Printing Requisition/Graphic Services

2509

Dept. No. 52079

Title or Description CS-79-16

Date 19 September/79 | **Date Required** (Tues/Wed) EARLY NEXT WEEK | **Account** 126-6454-41

Signature | **Signing Authority** Paul C

Department Computer Science | **Room** 5179 | **Phone** 3143 | **Delivery** ☒ Mail ☐ Pick-up ☐ Via Stores ☐ Other

1. Please complete unshaded areas on form as applicable. (4 part no carbon required).
2. Distribute copies as follows: White, Canary and Pink—Printing, Arts Library or applicable Copy Centre Goldenrod—**Retain.**
3. On completion of order, pink copy will be returned with printed material. Canary copy will be costed and returned to requisitioner, **Retain as a record of your charges.**
4. Please direct enquiries, quoting requisition number, to Printing/Graphic Services, Extension 3451.

Reproduction Requirements ☒ Offset ☐ Signs/Repro's ☐ Xerox

Number of Pages 10 | **Number of Copies** 50

Type of Paper Stock ☒ Bond ☐ Book ☐ Cover ☐ Bristol ☐ Supplied

NTS + Backs

Paper Size ☒ 8½ x 11 ☐ 8½ x 14 ☐ 11 x 17

Paper Colour ☒ White ☐ Other

Ink ☐ Black

Printing ☒ 1 Side ☐ 2 Sides

Numbering to

Binding/Finishing Operations ☒ Collating ☐ Corner Stitching ☐ 3 Ring ☐ Tape ☐ Plastic Ring ☐ Perforating

Folding Finished Size

Cutting Finished Size

Special Instructions

Please attache the backs and fronts to the report, and staple on top, middle and bottom.

Cost: Time/Materials	Fun.	Prod.Un.	Prod. Cl.	Opr. No.	Mins.	Total
Signs/Repro's	1					
Camera	2					
Correcting & Masking Negatives	3					
Platemaking	4					
Printing	5					
Bindery	6			21	20	
Sub. Total Time						

Film Qty	Size	Plates Qty	Size & Type	Sub. Total Materials	
Paper Qty	Size	**Plastic Rings** Qty	Size	Prov. Tax	
Outside Services				Total	

Graphic Services May 68 482-2

A NOTE ON ELEMENTARY HOMOMORPHISMS AND THE REGULARITY OF EQUALITY SETS

by

Juhani Karhumäki[+]

and

Imre Simon[*]

Research Report CS-79-16

Department of Computer Science

University of Waterloo
Waterloo, Ontario, Canada

[+] Department of Mathematics
University of Turku
Turku, Finland

[*] Instituto de Mathematica e Estatistica
Universidade de São Paulo
CP 20570
01000 São Paulo, Brazil

A NOTE ON ELEMENTARY HOMOMORPHISMS AND THE REGULARITY OF EQUALITY SETS

by

Juhani Karhumäki
Department of Mathematics
University of Turku
Turku, Finland

Imre Simon
Instituto de Mathematica e Estatistica
Universidade de São Paulo
CP 20570
01000 São Paulo
Brazil

The purpose of this note is two-fold. On the one hand we give a characterization for elementary homomorphisms (introduced in [EhRI]) by means of elementary biprefixes. On the other hand we discuss conditions which imply the regularity of the equality set of two homomorphisms.

Our characterization for elementary homomorphisms is not complete since it uses the notion of an elementary biprefix. However, it gives an alternative way to show that an elementary homomorphism is a code with bounded delay in both directions, see [EhRI] and [EhRII]. It also shows that to give a "real" characterization for elementary homomorphisms it suffices to characterize elementary biprefixes, a small subclass of all elementary homomorphisms.

An important feature of elementary homomorphisms is that their equality set is regular, see [EhRII]. Here we point out that the regularity of the equality set holds true even for a wider class of homomorphisms, namely for bounded delay homomorphisms. More precisely, we show that the equality set of two homomorphisms is regular if either both homomorphisms have bounded delay in the same direction or at least one of them has bounded delay in both directions. We also give an example showing that the equality set of two codes, i.e. injective homomorphisms, need not be even context-free.

Finally, we mention that the equality set of a periodic and an injective homomorphism is also regular, in fact a star of one word only. Moreover, this set can be effectively found.

Notations and Definitions

For a (finite) alphabet A we denote by A^* the free monoid generated by A. The cardinality of A is denoted by $|A|$. For an element u in A^* let $|u|$ be its length. For elements u and v in A^* $u \preceq v$ means that u is a prefix of v.

