78-13

A Purely Homomorphic Characterization of Recursively Enumerable Sets

K. CULIK II

University of Waterloo, Waterloo, Ontario, Canada

ABSTRACT. Characterizations of recursively enumerable sets as mappings of equality and minimal sets are given. An equality (minimal) set is the set of all (minimal) solutions of an instance of the Post correspondence problem where the solutions are viewed as strings. The main result is that every recursively enumerable set can be expressed (effectively) as a homomorphic image of a minimal set. From the algebraic point of view this seems to be the simplest characterization of recursively enumerable languages. A corollary of the main result is the solution of an open problem formulated by A. Salomaa. A purely homomorphic characterization of regular sets is derived. How such a characterization can be obtained for various time and space complexity classes for languages is outlined.

KEY WORDS AND PHRASES: formal languages, recursively enumerable sets, homomorphic characterization, equality sets

CR CATEGORIES: 5.22, 5.23, 5.25, 5.27

In several recent proofs of decidability results (e.g. [2, 3]) it turned out to be of crucial importance to effectively check whether two homomorphisms h_1, h_2 on a free monoid Σ^*, generated by an alphabet Σ, are equal on a certain subset of Σ^*, or alternatively, to find the language of all $w \in \Sigma^*$ for which $h_1(w) = h_2(w)$. Such languages were explicitly introduced as equality sets in [5] and further studied. We introduce here another group of languages called minimal sets which are defined as minima of equality sets using the terminology of [4], i.e. a minimal set is a subset of an equality set L containing all strings which have no proper prefix from L. These sets were also implicitly considered in [5], where it was noted that each equality set is a star language (star event in the sense of [1]) and where it was also shown that each minimal set is the minimal star root of an equality set.

Alternatively we can say that an equality set is the set of all solutions of an instance of the Post correspondence problem; a minimal set is the set of all its minimal solutions. Here we consider an instance with lists A, B of length n over alphabet Σ as homomorphisms A, B from $\{1, 2, \ldots, n\}^*$ to Σ^*, and its solution as a string over $\{1, \ldots, n\}$. Each equality set also contains ϵ, the empty string.

We are looking for characterizations of all recursively enumerable languages by mappings of minimal or equality sets. Our main result is that every recursively enumerable language can be expressed as a homomorphic image of a minimal set. As simple modifications of the proof of this main result we will obtain several other more complicated characterizations of recursively enumerable sets, one of them already shown in [5]. We will also solve an open problem from [5]. We show that the regular set is characterized by a

This research was supported by the National Research Council of Canada under Grant No. A7403.

Author's address: Department of Computer Science, University of Waterloo, Waterloo, Ontario N2L 3G1, Canada.

restriction on the pair of homomorphisms defining a minimal set. Finally, we outline how both time and space complexity classes for languages can be characterized by certain restrictions on homomorphisms used in our formula defining a language.

Definition. For homomorphisms $h_1, h_2 : \Sigma^* \to \Delta^*$,

$$\begin{aligned} E(h_1, h_2) &= \{w \in \Sigma^* : h_1(w) = h_2(w)\} \\ e(h_1, h_2) &= \{w \in \Sigma^+ : h_1(w) = h_2(w) \text{ and if } w = uv \text{ where} \\ &\qquad u \in \Sigma^+, v \in \Sigma^+ \text{ then } h_1(u) \neq h_2(u)\}. \end{aligned}$$

Note that using the notation of [4]

$$e(h_1, h_2) = \min(E(h_1, h_2)) - \{\epsilon\}.$$

If a language can be expressed as $E(h_1, h_2)$, then it is called an equality set [5]; if it can be expressed as $e(h_1, h_2)$, then it is called a minimal set.

Now, we show that every recursively enumerable set can be expressed as a homomorphic image of a minimal set.

THEOREM 1. *For each recursively enumerable set L (represented by a grammar), there effectively exist homomorphisms h_0, h_1, h_2 such that $L = h_0(e(h_1, h_2))$. The homomorphism h_0 maps each symbol either to itself or to ϵ.*

PROOF. Let $G = (N, T, P, S)$ be a phrase structure grammar (N are nonterminals, T are terminals, $P \subseteq N^+ \times (N \cup T)^*$, and $S \in N$) such that $L = L(G)$. Assume that $(S, S) \in P$ to assure that each string in $L(G)$ has a derivation of odd length. We will construct homomorphisms h_0, h_1, and h_2 such that $L = h_0(e(h_1, h_2))$ and outline the proof of the correctness of this construction.

