Acta Univ. M. Belii
Math no.3(1995), pp.81-87

STRONGLY IRREDUCIBLE STRINGS

Bohuslav Sivák

Abstract. The "strong irreducibility" of strings is defined and it is proved that certain special strings of 0's and 1's are strongly irreducible. This fact has found an application (see [1]) in the study of discrete dynamic systems by the methods of the symbolic dynamics.

1. Introduction

Let T be a finite set of symbols. The string over T is a finite sequence of symbols from the set T. The length of the string w is the number of symbols in the string w, this number will be denoted by $|w|$. For example, $|abac| = 4$. The empty string will be denoted by ε. Trivially, $|\varepsilon| = 0$. The concatenation of the strings u and v will be denoted by $u \cdot v$ or by uv. For example, the concatenation of the strings $u = 001$, $v = 10$ is the string $uv = 00110$. The concatenation of several identical strings will be written in the form of the formal power: $u^0 = \varepsilon$, $u^1 = u$, $u^2 = uu$, etc.

Definition. Let B be a string over T. The string B is called reducible iff it can be written in the form

$$B = W^k, \qquad k \geq 2.$$

The string B is called irreducible iff it is not reducible. The string B is called strongly irreducible iff the following two conditions are satisfied:

(1) - the string B is irreducible,
(2) - the string $A^m B$ is irreducible for every $m \geq 2$ and every irreducible string $A \neq B$.

Examples. Put $T = \{0, 1\}$. Following strings over T are reducible:

ε, 00, 11, 000, 111, 0000, 0101, 1010, 1111.

Following strings are irreducible:

0, 1, 01, 10, 001, 010, 011, 100, 101, 0010.

It is easy to see that the string 01 is strongly irreducible. The string 1001 is irreducible but not strongly irreducible. In fact,

---

1991 *Mathematics Subject Classification.* 20M35.
Key words and phrases. String, concatenation, prefix, postfix.

$$(010)^2 \cdot 1001 = (01001)^2.$$

Similarly, the string 10101011 is irreducible but not strongly irreducible. In fact,

$$(01)^3 \cdot 10101011 = (0101011)^2.$$

This example can be generalized: for every integer $m \geq 2$, the string $(10)^m \cdot 11$ is irreducible but not strongly irreducible.

Definition. Let U be a string of the length $\geq k$. The string $Pref_k(U)$ is the prefix of the length k. Similarly, the string $Postf_k(U)$ is the postfix of the length k.

Examples. The string 011010 has the following prefixes:

$$\varepsilon,\ 0,\ 01,\ 011,\ 0110,\ 01101,\ 011010.$$

The same string has the following postfixes:

$$\varepsilon,\ 0,\ 10,\ 010,\ 1010,\ 11010,\ 011010.$$

We leave to the reader the verification that the prefixes and the postfixes of the string 1111 are idenical.

Lemma 1.1. *Let D, E be arbitrary strings and let x, y be positive integers such that $D^x = E^y$. Then there exist positive integers i, j and a string Z such that*

$$D = Z^i,\ E = Z^j.$$

Remark. A common generalization of our Lemma 1.1, Lemma 1.2 and Lemma 1.3 is proved in [2].

Proof of Lemma 1.1. We can assume that the strings D, E are nonempty. Let $c = (x, y)$ be the greatest common divisor of the integers x, y. Then

$$x = c \cdot j,\ y = c \cdot i,\ (i, j) = 1.$$

By the assumption of the lemma,

$$D^{c \cdot j} = E^{c \cdot i},$$
$$c \cdot j \cdot |D| = c \cdot i \cdot |E|,$$
$$j \cdot |D| = i \cdot |E|.$$

By the last equality, there exists a positive integer h such that

$$|D| = h \cdot i, \quad |E| = h \cdot j.$$

The string $D^j = E^i$ can be uniquely written in the form

$$D^j = E^i = Z_1 \cdot Z_2 \ldots Z_{i \cdot j},$$
$$|Z_1| = |Z_2| = \cdots = |Z_{i \cdot j}| = h.$$