Following [EhRI] we say that a homomorphism $h: A^* \to B^*$ is <u>simplifiable</u> if it admits the factorization into two homomorphisms f and g:

$$\begin{tikzcd} A^* \arrow[rr, "h"] \arrow[dr, "f"'] & & B^* \\ & C^* \arrow[ur, "g"'] & \end{tikzcd}$$

with $|C| < |A|$. If h is not simplifiable then it is called <u>elementary</u>. A homomorphism $h : A^* \to B^*$ which is simplifiable via the alphabet of cardinality one, i.e. $Ah \subseteq z^*$ for some $z \in B^*$, is called <u>periodic</u>.

In our characterization of elementary homomorphisms the following two classes of homomorphisms turn out to be useful. We say that a homomorphism $h : A^* \to A^*$ is <u>atomic</u> if there exist a and a' in A such that

$$\text{either} \quad \begin{cases} ah = aa', & \\ xh = x & \text{if } x \neq a \end{cases} \qquad \text{or} \quad \begin{cases} ah = a'a, & \\ xh = x & \text{if } x \neq a. \end{cases}$$

Further $h : A^* \to A^*$ is called <u>quasiatomic</u> if it is a composition of atomic homomorphisms. It is obvious that any atomic (and hence also quasiatomic homomorphism) has bounded delay in both directions in the following sense.

We say that a homomorphism $h : A^* \to B^*$ has <u>bounded delay from left to right</u> if there exists a natural number k such that for all $u,v \in A^*$ and $a,b \in A$

$$\text{if} \quad |u| \geq k - 1 \quad \text{and} \quad (au)h \leqslant (bv)h \text{ ,} \quad \text{then} \quad a = b \text{ .}$$

The notion of a bounded delay from right to left is defined analogously. Clearly, the set of bounded delay homomorphisms in the same direction is closed under composition. It is also clear that a bounded delay homomorphism is a code. So the notion can be defined alternatively and equivalently as a code $h : A^* \to B^*$ satisfying the following property : there exists a natural number k such that for all w,u and v in A^*

$$\text{if} \quad |u| \geq k - 1 \quad \text{and} \quad (wu)h \leqslant vh \text{ ,} \quad \text{then } (wh)^{-1}(vh) \in (Ah)^*$$

where $(wh)^{-1}(vh)$ denotes the left difference of vh by wh .

Bounded delay homomorphisms from left to right (resp. from right to left) with $k = 1$ are called <u>prefixes</u> (resp. <u>suffixes</u>). By a <u>biprefix</u> we mean a homomorphism which is both a prefix and a suffix.

Finally, we define the notion of an <u>equality set</u> of two homomorphisms. For homomorphisms $h,g : A^* \to B^*$ their equality set is

$$E(h,g) = \{u \in A^* \mid h(u) = g(u)\}.$$

<u>A Characterization</u>

Here we give a characterization of elementary homomorphisms by means of elementary biprefixes.

<u>Theorem 1</u>: A homomorphism $h : A^* \to B^*$ is elementary if and only if there exists a quasiatomic homomorphism $\sigma : A^* \to A^*$ and an elementary biprefix $\pi : A^* \to B^*$ such that $h = \sigma\pi$.

<u>Proof</u>: Assume first that h is elementary. The existence of the required composition for h is proved by induction on $\rho_h = \sum_{a \in A} |ah|$.

The case $\rho_h = |A|$ is clear. So assume that $\rho_h > |A|$. If h is a biprefix there is nothing to be proved. Hence, assume that h is not a biprefix, say h is not a prefix (the other possibility being symmetric). Let $ah = (bh)u$. Define an atomic homomorphism $s : A^* \to A^*$ and a homomorphism $h' : A^* \to B^*$ by

$$\begin{aligned} as &= ba \ , \\ xs &= x \qquad \text{if } x \neq a \end{aligned}$$

and

$$\begin{aligned} ah' &= u \ , \\ xh' &= xh \qquad \text{if } x \neq a \ . \end{aligned}$$

Clearly $h = sh'$. Further h' is elementary (since h is) and $\rho_{h'} < \rho_h$. So the existence of σ and π follows from induction hypothesis.

