

If $L = L(10^*1)$, then $h(L) = L((ab)^*)$.

Closure Under Homomorphism

Theorem 4.14

If L is a regular language over the alphabet Σ and h is a homomorphism on Σ , then $h(L)$ is also regular.

Closure Under Homomorphism

Theorem 4.14

If L is a regular language over the alphabet Σ and h is a homomorphism on Σ , then $h(L)$ is also regular.

Proof plan

- ▶ Let E be a regular expression such that $L = L(E)$.
- ▶ Let $h(E)$ be the regular expression replacing each symbol $a \in \Sigma$ by $h(a)$ in the regular expression E .
- ▶ We need to prove that $L(h(E)) = h(L(E))$.

(continued on next slide)

Closure Under Homomorphism

Proving $L(h(E)) = h(L(E))$

► Basis step

► E is ε or \emptyset .

1. $h(E) = E$ (h does not affect E)
2. $h(L(E)) = L(E)$ ($L(E)$ is empty or only contains ε)
3. $L(h(E)) = L(E) = h(L(E))$ (by 1 and 2)

(continued on next slide)

Closure Under Homomorphism

Proving $L(h(E)) = h(L(E))$

► Basis step

► $E = a$

1. $L(E) = \{a\}$
2. $h(L(E)) = \{h(a)\}$
3. $h(E)$ is the regular expression that is the string of symbols $h(a)$
4. $L(h(E)) = \{h(a)\}$
5. $L(h(E)) = h(L(E))$ (by transitivity between 2 and 4)

(continued on next slide)

Closure Under Homomorphism

Proving $L(h(E)) = h(L(E))$ (continuation)

► Inductive step

► $E = F + G$

1. $L(E) = L(F) \cup L(G)$ (def. of $+$)
2. $h(E) = h(F + G) = h(F) + h(G)$ (def. of $h(E)$)
3. $L(h(E)) = L(h(F) + h(G)) = L(h(F)) \cup L(h(G))$ (def. of $+$)
4. $h(L(E)) = h(L(F) \cup L(G)) = h(L(F)) \cup h(L(G))$ (h is applied to a language by application to each of its strings)
5. $L(h(F)) = h(L(F))$ and $L(h(G)) = h(L(G))$ (IH)
6. $L(h(E)) = h(L(E))$

(continued on next slide)

Closure Under Homomorphism

Proving $L(h(E)) = h(L(E))$ (continuation)

- ▶ Inductive step
 - ▶ $E = FG$ (similar to the previous case)

Closure Under Homomorphism

Proving $L(h(E)) = h(L(E))$ (continuation)

► Inductive step

► $E = FG$ (similar to the previous case)

► $E = F^*$ (similar to the previous case)

1. $L(E) = (L(F))^*$ (def. of $*$)

2. $h(E) = h(F^*) = (h(F))^*$ (def. of $h(E)$)

3. $L(h(E)) = L((h(F))^*) = (L(h(F)))^*$ (def. of $*$)

4. $h(L(E)) = h((L(F))^*) = (h(L(F)))^*$ (h is applied to a language by application to each of its strings)

5. $L(h(F)) = h(L(F))$ (IH)

6. $L(h(E)) = h(L(E))$



Closure Under Homomorphism

Example

Let $\Sigma = \{0, 1, 2\}$. Prove that L is a language not regular.

$$L = \{ 0^i 1^j 2^k \mid i, j, k \in \mathbb{Z}^+ \text{ and } i \neq j \neq k \}.$$

Closure Under Homomorphism

Example

Let $\Sigma = \{0, 1, 2\}$. Prove that L is a language not regular.

$$L = \{ 0^i 1^j 2^k \mid i, j, k \in \mathbb{Z}^+ \text{ and } i \neq j \neq k \}.$$

Proof

1. We define the homomorphism

$$h(0) = 0, \quad h(1) = 1, \quad h(2) = \varepsilon.$$

Closure Under Homomorphism

Example

Let $\Sigma = \{0, 1, 2\}$. Prove that L is a language not regular.

$$L = \{ 0^i 1^j 2^k \mid i, j, k \in \mathbb{Z}^+ \text{ and } i \neq j \neq k \}.$$

Proof

1. We define the homomorphism

$$h(0) = 0, \quad h(1) = 1, \quad h(2) = \varepsilon.$$

2. The homomorphism h removes the 2^k s, so

$$h(L) = \{ 0^i 1^j \mid i, j \in \mathbb{Z}^+ \text{ and } i \neq j \}.$$

Closure Under Homomorphism

Example

Let $\Sigma = \{0, 1, 2\}$. Prove that L is a language not regular.

