Theoretical Computer Science 635 (2016) 85–93

Contents lists available at ScienceDirect

Theoretical Computer Science

www.elsevier.com/locate/tcs

Conjugacy relations of prefix codes

Yong He *, Zhenhe Cui, Zihan Yuan

School of Computer Science and Engineering, Hunan University of Science and Technology, Xiangtan 411201, Hunan, China

ARTICLE INFO

Article history:
Received 27 October 2014
Received in revised form 27 April 2016
Accepted 16 May 2016
Available online 18 May 2016
Communicated by J. Karhumäki

Keywords:
Conjugacy equation
Conjugacy relation
Initial part
Prefix code
Conjugacy problem

ABSTRACT

It is shown that, if X and Y are prefix codes and Z is a non-empty language satisfying the condition $XZ = ZY$, then Z is the union of a non-empty family $\{P_n\}_{i\in I}$ of pairwise disjoint prefix sets such that $XP_i = P_iY$ for all $i \in I$. Consequently, the conjugacy relations of prefix codes are explored and, under the restriction that both of X and Y are prefix codes, the solutions of the conjugacy equation $XZ = ZY$ for languages are determined. Also, the decidability of the conjugacy problem for finite prefix codes is confirmed.

1. Introduction

In the investigation of algebraic structures, equation is usually an interesting topic. On an algebraic structure G equipped with a binary operation, the *commutativity equation* $xy = yx$ is always a basic equation. For any $x, y \in G$, let

$$\mathscr{C}(x, y) = \{z \in G | xz = zy\}.$$

Then $\mathscr{C}(x, x)$ is exactly the centralizer of x. Therefore, the so called *conjugacy equation* $xz = zy$ is an expansion of the commutativity equation.

The conjugacy relation on a group induced by the conjugacy equation has played a fundamental role in the development of group theory. This motivates the consideration on the following analogous relations on an arbitrary monoid M:

the *right conjugacy relation*:	$\overset{r}{\sim} = \{(x, y) | (\exists z \in M)\, xz = zy,\ z \neq 0\}$,
the *left conjugacy relation*:	$\overset{l}{\sim} = \{(x, y) | (\exists z \in M)\, zx = yz,\ z \neq 0\}$,
the *two-sided conjugacy relation*:	$\overset{c}{\sim} = \overset{r}{\sim} \cap \overset{l}{\sim}$,
the *transposition relation*:	$\overset{t}{\sim} = \{(x, y) | (\exists w, u \in M)\, x = wu,\ y = uw\}$,
the *group conjugacy relation*:	$\overset{g}{\sim} = \{(x, y) | (\exists g \in U)\, xg = gy\}$,

* Corresponding author.
E-mail addresses: ynghe@263.net (Y. He), 2857727938@qq.com (Z. Cui), yuanzihan2000@163.com (Z. Yuan).

http://dx.doi.org/10.1016/j.tcs.2016.05.021
0304-3975/

where 0 is the possible zero element of M, and U the group of units of M. Together with the transitive closure $\overset{t^*}{\sim}$ of the transposition relation $\overset{t}{\sim}$, the relations on the monoid M defined as above are uniformly called the *conjugacy relations*. It is pointed out in [5] that the relations $\overset{r}{\sim}$ and $\overset{l}{\sim}$ are reflexive, transitive and mutually inverse, the relation $\overset{t}{\sim}$ is reflexive and symmetric, the relations $\overset{g}{\sim}$, $\overset{t^*}{\sim}$, $\overset{c}{\sim}$ are equivalences and, in general,

$$\overset{g}{\sim} \subseteq \overset{t}{\sim} \subseteq \overset{t^*}{\sim} \subseteq \overset{c}{\sim} \subseteq \overset{r}{\sim}, \overset{l}{\sim}. \tag{1}$$

The conjugacy relations on free monoids and free inverse monoids have been considered in [12] and [4], respectively, while the applications of conjugacy relations in the decomposition and representation of monoids are explored in [1].