To prove the converse it is sufficient to show that if $g : A^* \to B^*$ is elementary and $p : A^* \to A^*$ is atomic, then pg is elementary. Assume the contrary, that g is elementary but pg is not. Then pg has a factorization

$$\begin{tikzcd} A^* \arrow[rr, "pg"] \arrow[dr, "t"'] & & B^* \\ & C^* \arrow[ur, "f"'] & \end{tikzcd}$$

which $|C| < |A|$. We choose a factorization where $\rho_f = \sum_{c \in C} |cf|$ is the smallest possible. Let $|ap| = 2$. By symmetry we may assume that $ap = a'a$. Since $Apg \in (Cf)^*$ we may write

$$(*) \qquad \begin{aligned} apg &= (a'a)g = \alpha_1 \ldots \alpha_n \ , \\ a'pg &= a'g \quad\;\; = \beta_1 \ldots \beta_m \ , \end{aligned}$$

where $\alpha_1, \ldots, \alpha_n, \beta_1, \ldots, \beta_m \in Cf$. We will show that

$$(**) \qquad n \geq m \quad \text{and} \quad \alpha_i = \beta_i \quad \text{for} \quad i \leq m \ ,$$

from which the theorem follows. Indeed (*) and (**) imply that $Ag \subseteq (Cf)^*$ and hence g is not elementary, a contradiction.

Assume that (**) is not true. Then there exists $j \leq m$ such that

$$\begin{aligned} \alpha_j &\neq \beta_j \ , \\ apg &= \alpha_1 \ldots \alpha_{j-1} \alpha_j \ldots \alpha_n \ , \\ a'pg &= \alpha_1 \ldots \alpha_{j-1} \beta_j \ldots \beta_m \ . \end{aligned}$$

Since $a'pg \not< apg$ it follows that

$$\text{either} \quad \alpha_j < \beta_j \quad \text{or} \quad \beta_j < \alpha_j \ .$$

But both α_j and β_j are in Cf , so by the first part of this proof f has a factorization $f = s'f'$ with $\rho_{f'} < \rho_f$. This contradicts with the minimality of f and so (*) is true and the theorem is proved. □

As a consequence of Theorem 1 we conclude, cf. [EhRI] and [EhRII] .

<u>Corollary</u>: Any elementary homomorphism is a code with bounded delay in both directions. □

Employing the proof technique used above the following representation for elementary homomorphisms can also be proved. We leave the proof for the reader.

<u>Theorem 2</u>: Any elementary homomorphism $h : A^* \to B^*$ has

factorizations

$$\begin{tikzcd} A^* \arrow[r, "p"] \arrow[d, "s'"'] \arrow[dr, "h"] & C^* \arrow[d, "s"] \\ D^* \arrow[r, "p'"'] & B^* \end{tikzcd}$$

where p and p' are prefixes and s and s' are suffixes such that all of them has bounded delay in both directions. □

It is clear that Theorem 2 does not characterize elementary homomorphisms. Indeed, any biprefix has these factorizations but they are not all elementary.

Regularity of Equality Sets

A remarkable property of elementary homomorphisms is that their equality set is regular. This has been shown in [EhRII], and generalized in [EnR] to cover also the case when only one of the homomorphisms is elementary (or, in fact, a composition of elementary homomorphisms). In general equality sets are far from being regular, see e.g. [C].

Here we will point out that the regularity of an equality set follows also from suitable bounded delay properties. These results are direct consequences of the proof of the fact that $E(h,g)$ is regular for elementary h and g, see [EhRII]. The basic idea behind this proof is that elementary homomorphisms have "a unique continuation property" in the following sense. Assume that h and g are elementary and $u \prec v \in E(h,g)$. Then if $||uh| - |ug||$ is large enough there exists a unique letter a such that ua is a prefix of a word in $E(h,g)$. But obviously this condition is valid for homomorphisms having bounded delay from left to right. So, by symmetry, we may formulate

<u>Theorem 3</u>: Let h and g be homomorphisms with bounded delay in the same direction. Then $E(h,g)$ is regular. □

Corresponding to a result in [EnR] we also have

Theorem 4: Let h and g be homomorphisms. If h has bounded delay in both directions, then $E(h,g)$ is regular.

Proof: As above we should conclude that the pair (h,g) has a unique continuation property. For situations where g is "ahead", i.e. $uh \prec ug$, this is clearly true, since h has bounded delay from left to right. So let $u \prec v \in E(h,g)$ and $ug \prec uh$. We should show that if $|uh| - |ug|$ is large enough there exists a unique a such that ua is a prefix of a word in $E(h,g)$. Let $uaw \in E(h,g)$. Now if $|uh| - |ug|$ is large then also $|wg| - |wh|$ is large and since wh is a suffix of wg the uniqueness of a follows from the fact that h (which is now "slower") has bounded delay from right to left, too. So the theorem is established. $\square$

It is interesting to note that the equality set of two codes need not be regular, or even context-free. In fact, as regards bounded delay properties the assumptions of Theorems 3 and 4 can not be weakened. This follows from the next example where we introduce a prefix p and a suffix s such that $E(p,s)$ is non context-free.