Let $\Gamma = \{\#\} \cup N \cup T \cup \{\langle \alpha, \beta \rangle : (\alpha, \beta) \in P\}$, $\bar{\Gamma} = \{\bar{a} : a \in \Gamma\}$, and $\hat{T} = \{\hat{a} : a \in T\}$. Let $\Sigma = \Gamma \cup \bar{\Gamma} \cup \hat{T} \cup \{\vdash, \$, 0, 2, 3\}$ and $\Delta = \Gamma \cup \bar{\Gamma} \cup \{0, 1, 2, 3, \vdash, \dashv\}$. For $x \in T^*$ let $\hat{x}$ denote the string obtained from x by "hooding" each symbol.

Homomorphisms h_1, h_2 from Σ^* to Δ^* are defined by the following table:

ξ	$\vdash$	X	$\bar{X}$	$\langle \alpha, \beta \rangle$	$\overline{\langle \alpha, \beta \rangle}$	\$	$\hat{a}$	0	2	3
$h_1(\xi)$	$\vdash$	X	$\bar{X}$	α	$\bar{\alpha}$	$\#$	$\bar{a}$	10	$\dashv$	123
$h_2(\xi)$	$\vdash S\#$	$\bar{X}$	X	$\bar{\beta}$	β	$\dashv 1$	01	ϵ	2	3

for all $X \in N \cup T \cup \{\#\}$, $(\alpha, \beta) \in P$, and $a \in T$. Let $P = \{p_1, p_2, \ldots, p_m\}$. Consider a derivation of grammar G

$$S = w_0 \Rightarrow w_1 \Rightarrow \cdots \Rightarrow w_n$$

where $w_n \in T^*$, n is odd and $w_{i-1} \Rightarrow w_i$ is obtained using production $p_{k_i} = (\alpha_i, \beta_i)$ from $P = \{p_1, \ldots, p_m\}$ for $i = 0, \ldots, n-1$. Then there are u_i, v_i such that $w_i = u_i \alpha_i v_i$ for $i = 0, \ldots, n-1$ and $w_i = u_{i-1}\beta_{i-1}v_{i-1}$ for $i = 1, \ldots, n$. Consider the string γ in Σ^* of the form

$$\vdash \langle S, \beta_0 \rangle \#\bar{u}_1\overline{\langle \alpha_1, \beta_1 \rangle}\bar{v}_1\#u_2\langle \alpha_2, \beta_2 \rangle v_2\#\bar{u}_3\overline{\langle \alpha_3, \beta_3 \rangle}\bar{v}_3\# \cdots \cdots \#u_{n-1}\langle \alpha_{n-1}, \beta_{n-1} \rangle v_{n-1}\$\hat{w}_n 2\,(0)^{|w_n|}\,3, \quad (*)$$

where $|w_n|$ is the length of w_n. We have

$$h_1(u_i\langle \alpha_i, \beta_i \rangle v_i) = u_i\alpha_i v_i, \quad h_2(u_i\langle \alpha_i, \beta_i \rangle v_i) = \bar{u}_i\bar{\beta}_i\bar{v}_i \quad \text{for} \quad i = 0, 2, \ldots, n-1$$

$$h_1(\bar{u}_i\overline{\langle \alpha_i, \beta_i \rangle}\bar{v}_i) = \bar{u}_i\bar{\alpha}_i\bar{v}_i, \quad h_2(\bar{u}_i\overline{\langle \alpha_i, \beta_i \rangle}\bar{v}_i) = u_i\beta_i v_i \quad \text{for} \quad i = 1, 3, \ldots, n-2.$$

