On the centralizer of a language

Paolo Massazza[1]

[1]Dipartimento di Informatica e Comunicazione
Università degli Studi dell'Insubria
Varese Italy

Varese July 18th 2006

Outline

Outline

Introduction

Goal: to study particular cases of the language equation

$$XL = LX, \quad L \subseteq \Sigma^\star, \Sigma = \{\sigma_1, \ldots, \sigma_n\}.$$

Definition (Centralizer)

$\mathcal{C}(L)$ is the largest solution of $XL = LX$.

Problem (Conway 1971)

Is it true that if L is rational then $\mathcal{C}(L)$ is rational?

Our Result: some sufficient conditions for $\mathcal{C}(L) = L^\star$.

Introduction

Goal: to study particular cases of the language equation

$$XL = LX, \quad L \subseteq \Sigma^\star, \Sigma = \{\sigma_1, \dots, \sigma_n\}.$$

Definition (Centralizer)

$\mathcal{C}(L)$ is the largest solution of $XL = LX$.

Problem (Conway 1971)

Is it true that if L is rational then $\mathcal{C}(L)$ is rational?

Our Result: some sufficient conditions for $\mathcal{C}(L) = L^\star$.

Introduction

Goal: to study particular cases of the language equation

$$XL = LX, \quad L \subseteq \Sigma^\star, \Sigma = \{\sigma_1, \ldots, \sigma_n\}.$$

Definition (Centralizer)

$\mathcal{C}(L)$ is the largest solution of $XL = LX$.

Problem (Conway 1971)

Is it true that if L is rational then $\mathcal{C}(L)$ is rational?

Our Result: some sufficient conditions for $\mathcal{C}(L) = L^\star$.

Introduction

Goal: to study particular cases of the language equation

$$XL = LX, \quad L \subseteq \Sigma^\star, \Sigma = \{\sigma_1, \ldots, \sigma_n\}.$$

Definition (Centralizer)

$\mathcal{C}(L)$ is the largest solution of $XL = LX$.

Problem (Conway 1971)

Is it true that if L is rational then $\mathcal{C}(L)$ is rational?

Our Result: some sufficient conditions for $\mathcal{C}(L) = L^\star$.

Known results

About codes

- If L is a prefix code then $\mathcal{C}(L) = \rho(L)^{\star}$ (Ratoandromanana 1989).
- If L is a rational code then $\mathcal{C}(L)$ is rational (Karhumäki, Latteux and Petre 2005).

About finite sets

- If L has at most 3 words then $\mathcal{C}(L)$ is rational (Karhumäki et al. 2002).
- There exists a finite language L such that $\mathcal{C}(L)$ is not recursively enumerable (Kunc 2005).

Known results

About codes

- If L is a prefix code then $\mathcal{C}(L) = \rho(L)^\star$ (Ratoandromanana 1989).
- If L is a rational code then $\mathcal{C}(L)$ is rational (Karhumäki, Latteux and Petre 2005).

About finite sets

- If L has at most 3 words then $\mathcal{C}(L)$ is rational (Karhumäki et al. 2002).
- There exists a finite language L such that $\mathcal{C}(L)$ is not recursively enumerable (Kunc 2005).

Outline

1 Introduction

2 Preliminary notions
- Centralizer
- Words and languages

3 The result
- Theorems and corollaries
- Conclusions and open problems

The centralizer: elementary properties

Properties of $\mathcal{C}(L)$

1. If $A \subseteq \Sigma^\star$ commutes with $L \subseteq \Sigma^\star$ then $A \subseteq \mathcal{C}(L)$
2. $L^\star \subseteq \mathcal{C}(L) \subseteq \mathrm{Pref}(L^\star) \cap \mathrm{Suf}(L^\star)$
3. $\mathcal{C}(L)$ is a monoid
4. If $\epsilon \in L$ then $\mathcal{C}(L) = \Sigma^\star$

The centralizer: elementary properties

Properties of $\mathcal{C}(L)$

1. If $A \subseteq \Sigma^\star$ commutes with $L \subseteq \Sigma^\star$ then $A \subseteq \mathcal{C}(L)$
2. $L^\star \subseteq \mathcal{C}(L) \subseteq \mathrm{Pref}(L^\star) \cap \mathrm{Suf}(L^\star)$
3. $\mathcal{C}(L)$ is a monoid
4. If $\epsilon \in L$ then $\mathcal{C}(L) = \Sigma^\star$