By these equalities,

$$D = Z_1 \ldots Z_i,$$
$$D = Z_{i+1} \ldots Z_{2 \cdot i},$$
$$\ldots\ldots\ldots$$
$$D = Z_{i \cdot (j-1)+1} \ldots Z_{i \cdot j},$$

Consequently, $Z_p = Z_q$ whenever the difference of the indexes p, q is a multiple of i. Similarly,

$$E = Z_1 \dots Z_j,$$
$$E = Z_{j+1} \dots Z_{2\cdot j},$$
$$\dots\dots\dots$$
$$E = Z_{(i-1)\cdot j+1} \dots Z_{i\cdot j}$$

and $Z_p = Z_q$ whenever the difference of the indexes p, q is a multiple of j. We know that $(i, j) = 1$. It follows that all of the strings $Z_1,\ Z_2, \dots, Z_{i\cdot j}$ are identical.

Lemma 1.2. *Let D, E be irreducible strings and let x, y be positive integers such that $D^x = E^y$. Then*

$$x = y, \quad D = E.$$

Proof. It suffices to apply Lemma 1.1.

Lemma 1.3. *Let D, E be irreducible strings and let x, y be positive integers such that $D^x \cdot E^y = E^y \cdot D^x$. Then $D = E$.*

Proof. Suppose the assertion of this lemma is false. Then we can choose a counterexample (D, E, x, y) such that the length of the string $D^x \cdot E^y$ is minimal. We can assume the inequality $|D| < |E|$ in this counterexample. Several powers of the string D can be prefixes of the string E (trivially, D^0 is a prefix of the string E). Let q be the maximal integer such that the string D^q is a prefix of the string E. Then we can write E in the form

$$E = D^q \cdot F, \qquad D \text{ is not a prefix of } F.$$

Substituting the last equation into the assumption of lemma, we obtain

$$D^x \cdot (D^q \cdot F)^y = (D^q \cdot F)^y \cdot D^x.$$

Put $W = (D^q \cdot F)^{y-1}$. Then we can write

$$D^{x+q} \cdot F \cdot W = D^q \cdot F \cdot W \cdot D^x,$$
$$D^x \cdot (F \cdot W) = (F \cdot W) \cdot D^x.$$

The string D is not a prefix of F, and so the string F is a prefix of D. Now it is easy to check that $|F \cdot W| < |E^y|$. By our assumption of the minimality, the string $F \cdot W$ is a power of the string D:

$$F \cdot (D^q \cdot F)^{y-1} = D^z, \quad z \geq 1,$$
$$E^y = (D^q \cdot F)^y = D^q \cdot F \cdot W = D^{q+z},$$

contrary to Lemma 1.2.

Lemma 1.4. *Let D, E be arbitrary strings such that the string DE is irreducible and $DE = ED$. Then exactly one of the strings D, E is empty.*

Proof. It suffices to apply Lemma 1.3.

2. The fundamental theorem

Lemma 2.1. *Let A, C be irreducible strings and let B be arbitrary string such that*

$$A^m \cdot B = C^k, \quad m \geq 2, \quad k \geq 2, \quad |A| < |C| < m \cdot |A|.$$

Then there exist non-empty strings F, G and a non-negative integer s such that

$$A = (FG)^{s+1} \cdot F, \quad C = A^{m-1} \cdot FG,$$
$$B = GF \cdot A^{m-2} \cdot FG \cdot C^{k-2}.$$

Proof. The string A is a prefix of the string C and the string C is a prefix of the string A^m. Therefore we can write the string C in the following form:

$$C = A^r \cdot D, \quad 0 < |D| < |A|, \quad 1 \leq r < m.$$

(The equality $D = \varepsilon$ would contradict the irreducibility of C.) Substituting the equality $C = A^r \cdot D$ into the assumption of lemma, we obtain

$$A^m \cdot B = (A^r \cdot D)^k,$$
$$A^m \cdot B = A^r \cdot D \cdot (A^r \cdot D)^{k-1},$$
$$A^{m-r} \cdot B = D \cdot (A^r \cdot D)^{k-1}.$$