Observe that the set $\mathbb{L}(A)$ of languages on an alphabet A forms a monoid with respect to concatenation. The investigation of the conjugacy equation and conjugacy relations on $\mathbb{L}(A)$ is initiated by Perrin who considered the $\overset{t}{\sim}$-relation of codes [14]. It is observed that the relations $\overset{t}{\sim}$, $\overset{c}{\sim}$, $\overset{r}{\sim}$ and $\overset{l}{\sim}$ on $\mathbb{L}(A)$ are pairwise distinct when A contains at least two letters [2]. The solutions of the conjugacy equation $XZ = ZY$ on $\mathbb{L}(A)$ satisfying the following additional conditions are determined or partially determined in [2,3,7,10]:

(i) $|Z| = 1$;
(ii) Z is a biprefix code;
(iii) X, Y, Z are prefix codes;
(iv) $|X| + |Y| \leq 5$;
(v) X, Y are finite biprefix codes (esp., uniform codes).

The aim of the present paper is to characterize the conjugacy relations of prefix codes, determine the solutions of the conjugacy equation $XZ = ZY$ under the restriction that X and Y are prefix codes, and then consider the conjugacy problem for finite prefix codes. The technique used in this paper is hungered by the authors of [3] when they investigated the conjugacy of finite biprefix codes. For the terminologies and notations without explanation, the reader is referred to [1,17].

2. Preliminaries

In what follows, all words and languages considered are on a given alphabet A. The free monoid, the free semigroup and the set of languages on A are denoted by A^*, A^+ and $\mathbb{L}(A)$, respectively. The empty word is written as ϵ. The length of a word w is denoted by $\lg(w)$. Then $\lg(w) = 0$ precisely when $w = \epsilon$. The group conjugacy relations $\overset{g}{\sim}$ on A^* and $\mathbb{L}(A)$ are to be out of considerations since they are trivial. The following achievement of Lentin and Schützenberger is interesting.

Lemma 2.1. *([12]) On the monoid A^*, $\overset{t}{\sim} = \overset{t^*}{\sim} = \overset{c}{\sim} = \overset{r}{\sim} = \overset{l}{\sim}$.*

A non-empty word is said to be *primitive* if it is not a power of another word. Observe that each non-empty word w is the power of a unique primitive word which is called the *primitive root* of w and denoted by $\sqrt{w}$. A description for the relation $\overset{r}{\sim}$ (and hence for the other conjugacy relations) on A^* is given as below:

Lemma 2.2. *([13]) Let x and y be two non-empty words. Then $x \overset{r}{\sim} y$ if and only if $\sqrt{x} \overset{r}{\sim} \sqrt{y}$ and $x = (\sqrt{x})^k$, $y = (\sqrt{y})^k$ for some $k \in \mathbb{N}^+$. Moreover, if this is the case, there exists a unique pair $(p, q) \in A^* \times A^+$ such that $\sqrt{x} = pq$, $\sqrt{y} = qp$ and $\mathscr{C}(x, y) = (pq)^* p$.*

Let L be an arbitrary language. The cardinality of L is denoted by $|L|$. If L is the union of a family $\{L_i\}_{i \in I}$ of pairwise disjoint languages, then we use the notation

$$L = \biguplus_{i \in I} L_i.$$

For each $n \in \mathbb{N}$, denote the set of all words of length n in L by $L^{[n]}$. It is clear that

$$L = \biguplus_{n \in \mathbb{N}} L^{[n]},$$

and hence a language K coincides with L if and only if $L^{[n]} = K^{[n]}$ for all $n \in \mathbb{N}$. For any subset I of $\mathbb{N}$, put

$$L^I = \bigcup_{i \in I} L^i.$$

Then $L^{\mathbb{N}} = L^*$ where * is the Kleene star operation on languages. The language

$$\bar{L} = L - LA^+$$