The centralizer: elementary properties

Properties of $\mathcal{C}(L)$

1. If $A \subseteq \Sigma^\star$ commutes with $L \subseteq \Sigma^\star$ then $A \subseteq \mathcal{C}(L)$
2. $L^\star \subseteq \mathcal{C}(L) \subseteq \mathrm{Pref}(L^\star) \cap \mathrm{Suf}(L^\star)$
3. $\mathcal{C}(L)$ is a monoid
4. If $\epsilon \in L$ then $\mathcal{C}(L) = \Sigma^\star$

The centralizer: elementary properties

Properties of $\mathcal{C}(L)$

1. If $A \subseteq \Sigma^\star$ commutes with $L \subseteq \Sigma^\star$ then $A \subseteq \mathcal{C}(L)$
2. $L^\star \subseteq \mathcal{C}(L) \subseteq \mathrm{Pref}(L^\star) \cap \mathrm{Suf}(L^\star)$
3. $\mathcal{C}(L)$ is a monoid
4. If $\epsilon \in L$ then $\mathcal{C}(L) = \Sigma^\star$

Basic notions: words

- x is a prefix of y ($x \leq y$) iff $y = xz$
- x is prefix distinguishable in L iff for $y \in L$, $y \neq x$, $x \not\leq y$ and $y \not\leq x$
- u is a root of w if $w = u^r$
- w is primitive if $w = u^r$ implies $w = u$ and $r = 1$
- a root which is primitive is called primitive root
- every w admits exactly one primitive root $\rho(w)$
- $w \in L$ is root prefix distinguishable in L if $\rho(w)$ is prefix distinguishable in $L \setminus \{w\}$
- $x \leq_{\text{lex}} y$ iff $x \leq y$ or $x = \alpha\sigma u$ and $y = \alpha\tau v$ with $\sigma < \tau$

Basic notions: languages

- R is a root of L if $L = R^e$
- a root R is minimal if $R \neq Y^k$ $(Y \neq R)$
- R is the primitive root of L, $R = \rho(L)$, if R is the only minimal root of L
- L is periodic if $L \subseteq u^\star$
- L is branching if $\exists v, w \in L$ such that $\mathrm{pref}_1(v) \neq \mathrm{pref}_1(v)$
- The circular shift of a nonbranching language $L = aL_1$ is $L^{\hookrightarrow} = L_1 a$

Branching languages: some results

Theorem

Let $L \subseteq \Sigma^+$ be a nonperiodic language. Then there is a branching language $\hat{L} \subseteq \Sigma^+$ such that

$$\mathcal{C}(L) = L^\star \iff \mathcal{C}(\hat{L}) = \hat{L}^\star.$$

Lemma

Let $L \subseteq \Sigma^+$ be nonperiodic and nonbranching, $L = aL_1$. Then

$$\mathcal{C}(L) = (a\mathcal{C}(L^{\hookrightarrow}))a^{-1}.$$

Branching languages: some results

Theorem

Let $L \subseteq \Sigma^+$ be a nonperiodic language. Then there is a branching language $\hat{L} \subseteq \Sigma^+$ such that

$$\mathcal{C}(L) = L^\star \qquad \Longleftrightarrow \qquad \mathcal{C}(\hat{L}) = \hat{L}^\star.$$

Lemma

Let $L \subseteq \Sigma^+$ be nonperiodic and nonbranching, $L = aL_1$. Then

$$\mathcal{C}(L) = (a\mathcal{C}(L^{\hookrightarrow}))a^{-1}.$$

Outline

Main theorems

Theorem

Let $L \subseteq \Sigma^+$ be a language such that $\sharp L > 1$ and $u = \min_{lex}(L)$ is root prefix distinguishable in L. Then $\mathcal{C}(L) = L^\star$.

This generalizes a previous result:

Theorem

Let $L \subseteq \Sigma^+$ be a language such that $u = \min_{lex}(L)$ is primitive and prefix distinguishable in L. Then $\mathcal{C}(L) = L^\star$.

Main theorems

Theorem

Let $L \subseteq \Sigma^+$ be a language such that $\sharp L > 1$ and $u = \min_{lex}(L)$ is root prefix distinguishable in L. Then $\mathcal{C}(L) = L^\star$.

This generalizes a previous result:

Theorem

Let $L \subseteq \Sigma^+$ be a language such that $u = \min_{lex}(L)$ is primitive and prefix distinguishable in L. Then $\mathcal{C}(L) = L^\star$.