It follows immediately that the string D is a prefix of A:

$$A = D \cdot E, \quad |E| > 0,$$
$$DE \cdot (DE)^{m-r-1} \cdot B = D \cdot (DE)^r \cdot D^{k-1},$$
$$E \cdot (DE)^{m-r-1} \cdot B = (DE)^r \cdot D^{k-1},$$

The inequality $m - r - 1 > 0$ would contradict Lemma 1.4. It follows that $r = m - 1$ and

$$E \cdot B = (DE)^{m-1} \cdot D^{k-1}.$$

Let s be the maximal integer such that the string D^s is a prefix of the string E. Then there exists a string F such that

$$E = D^s \cdot F, \quad D \text{ is not a prefix of } F.$$

The strings A, C, $E \cdot B$ can be written as follows:

$$A = D \cdot E = D^{s+1} \cdot F,$$
$$C = A^{m-1} \cdot D,$$
$$D^s \cdot F \cdot B = ((D^{s+1} \cdot F)^{m-1} \cdot D)^{k-1}.$$

From this we conclude that

$$F \cdot B = DF \cdot (D^{s+1} \cdot F)^{m-2} \cdot D \cdot ((D^{s+1} \cdot F)^{m-1} \cdot D)^{k-2}.$$

We know that the string D is not a prefix of F, and so the string F is a (proper) prefix of D:

$$D = F \cdot G, \quad 0 < |G| < |D|.$$

Substituting this equality into the preceding ones, we complete the proof.

Lemma 2.2. *Let A, C be irreducible strings and let B be arbitrary string such that*

$$A^m \cdot B = C^k, \quad m \geq 2, \quad k \geq 2, \quad |C| > m \cdot |A|.$$

Then there exists a non-empty string D such that

$$C = A^m \cdot D, \quad B = D \cdot C^{k-1}.$$

Proof. The string A^m is a prefix of the string C. It follows that there exists a string D such that $C = A^m \cdot D$. (The string D is non-empty because C is irreducible.) Substituting this equation into the assumption of lemma, we obtain

$$A^m \cdot B = (A^m \cdot D)^k,$$
$$A^m \cdot B = A^m \cdot D \cdot (A^m \cdot D)^{k-1},$$
$$B = D \cdot (A^m \cdot D)^{k-1} = D \cdot C^{k-1}.$$

Theorem 2.1. *Let A, C be irreducible strings and let B be arbitrary string such that*

$$A^m \cdot B = C^k, \quad m \geq 2, \quad k \geq 2.$$

Then there is satisfied exactly one of the following three conditions:

(1) $C = A$, $B = A^{k-m}$,

(2) *there exist non-empty strings F, G and a non-negative integer s such that*

$$A = (FG)^{s+1} \cdot F, \quad C = A^{m-1} \cdot FG,$$
$$B = GF \cdot A^{m-2} \cdot FG \cdot C^{k-2},$$

(3) *there exists a non-empty string D such that*

$$C = A^m \cdot D, \quad B = D \cdot C^{k-1}.$$

Proof. According to Lemma 2.1 and Lemma 2.2, we can suppose that

$$0 < |C| < |A|.$$

The string A is a prefix of the string C^k and so it can be written in the form

$$A = C^r \cdot D, \quad r \geq 1, \quad 0 < |D| < |C|.$$

Substituting this equality into the assumption of the lemma, we obtain

$$(C^r \cdot D)^m \cdot B = C^k,$$
$$C^r \cdot D \cdot (C^r \cdot D)^{m-1} \cdot B = C^k,$$
$$D \cdot (C^r \cdot D)^{m-1} \cdot B = C^{k-r}.$$

Therefore the string D is a proper prefix of the string C:

$$C = D \cdot E, \quad |E| > 0,$$
$$D \cdot ((DE)^r \cdot D)^{m-1} \cdot B = (DE)^{k-r},$$
$$D \cdot ((DE)^r \cdot D)^{m-1} \cdot B = DE \cdot (DE)^{k-r-1},$$
$$((DE)^r \cdot D)^{m-1} \cdot B = E \cdot (DE)^{k-r-1}.$$

Applying the prefix of the length $|DE|$, we obtain $DE = ED$, contrary to Lemma 1.4.