Therefore

$$h_1(\vdash\langle S, \beta_0 \rangle\#) = h_2(\vdash) = \vdash S\#,$$

$$h_1(u_i\langle \alpha_i, \beta_i \rangle v_i) = h_2(\bar{u}_{i-1}\overline{\langle \alpha_{i-1}, \beta_{i-1} \rangle}\bar{v}_{i-1}) = w_i \quad \text{for} \quad i = 2, 4, \ldots, n-1,$$

$$h_1(\bar{u}_i\overline{\langle \alpha_i, \beta_i \rangle}\bar{v}_i) = h_2(u_{i-1}\langle \alpha_{i-1}, \beta_{i-1} \rangle v_{i-1}) = \bar{w}_i \quad \text{for} \quad i = 1, 3, \ldots, n-2,$$

$$h_1(\hat{w}_n) = h_2(u_{n-1}\langle \alpha_{n-1}, \beta_{n-1} \rangle v_{n-1}) = \bar{w}_n,$$

and finally

$$h_1(20^{|u_n|}3) = h_2(\$\hat{w}_n 20^{|u_n|}3) = \dashv 1(01)^{|u_n|}23.$$

When combining the above with the images of markers $\#, \bar{\#}$ we get $h_1(\gamma) = h_2(\gamma)$ as shown in Figure 1, we also see that $h_1(\gamma') \neq h_2(\gamma')$ for each proper prefix γ' of γ.

Now, we define homomorphism h_0 by $h_0(\hat{a}) = a$ for $a \in T$ and $h_0(b) = \epsilon$ for $b \in \Sigma - \hat{T}$. Clearly, we have shown that $L(G) \subseteq h_0(e(h_1, h_2))$.

Consider now any string w in Σ^+ for which $h_1(w) = h_2(w)$ and $h_1(u) \neq h_2(u)$ for each proper prefix of w. We will analyze the form of w and its images. We start from the right end.

(0) First, we note that neither the first nor the last symbol of w is "0" since "1" cannot begin $h_2(w)$ and "0" cannot end $h_2(w)$.

(1) The only symbol $\xi \in \Sigma - \{0\}$ for which $h_1(\xi)$ is a suffix of $h_2(\xi)$ or vice versa is "3", which thus must be the last symbol of w. Since "3" in both homomorphisms cannot be obtained from any other symbol, the occurrences of "3" produced from the same "3" must match. Since w is a minimal solution we conclude that it has exactly one occurrence of 3 as its last symbol.

(2) By (1) $h_1(w)$ has a suffix 123. With each occurrence of "0" in $h_1(w)$ there must be a left-adjacent occurrence of "1" and in $h_2(w)$ with each occurrence of "0" there must be a right-adjacent occurrence of "1". Since there is exactly one occurrence of "1" in $h_1(w)$ not paired with "0" coming from 3 we conclude that $h_1(w)$ has the form $t(01)^k123$ for some $k \geq 0$, and $t \in \Delta^*$.

Now, it is easy to deduce more information step by step about the form of w and $h_1(w)$. We get

(3) w has a suffix 20^k3 for some $k \geq 0$.
(4) $h_1(w)$ has a suffix $\dashv (10)^k123$ for some $k \geq 0$.
(5) w has a suffix $\$\hat{x}20^k3$ for some $x \in T^*$, $|x| = k$.
(6) $h_1(w)$ has a suffix $\#\bar{x} \dashv (10)^k123$ for some $x \in T^*$, $|x| = k$.

Now, we will analyze the form of w and $h_1(w)$ from the left end:

(7) The only symbol $\xi \in \Sigma - \{0\}$ for which $h_1(\xi)$ is a prefix of $h_2(\xi)$, or vice versa, is $\vdash$. Therefore the first symbol of w must be $\vdash$. Thus $h_2(w)$ has a prefix $\vdash S\#$.

(8) By (7) we must have either prefix $\vdash \langle S, \beta_0\rangle\#$ or prefix $\vdash S\#$ or prefix $\vdash S\$$. In the first two cases $h_2(w)$ has a prefix $\vdash S\#\bar{y}_1\bar{\#}$ where $S \underset{G}{\Rightarrow}^* y_1$ and in the third case $h_2(w)$ has a prefix $\vdash S\#\bar{y}_1 \dashv$ where $S \underset{G}{\Rightarrow}^* \bar{y}_1$.

(9) In the two first cases of (8), the next symbols after $\#$ will be from $\bar{\Gamma}$ ending with $\bar{\#}$ and after that symbols from $\Gamma \cup \{\$\}$ ending with $\#$ or $\$$. This reasoning is repeated until $\$$ appears.