Example Let $p,s : \{a,b,c,d,e,f\}^* \to \{1,2,3,4,5\}^*$ be homomorphisms defined by the table

	a	b	c	d	e	f
p	1234	2323	4	24	32	5
s	1	23	4	42	3232	4325

The word $abcb^2de^2cef$ is in $E(p,s)$ since

p: a b c b b d e e c e f

1 2 3 4 2 3 2 3 4 2 3 2 3 2 3 2 3 2 4 3 2 3 2 4 3 2 5

s: a b c b b d e e c e f

It is straightforward to see that

$$E(p,s) = (\{abcb^2c\ldots cb^{2^n}de^{2^n}c\ldots ce^2cef \mid n \geq 0\} \cup \{c\})^* .$$

Hence $E(p,s)$ is not context-free.

After establishing that the equality sets of bounded delay codes (or elementary homomorphisms) is regular some natural questions arise. Is it effectively regular? Is the Post Correspondence Problem decidable for these homomorphisms? Or is the Post Correspondence Problem decidable or undecidable for codes in general? We do not know the answers. Our example, however, indicates that injective homomorphisms are probably much more powerful than bounded delay (or elementary) homomorphisms to generate equality languages. So it is likely that the Post Correspondence Problem for codes is more difficult than for example for elementary homomorphisms. Related topics, especially the Post Correspondence Problem in the binary case, are discussed in [CK].

We finish this note by mentioning another situation when the equality set is regular.

Theorem 5: Let h and g be homomorphisms. If h is periodic and g is injective, then there exists effectively a word w (possibly empty) such that $E(h,g) = w^*$.

Proof: The existence of w has been proved in [CS] (although using a different formulation). The effectiveness follows from

Lemma: It is decidable whether $E(h,g)$ is empty for two homomorphisms h and g one of which is periodic.

Proof: Let $h : A^* \to B^*$ such that $Ah \subseteq z^*$ for some $z \in B^*$. Define

$$L_1 = (A^*g \cap z^*)g^{-1}$$

and

$$L_2 = \{u \in A^* \mid |uh| = |ug|\}.$$

Then

$$E(h,g) = L_1 \cap L_2 \ .$$

Clearly, L_1 regular and $L_2\psi$, where $\psi : A^* \to N^{|A|}$ is the Parikh mapping, is semilinear in the sense of [G]. Since $L_2 = L_2\psi\psi^{-1}$ it follows that $L_1 \cap L_2$ is empty if and only if $\psi(L_1) \cap \psi(L_2)$ is empty, so the proof is completed. □

In conclusion we mention that the above lemma represents (besides some trivialities) the only case known to the authors when the Post Correspondence Problem is decidable. Are there any others? Are there any "reasonable" classes of homomorphisms with the decidable Post Correspondence Problem?

Acknowledgement

The authors are grateful to Karel Culik II for useful comments and for arranging excellent working conditions during the authors' visit in the University of Waterloo. The work was supported by the Natural Sciences and Engineering Council of Canada, Grant No. A7403.

References

[C] K. Culik II, A purely homomorphic characterization of recursively enumerable sets, to appear in JACM.

[CK] K. Culik and J. Karhumäki, On the equality sets for homomorphisms on free monoids with two generators, University of Waterloo, Research Report CS-79-17.

[CS] K. Culik II and A. Salomaa, Test sets and checking words for homomorphism equivalence, University of Waterloo, Research Report CS-79-04.

[EhRI] A. Ehrenfeucht and G. Rozenberg, Simplifications of homomorphisms, Information and Control 38 (1978) 298-305.

[EhRII] A. Ehrenfeucht and G. Rozenberg, Elementary homomorphisms and a solution of the DOL sequence equivalence problem, Theoretical Computer Science 7 (1978) 169-183.

[EnR] J. Engelfriet and G. Rozenberg, Equality languages and fixed point languages, University of Antwerpen, Research Report 78-05.

[G] S. Ginsburg, The mathematical theory of context-free languages, McGraw-Hill, New York, 1966.