Theorem 2.2. *Let the string B be irreducible but not strongly irreducible. Then B can be written in the form*

$$B = G \cdot F \cdot H \cdot F \cdot G,$$

where H is an arbitrary string and F, G are non-empty strings.

Proof. Apply Theorem 2.1.

Corollary. *Every string of the length 3 containing at least two different symbols is strongly irreducible.*

Remark. The condition in Theorem 2.2 is necessary but not sufficient. For example, the string 10101 is strongly irreducible and the string 11111 is not irreducible.

3. Applications to concrete strings

Put $T = \{0, 1\}$. Every non-negative integer $j < 2^n$ can be uniquely written in the form of a string over T of the length n. This string will be denoted by $Cod(n, j)$. For $z \in T$, put

$$B(n, z) = Cod(n, 0) \cdot Cod(n, 1) \ldots Cod(n, 2^n - 1) \cdot z.$$

Example. Put $n = 3$. Then

$$\begin{aligned} Cod(3,0) &= 000, \quad Cod(3,1) = 001,\\ Cod(3,2) &= 010, \quad Cod(3,3) = 011,\\ Cod(3,4) &= 100, \quad Cod(3,5) = 101,\\ Cod(3,6) &= 110, \quad Cod(3,7) = 111,\\ B(3,0) &= 0000010100111001011101110,\\ B(3,1) &= 0000010100111001011101111. \end{aligned}$$

Lemma 3.1. *The string $B(n, z)$ is irreducible.*

Proof. For any string w over T, the number of the occurences of the symbols x in w is usually denoted by $\#_x(w)$. It is evident that

$$|\#_0(B(n, z)) - \#_1(B(n, z))| = 1.$$

If the string $B(n, z)$ would not be irreducible, we could write

$$B(n, z) = K^m, \quad m \geq 2$$

and both numbers $\#_0(B(n, z))$, $\#_1(B(n, z))$ would be multiples of m, a contradiction.

Lemma 3.2. *The string $B(n, z)$ can not be written in the form*

$$B(n, z) = G \cdot F \cdot H \cdot F \cdot G,$$

where H is an arbitrary string and F, G are non-empty strings.

Proof. Suppose, contrary to our claim, that the string $B(n,z)$ can be written in the form

$$B(n,z) = G \cdot F \cdot H \cdot F \cdot G,$$

where H is an arbitrary string and F, G are non-empty strings. Let us denote

$$d = |GF| = |FG|.$$

The symbol 1 occurs in the strings GF and FG. Consequently, $d \geq 2n$. Moreover, it is obvious that

$$Pref_{2n}(B(n,z)) = 0^{2n-1} \cdot 1,$$
$$Postf_{2n+1}(B(n,z)) \in \{1^{n-1}01^n0, 1^{n-1}01^{n+1}\}$$

and the string 0^{2n-1} has only one occurence in $B(n,z)$ - in the role of the prefix. The rest of this proof is left to the reader.

Theorem 3.1. *For every positive integer n and every $z \in \{0,1\}$, the string $B(n,z)$ is strongly irreducible.*

Proof. It suffices to apply Theorem 2.2, Lemma 3.1 amd Lemma 3.2.

References

[1] Ll.Alseda, S.Kolyada, J.Llibre, Ľ.Snoha, *Entropy and periodic points for transitive maps*, Preprint CRM # 305 (1995).

[2] Yu.I.Chmelevskii, *Uravnenia v svobodnoi polugruppe*, Trudy matem. instituta imeni V.A.Steklova (1971).

Department of Computer Science, Faculty of Natural Sciences

Matej Bel University, Tajovského 40, 974 01 Banská Bystrica, Slovakia

E-mail address: sivak@fhpv.umb.sk

(Received October 3, 1995)