When continuing the analysis of w and its images in the way outlined we can verify that $h_1(w)$ $(=h_2(w))$ must have the form $\vdash y_0\#\bar{y}_1\bar{\#}y_2\#\bar{y}_3\bar{\#} \cdots \#\bar{y}_r \dashv (10)^{|y_r|}123$ where $r \geq 1$,

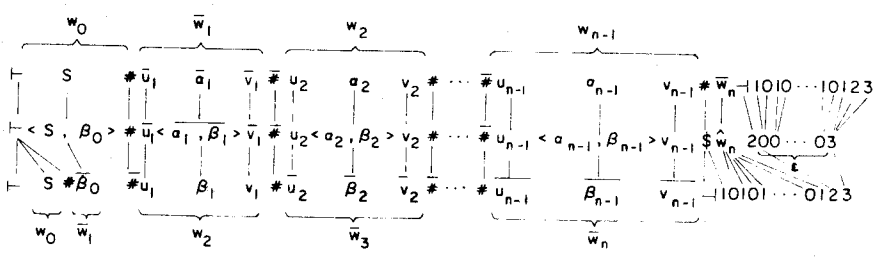

FIG. 1

$y_0 = S$ and $y_{i-1} \underset{G}{\Rightarrow}^* y_i$ for $i = 1, 2, \ldots, r$. Hence, w must be of the form

$$\vdash z\$\hat{y}20^{|y|}3 \qquad (**)$$

for some z in $(\Gamma \cup \bar{\Gamma})^*$ and y in T^* such that $S \underset{G}{\Rightarrow}^* y$. Since each w in $e(h_1, h_2)$ has the form (**) we conclude that $h_0(e(h_1, h_2)) \subseteq L(G)$.

By interchanging the roles of T and $\hat{T}$ in our construction we can assure that the homomorphism h_0 satisfies the condition of mapping each symbol either to itself or to ϵ. We have not done it in order to simplify the notation. □

COROLLARY 1. *A language over alphabet Σ is recursively enumerable iff there exist an alphabet Δ and two homomorphisms h_1, h_2 on $(\Sigma \cup \Delta)^*$ so that $L = h(e(h_1, h_2))$, where h is the homomorphism mapping each symbol of Σ to itself and each symbol of Δ to ϵ.*

PROOF. We can suitably rename the symbols and extend the homomorphisms h_1, h_2 from the proof of Theorem 1 to a common alphabet so that we do not get any additional strings on which h_1 and h_2 are equal. It is, of course, obvious that each language expressed in this way is recursively enumerable. □

From Theorem 1 immediately follows the solution of an open problem from [5]:

COROLLARY 2. *For every recursively enumerable language L, there effectively exist homomorphisms h_0, h_1, h_2, where h_0 maps each symbol to itself or to ϵ, so that $L^* = h_0(E(h_1, h_2))$.*

By modifying the above construction we can show that every recursively enumerable language can be expressed as a gsm (generalized sequential machine) mapping of an equality set. This result has been shown in [5].

If we would like to use $E(h_1, h_2)$ instead of $e(h_1, h_2)$ in the above construction, then we have to modify the mapping h_0 to erase everything after the first occurrence of the "endmarker" 3 since, clearly, $E(h_1, h_2) = (e(h_1, h_2))^*$. This cannot be done anymore by any homomorphism but can be done easily by a deterministic gsm mapping.

THEOREM 2. *For each recursively enumerable set L, there effectively exist homomorphisms h_1, h_2 and a deterministic gsm mapping g so that $L = g(E(h_1, h_2))$.*

PROOF. Consider h_0, h_1, h_2 and G from the proof of Theorem 1. Using the notation of [4], let gsm mapping g be defined by gsm machine $M = (\{q_0, q_f\}, \Sigma, \Delta, \delta, q_0, \{q_f\})$, where $\delta(q_0, x) = (q_0, h_0(x))$ for $x \in \Sigma - \{3\}$, $\delta(q_0, 3) = (q_f, \epsilon)$, and $\delta(q_f, x) = (q_f, \epsilon)$ for each $x \in \Sigma$. Clearly, $g(E(h_1, h_2)) = L(G)$. □

We can further exploit the construction from the proof of Theorem 1 to show that each recursively enumerable set can be obtained from a minimal set by operations of left and right quotient with regular sets (see [4]) or from an equality set using the same operations plus intersection with a regular set.

THEOREM 3. *For each recursively enumerable set L there exist homomorphisms h_1, h_2 and regular sets R_1, R_2, R_3 so that*

(a) $L = R_1 \backslash e(h_1, h_2) / R_2$,

(b) $L = (R_1 \backslash E(h_1, h_2) / R_2) \cap R_3$.