Corollaries

Corollary

Let $L \subseteq \Sigma^+$ ($\sharp L > 1$). If there is $w \in L$ with $pref_1(w) \neq pref_1(y)$ for any $y \in L \setminus \{w\}$, then $\mathcal{C}(L) = L^\star$.

Corollary

Let $L \subseteq \Sigma^+$ be a three-word language which is not periodic. Then $\mathcal{C}(L) = L^\star$.

Corollaries

Corollary

Let $L \subseteq \Sigma^+$ ($\sharp L > 1$). If there is $w \in L$ with $pref_1(w) \neq pref_1(y)$ for any $y \in L \setminus \{w\}$, then $\mathcal{C}(L) = L^\star$.

Corollary

Let $L \subseteq \Sigma^+$ be a three-word language which is not periodic. Then $\mathcal{C}(L) = L^\star$.

Conclusions

Fact

The conditions we have shown are not not necessary for

$$\mathcal{C}(L) = L^{\star}$$

Example

1. $L = \{a, aba, ababa\}$.
   $\mathcal{C}(L) = a\{a, baa, babaa\}^{\star}a^{-1} = \{a, aba, ababa\}^{\star} = L^{\star}$
2. $L = \{a^i ba^i, b^i ab^i \mid i > 0\}$. $\min_{\text{lex}}(L) = \perp$, L is branching. Since L is a prefix code and primitive $\mathcal{C}(L) = L^{\star}$.
3. $L = \{aaa, bbb, ab, ba\}$. $\min_{\text{lex}}(L)$ is not root prefix distinguishable, L is branching. Nevertheless, $\mathcal{C}(L) = L^{\star}$ (L is a prefix code and primitive).

Conclusions

Fact

The conditions we have shown are not not necessary for

$$\mathcal{C}(L) = L^\star$$

Example

1. $L = \{a, aba, ababa\}$.
   $\mathcal{C}(L) = a\{a, baa, babaa\}^\star a^{-1} = \{a, aba, ababa\}^\star = L^\star$
2. $L = \{a^i ba^i, b^i ab^i \mid i > 0\}$. $\min_{\text{lex}}(L) = \perp$, L is branching. Since L is a prefix code and primitive $\mathcal{C}(L) = L^\star$.
3. $L = \{aaa, bbb, ab, ba\}$. $\min_{\text{lex}}(L)$ is not root prefix distinguishable, L is branching. Nevertheless, $\mathcal{C}(L) = L^\star$ (L is a prefix code and primitive).

Conclusions

Fact

The conditions we have shown are not not necessary for

$$\mathcal{C}(L) = L^\star$$

Example

1. $L = \{a, aba, ababa\}$.
   $\mathcal{C}(L) = a\{a, baa, babaa\}^\star a^{-1} = \{a, aba, ababa\}^\star = L^\star$
2. $L = \{a^i ba^i, b^i ab^i \mid i > 0\}$. $\min_{\text{lex}}(L) = \perp$, L is branching. Since L is a prefix code and primitive $\mathcal{C}(L) = L^\star$.
3. $L = \{aaa, bbb, ab, ba\}$. $\min_{\text{lex}}(L)$ is not root prefix distinguishable, L is branching. Nevertheless, $\mathcal{C}(L) = L^\star$ (L is a prefix code and primitive).

Conclusions

Fact

The conditions we have shown are not not necessary for

$$\mathcal{C}(L) = L^\star$$

Example

1. $L = \{a, aba, ababa\}$.
   $\mathcal{C}(L) = a\{a, baa, babaa\}^\star a^{-1} = \{a, aba, ababa\}^\star = L^\star$
2. $L = \{a^i ba^i, b^i ab^i \mid i > 0\}$. $\min_{\text{lex}}(L) = \perp$, L is branching.
   Since L is a prefix code and primitive $\mathcal{C}(L) = L^\star$.
3. $L = \{aaa, bbb, ab, ba\}$. $\min_{\text{lex}}(L)$ is not root prefix distinguishable, L is branching. Nevertheless, $\mathcal{C}(L) = L^\star$ (L is a prefix code and primitive).