$$L = \{ 0^i 1^j 2^k \mid i, j, k \in \mathbb{Z}^+ \text{ and } i \neq j \neq k \}.$$

Proof

1. We define the homomorphism

$$h(0) = 0, \quad h(1) = 1, \quad h(2) = \varepsilon.$$

2. The homomorphism h removes the 2^k s, so

$$h(L) = \{ 0^i 1^j \mid i, j \in \mathbb{Z}^+ \text{ and } i \neq j \}.$$

3. We know that $h(L)$ is not regular, so L is not regular. ■

Closure Under Homomorphism

Example

Let L be a regular language and h a homomorphism on L . Define $h^*(L)$ by

$$h^*(L) = L \cup h(L) \cup h(h(L)) \cup h(h(h(L))) \cup \dots$$

Is $h^*(L)$ necessarily regular?

[†]From somewhere in Internet (I don't remember).

Closure Under Homomorphism

Example

Let L be a regular language and h a homomorphism on L . Define $h^*(L)$ by

$$h^*(L) = L \cup h(L) \cup h(h(L)) \cup h(h(h(L))) \cup \dots$$

Is $h^*(L)$ necessarily regular?

Solution

No. Let $L = \{01\}$ and h defined as $h(0) = 00$ and $h(1) = 11$. Then

$$\begin{aligned} h^*(L) &= \{01, 0011, 00001111, \dots\} \\ &= \{0^n 1^n \mid n = 2^k \text{ for } k \geq 0\}, \end{aligned}$$

which is a language not regular.[†]

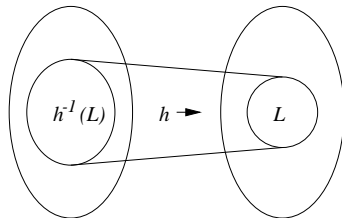
[†]From somewhere in Internet (I don't remember).

Closure Under Inverse Homomorphism

Definition

Let $h : \Sigma^* \rightarrow \Gamma^*$ be a homomorphism and $L \subseteq \Gamma^*$ a language. The **application** of h^{-1} to L , denoted $h^{-1}(L)$, is defined by[†]

$$h^{-1}(L) := \{ w \in \Sigma^* \mid h(w) \in L \}.$$



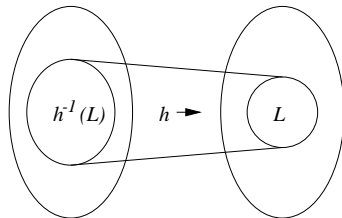
[†]Figure from Hopcroft, Motwani and Ullman [(1979) 2007, Fig. 4.5b].

Closure Under Inverse Homomorphism

Definition

Let $h : \Sigma^* \rightarrow \Gamma^*$ be a homomorphism and $L \subseteq \Gamma^*$ a language. The **application** of h^{-1} to L , denoted $h^{-1}(L)$, is defined by[†]

$$h^{-1}(L) := \{ w \in \Sigma^* \mid h(w) \in L \}.$$



Observation

Note that h^{-1} is a relation but it is not necessarily a function.

[†]Figure from Hopcroft, Motwani and Ullman [(1979) 2007, Fig. 4.5b].

Closure Under Inverse Homomorphism

Example

Let $h : \{a, b\} \rightarrow \{0, 1\}^*$ a homomorphism defined by

$$h(a) = 01, \quad h(b) = 10,$$

and let L be the language denoted by the regular expression $(\mathbf{00} + \mathbf{1})^*$, i.e.

$$L = \{ w \in \{0, 1\}^* \mid \text{all the } 0\text{'s occur in adjacent pairs} \}.$$

Closure Under Inverse Homomorphism

Example

Let $h : \{a, b\} \rightarrow \{0, 1\}^*$ a homomorphism defined by

$$h(a) = 01, \quad h(b) = 10,$$

and let L be the language denoted by the regular expression $(00 + 1)^*$, i.e.

$$L = \{ w \in \{0, 1\}^* \mid \text{all the } 0\text{'s occur in adjacent pairs} \}.$$

Then

$$h^{-1}(L) = L((ba)^*).$$

Note that h^{-1} is not a function, but a relation.

It is necessary to prove $h(w) \in L \Leftrightarrow w = baba \cdots ba$.

Closure Under Inverse Homomorphism

Theorem 4.16

Let $h : \Sigma^* \rightarrow \Gamma^*$ be a homomorphism and $L \subseteq \Gamma^*$ a regular language. Then $h^{-1}(L)$ is regular (proof using automata).