PROOF. Consider L, h_1, h_2 from the proof of Theorem 1, where the role of T and $\hat{T}$ is interchanged. Let $R_1 = \Sigma^*\{\$\}$, $R_2 = \{2\}\Sigma^*$, and $R_3 = T^*$. Clearly, (a) and (b) are satisfied for this choice. □

We conclude with a purely homomorphic characterization of regular sets. For that purpose we need the notion of balance originally introduced in [2].

Definition. Consider two fixed homomorphisms h_1 and h_2 from Σ^* to Δ^* and a word w in Σ^*. The *balance* of w is defined by

$$B(w) = |h_1(w)| - |h_2(w)|.$$

We say that the pair (h_1, h_2) has k-bounded balance on a given language L for some $k \geq 0$ if $|B(u)| \leq k$ holds for each prefix u of every word in L. We say that the pair (h_1, h_2) has

bounded balance on L if it has k-bounded balance on L for some $k \geq 0$.

For $k \geq 0$, we denote by $E_k(h_1, h_2)$, $e_k(h_1, h_2)$ the largest subsets of $E(h_1, h_2)$, $e(h_1, h_2)$ such that the pair (h_1, h_2) has k-bounded balance on $E_k(h_1, h_2)$, $e_k(h_1, h_2)$, respectively. In [6], and essentially already in [2], it was shown that $E_k(h_1, h_2)$ is regular. Using almost the same argument we have the following:

LEMMA 1. *For homomorphisms h_1, h_2, and $k \geq 0$, $e_k(h_1, h_2)$ is regular.*

THEOREM 4. *A language L is regular iff there are homomorphisms h_0, h_1, h_2 such that $L = h_0(e(h_1, h_2))$ and the pair (h_1, h_2) has bounded balance on $e(h_1, h_2)$.*

PROOF. Assume that $L = h_0(e(h_1, h_2))$ and the pair (h_1, h_2) has k-bounded balance for some $k > 0$. Then $e(h_1, h_2) = e_k(h_1, h_2)$ and by Lemma 1 $e(h_1, h_2)$ is regular. Thus also its homomorphic image $h_0(e(h_1, h_2))$ is regular.

Assume that L is regular. Then there is a (nondeterministic) finite automaton with only one final state q_f (possibly $q_0 = q_f$). $A = (K, \Sigma, \delta, q_0, \{q_f\})$ such that $L = L(A)$. Let $\pi = \{\langle q, a, p\rangle : p \in \delta(q, a), q, p \in K, a \in \Sigma\}$, π is a finite alphabet. Let $\bar{\pi} = \{\bar{\alpha} : \alpha \in \pi\}$ *and* $\Gamma = \pi \cup \bar{\pi} \cup \{\vdash, \dashv, \$\}$.

Let homomorphisms h_0, h_1, h_2 be given by the table

ξ	$\vdash$	$\langle q, a, p\rangle$	$\overline{\langle q, a, p\rangle}$	$\dashv$	$\$$
$h_0(\xi)$	ϵ	a	a	ϵ	ϵ
$h_1(\xi)$	$\vdash$	q	$\bar{q}$	$q_f\dashv$	$\bar{q}_f\dashv$
$h_2(\xi)$	$\vdash q_0$	$\bar{p}$	p	$\dashv$	$\dashv$

for all $a \in \Sigma$, $p, q \in K$ such that $p \in \delta(q, a)$. It is easy to verify that $L = h_0(e(h_1, h_2))$. This fact is illustrated in Figure 2 where n is even. For computations of odd length the endmarker \$ would be used instead of $\dashv$. □

The condition of bounded balance in Theorem 4 is noneffective. However, we can give the following stronger and effective conditions on the homomorphisms.