Conclusions

Fact

The conditions we have shown are not not necessary for

$$\mathcal{C}(L) = L^\star$$

Example

1. $L = \{a, aba, ababa\}$.
   $\mathcal{C}(L) = a\{a, baa, babaa\}^\star a^{-1} = \{a, aba, ababa\}^\star = L^\star$
2. $L = \{a^i ba^i, b^i ab^i \mid i > 0\}$. $\min_{\text{lex}}(L) = \perp$, L is branching. Since L is a prefix code and primitive $\mathcal{C}(L) = L^\star$.
3. $L = \{aaa, bbb, ab, ba\}$. $\min_{\text{lex}}(L)$ is not root prefix distinguishable, L is branching. Nevertheless, $\mathcal{C}(L) = L^\star$ (L is a prefix code and primitive).

Conclusions

Fact

The conditions we have shown are not not necessary for

$$\mathcal{C}(L) = L^\star$$

Example

1. $L = \{a, aba, ababa\}$.
   $\mathcal{C}(L) = a\{a, baa, babaa\}^\star a^{-1} = \{a, aba, ababa\}^\star = L^\star$
2. $L = \{a^i ba^i, b^i ab^i \mid i > 0\}$. $\min_{\text{lex}}(L) = \perp$, L is branching. Since L is a prefix code and primitive $\mathcal{C}(L) = L^\star$.
3. $L = \{aaa, bbb, ab, ba\}$. $\min_{\text{lex}}(L)$ is not root prefix distinguishable, L is branching. Nevertheless, $\mathcal{C}(L) = L^\star$ (L is a prefix code and primitive).

Conclusions

Fact

The conditions we have shown are not not necessary for

$$\mathcal{C}(L) = L^\star$$

Example

1. $L = \{a, aba, ababa\}$.
   $\mathcal{C}(L) = a\{a, baa, babaa\}^\star a^{-1} = \{a, aba, ababa\}^\star = L^\star$
2. $L = \{a^i ba^i, b^i ab^i \mid i > 0\}$. $\min_{\text{lex}}(L) = \perp$, L is branching. Since L is a prefix code and primitive $\mathcal{C}(L) = L^\star$.
3. $L = \{aaa, bbb, ab, ba\}$. $\min_{\text{lex}}(L)$ is not root prefix distinguishable, L is branching. Nevertheless, $\mathcal{C}(L) = L^\star$ (L is a prefix code and primitive).

Open problems

Problems

- find weaker sufficient conditions
- study the centralizer of a 4-word language
- characterize the languages L such that $\mathcal{C}(L) = L^\star$

Example

$L = \{aa, ab, ba, bb\}$ has a root prefix distinguishable word ab but $\mathcal{C}(L) = \rho(L)^\star$ with $\rho(L) = \{a, b\}$

Open problems

Problems

- find weaker sufficient conditions
- study the centralizer of a 4-word language
- characterize the languages L such that $\mathcal{C}(L) = L^{\star}$

Example

$L = \{aa, ab, ba, bb\}$ has a root prefix distinguishable word ab but $\mathcal{C}(L) = \rho(L)^{\star}$ with $\rho(L) = \{a, b\}$

Open problems

Problems

- find weaker sufficient conditions
- study the centralizer of a 4-word language
- characterize the languages L such that $\mathcal{C}(L) = L^\star$

Example

$L = \{aa, ab, ba, bb\}$ has a root prefix distinguishable word ab but $\mathcal{C}(L) = \rho(L)^\star$ with $\rho(L) = \{a, b\}$

Open problems

Problems

- find weaker sufficient conditions
- study the centralizer of a 4-word language
- characterize the languages L such that $\mathcal{C}(L) = L^\star$

Example

$L = \{aa, ab, ba, bb\}$ has a root prefix distinguishable word ab but $\mathcal{C}(L) = \rho(L)^\star$ with $\rho(L) = \{a, b\}$

Open problems

Problems

- find weaker sufficient conditions
- study the centralizer of a 4-word language
- characterize the languages L such that $\mathcal{C}(L) = L^\star$

Example

$L = \{aa, ab, ba, bb\}$ has a root prefix distinguishable word *ab* but $\mathcal{C}(L) = \rho(L)^\star$ with $\rho(L) = \{a, b\}$

Papers

- P. Massazza. On the equation $XL = LX$,
  Proc. of WORDS 2005.
  Publications du Laboratoire de Combinatoire et d'Informatique Mathématique, Montreal, 36 (2005), p. 315–322.
- P. Massazza, P. Salmela.
  On the simplest centralizer of a language.
  to appear in RAIRO Theoretical Informatics and Applications.