Closure Under Inverse Homomorphism

Example

Prove that $L = \{ 0^n 1^{2n} \mid n \geq 0 \}$ is a language not regular.

Closure Under Inverse Homomorphism

Example

Prove that $L = \{ 0^n 1^{2n} \mid n \geq 0 \}$ is a language not regular.

Proof

1. Given the homomorphism

$$h(0) = 0, \quad h(1) = 11,$$

then

$$h^{-1}(L) = \{ 0^n 1^n \mid n \geq 0 \}.$$

Closure Under Inverse Homomorphism

Example

Prove that $L = \{ 0^n 1^{2n} \mid n \geq 0 \}$ is a language not regular.

Proof

1. Given the homomorphism

$$h(0) = 0, \quad h(1) = 11,$$

then

$$h^{-1}(L) = \{ 0^n 1^n \mid n \geq 0 \}.$$

2. Since $h^{-1}(L)$ is not regular, then L is not regular. ■

Some Exercises

Exercise 4.2.2

If L is a language, and a is a symbol, then L/a , the quotient of L and a , is the set of strings w such that wa is in L . For example, if $L = \{a, aab, baa\}$, then $L/a = \{\varepsilon, ba\}$. Prove that if L is regular, so is L/a . *Hint:* Start with a DFA for L and consider the set of accepting states.

Some Exercises

Exercise 4.2.2

If L is a language, and a is a symbol, then L/a , the quotient of L and a , is the set of strings w such that wa is in L . For example, if $L = \{a, aab, baa\}$, then $L/a = \{\varepsilon, ba\}$. Prove that if L is regular, so is L/a . *Hint:* Start with a DFA for L and consider the set of accepting states.

Proof (Hopcroft, Motwani and Ullman [(1979) 2007] solution)

Start with a DFA A for L . Construct a new DFA B , that is exactly the same as A , except that state q is an accepting state of B if and only if $\delta(q, a)$ is an accepting state of A . Then B accepts input string w if and only if A accepts wa ; that is, $L(B) = L/a$. ■

Closure Properties

Exercise 4.2.3

If L is a language, and a is a symbol, then $a \backslash L$ is the set of strings w such that aw is in L . For example, if $L = \{a, aab, baa\}$, then $a \backslash L = \{\varepsilon, ab\}$. Prove that if L is regular, so is $a \backslash L$. *Hint:* Start with a DFA for L and consider its start state.

Closure Properties

Exercise 4.2.3

If L is a language, and a is a symbol, then $a \backslash L$ is the set of strings w such that aw is in L . For example, if $L = \{a, aab, baa\}$, then $a \backslash L = \{\varepsilon, ab\}$. Prove that if L is regular, so is $a \backslash L$. *Hint:* Start with a DFA for L and consider its start state.

Proof (Hopcroft, Motwani and Ullman [(1979) 2007] solution)

Start with a DFA A for L . Construct a new DFA B , that is exactly the same as A , except that its start state is $\delta(q_0, a)$ where q_0 is the start state of A . Then B accepts input string w if and only if A accepts aw ; that is, $L(B) = L \backslash a$. ■

Closure Properties

Exercise 4.2.13.b

We can use closure properties to help prove certain languages are not regular. Start with the fact that the language

$$L_{0n1n} = \{0^n 1^n \mid n \geq 0\}$$

is not a regular set. Prove that the following language not to be regular by transforming it, using operations known to preserve regularity, to L_{0n1n} :

$$L = \{0^n 1^m 2^{n-m} \mid n \geq m \geq 0\}.$$

References



Birkhoff, G. (1946). Universal Algebra. In: Comptes Rendus du Premier Congrès Canadien de Mathématiques. University of Toronto Press, pp. 310–326 (cit. on pp. 26–28).



— (1987). Universal Algebra. In: Selected Papers on Algebra and Topology by Garrett Birkhoff. Ed. by Rota, G.-C. and Oliveira, J. S. Birkhäuser, pp. 146–162 (cit. on pp. 26–28).



Cohn, P. M. [1965] (1981). Universal Algebra. Revised edition. Vol. 6. Mathematics and Its Applications. D. Reidel Publishing Company (cit. on p. 33).



Hopcroft, J. E., Motwani, R. and Ullman, J. D. [1979] (2007). Introduction to Automata Theory, Languages, and Computation. 3rd ed. Pearson Education (cit. on pp. 2, 9, 36, 51, 52, 59–62).