THEOREM 5. *A language L is regular iff there are homomorphisms h_0, h_1, h_2 such that $L = h_0(e(h_1, h_2))$ where h_0 is 1-limited on $e(h_1, h_2)$ (cf. [4]) and the pair (h_1, h_2) has 1--bounded balance on $e(h_1, h_2)$. If ϵ is not in L, then h_0 can be chosen to be a letter-to-letter homomorphism.*

PROOF. The if part follows by Theorem 4. Also the first statement of the only-if-part has been proved in the proof of Theorem 4 since h_0 erases only the first and the last of each string in $e(h_1, h_2)$ and, clearly, $B(w) = 1$ for all proper prefixes of strings in $e(h_1, h_2)$. It is easy to see that if $\epsilon \notin L$ we can modify the construction of automaton A by merging $\vdash$ with $\langle q_0, a, q\rangle$ and $\dashv$ or \$ with $\langle q, a, q_f\rangle$ for each $a \in \Sigma$ and $q \in K$. We get larger but still a finite alphabet. The homomorphism h_0 is then defined only on triples $\langle q, a, p\rangle$ and therefore letter-to-letter. □

We outline some further results which will be studied in detail in a separate paper. They show that certain restrictions on h_0 and the pair (h_1, h_2) in the formula $h_0(e(h_1, h_2))$ give a characterization of time and space complexity classes for the languages.

The time complexity characterization was suggested by one of the referees of this paper. The notion of k-limited erasing (see [4]) is generalized as follows. For a monotone function f on integers we say that erasing h is f-bounded on a language L if for each w in L at most $f(|w|)$ consecutive symbols of w may be erased. We get the following result which comes from the proof of Theorem 1 and known results.

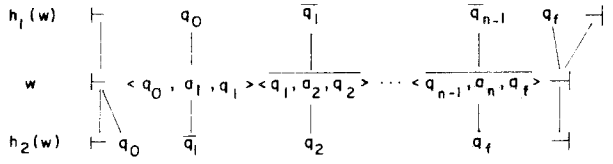

FIG. 2

Let $\mathcal{F}$ be a class of functions closed under squaring and composition and let $\mathcal{L}_{\mathcal{F}}$ be the class of languages accepted by nondeterministic Turing machines that operate with time bounds in $\mathcal{F}$. Then L is in $\mathcal{L}_{\mathcal{F}}$ iff there exist homomorphisms h_0, h_1, h_2 such that $h_0(e(h_1, h_2)) = L$ and h_0 is f-bounded erasing on $e(h_1, h_2)$ for some of $f \in \mathcal{F}$.

As special cases we have, for example, the following. A language L is in NP (is primitive recursive, recursive) iff there exist homomorphisms h_0, h_1, h_2 such that $h_0(e(h_1, h_2)) = L$ and h_0 is polynomial- (primitive recursive-, recursive-) bounded erasing on $e(h_1, h_2)$.

Our Theorem 4 shows that a constant bound on the balance of the pair (h_1, h_2) on $e(h_1, h_2)$ characterizes the regular sets. In a similar way as we generalized the notion of k-limited (k-bounded) erasing, we also can generalize the notion of k-bounded balance. Given a monotone function f on integers, a language $L \subseteq \Sigma^*$, and an erasing h on Σ^*, we say that a pair of homomorphisms (h_1, h_2) has f-bounded balance on L with respect to h if for each x in L and each prefix w of x we have $|B(w)| \leq f(|h(x)|)$. It can be shown for certain classes of functions $\mathcal{F}$ that a language L is of space complexity $\mathcal{F}$ iff there exist an erasing h_0 and homomorphisms h_1, h_2 such that $L = h_0(e(h_1, h_2))$ and the pair (h_1, h_2) has f-bounded balance on $e(h_1, h_2)$ with respect to h_0 for some $f \in \mathcal{F}$. For example, we conjecture that the context sensitive languages are exactly those which can be expressed in the form $h_0(e(h_1, h_2))$ where the pair (h_1, h_2) has linearly bounded balance on $e(h_1, h_2)$ with respect to h_0.

ACKNOWLEDGMENTS. The author is grateful to J. Opatrny for discussions of the characterization of regular sets and to M. Penttonen for remarks on a draft of this paper.

REFERENCES

1. BRZOZOWSKI, J.A. Roots of star events. *J. ACM 14*, 3 (July 1967), 466–477.
2. CULIK II, K., AND FRIS, I. The decidability of the equivalence problem for DOL-systems. *Inform. and Control 35* (1977), 20–39.
3. CULIK II, K., AND SALOMAA, A. On the decidability of homomorphism equivalence for languages. To appear in *J. Comptr. Syst. Sci. 17* (1978).
4. HOPCROFT, J.E., AND ULLMAN, J.D. *Formal Languages and Their Relations to Automata.* Addison-Wesley, Reading, Mass., 1969.
5. SALOMAA, A. *Equality Sets for Homomorphisms of Free Monoids.* To appear in *Acta Cybernetica.*
6. SALOMAA, A. DOL equivalence: The problem of iterated homomorphisms. *Bull. EATCS* (European Assoc. Theoret. Comptr. Sci.) *4* (Jan. 1978), 5–12.

RECEIVED JANUARY 1978; REVISED MAY 1978

A Purely Homomorphic Characterization of Recursively Enumerable Sets

K. Culik II

Department of Computer Science
University of Waterloo
Waterloo, Ontario, Canada

Research Report CS-78-13

January 1978

A Purely Homomorphic Characterization of Recursively Enumerable Sets

K. Culik II
Department of Computer Science
University of Waterloo
Waterloo, Ontario, Canada

January 1978

Abstract

We give characterizations of recursively enumerable sets as mappings of equality and minimal sets. An equality (minimal) set is the set of all (minimal) solutions of an instance of the Post correspondence problem where the solutions are viewed as strings. The main result is that every recursively enumerable set can be expressed (effectively) as a homomorphic image of a minimal set. From the algebraic point of view this seems to be the simplest characterization of recursively enumerable languages. As a corollary of our main result we get the solution of an open problem formulated by A. Salomaa. We also get a purely homomorphic characterization of regular sets.

In several recent proofs of decidability results (see e.g. [2, 3]) it turned out to be of crucial importance to effectively check whether two homomorphisms h_1, h_2 on a free monoid Σ^*, generated by an alphabet Σ, are equal for certain subset of Σ^*, or alternatively, to find the language of all $w \in \Sigma^*$ for which $h_1(w) = h_2(w)$. Such languages were explicitly introduced as equality sets in [5] and further studied. We introduce here another type of languages called minimal sets which are defined as minima of equality sets using the terminology of [4], i.e. a minimal set is a subset of an equality set L containing all strings which have no proper prefix from L. These sets were also implicitly considered in [5], where it was noted that each equality set is a star language (star event in the sense of [1]) and also shown that each minimal set is the minimal star root of a equality set.

Alternatively we can say that an equality set is the set of all solutions of an instance of the Post correspondence problem, a minimal set is the set of all its minimal solutions. Here we consider an instance with lists A, B of length n over alphabet Σ as homomorphisms A, B from $\{1, 2, \ldots, n\}^*$ to Σ^*, and its solution as a string over $\{1, \ldots, n\}$. Each equality set also contains ε, the empty string.

We are looking for characterizations of all recursively enumerable languages by mappings of minimal or equality sets. Our main result is that every recursively enumerable language can be expressed as a homomorphic image of a minimal set. As easy modifications of the proof of this main result we will obtain several other more complicated characterizations of recursively enumerable sets, one of them was already

shown in [5]. We will also solve an open problem from [5]. Finally we get a restriction on h_1, h_2 to characterize regular sets.

Definition For homomorphisms $h_1, h_2 : \Sigma^* \to \Delta^*$

$$E(h_1, h_2) = \{w \in \Sigma^* : h_1(w) = h_2(w)\}$$

$$e(h_1, h_2) = \{w \in \Sigma^+ : h_1(w) = h_2(w) \text{ and if } w = uv \text{ where } u \in \Sigma^+, v \in \Sigma^+ \text{ then } h_1(v) \neq h_2(v)\} .$$

Note that using the notation of [4]

$$e(h_1, h_2) = \text{Min}(E(h_1, h_2)) - \{\varepsilon\}$$

If a language can be expressed as $E(h_1, h_2)$ it is called an equality set [5], if it can be expressed as $e(h_1, h_2)$ it is called a minimal set.

Now, we show that every recursively enumerable set can be expressed as a homomorphic image of a minimal set.

Theorem 1 For each recursively enumerable set L (represented by a grammar), there effectively exist homomorphisms h_0, h_1, h_2 so that $L = h_0(e(h_1, h_2))$. The homomorphism h_0 maps each symbol either to itself or to ε .

Proof Let $G = (N, T, P, S)$ be a phrase structure grammar such that $L = L(G)$. Assume that $S \to S \in P$ to assure that each string in $L(G)$ has a derivation of odd length.

Let $\Gamma = \{\#\} \cup N \cup T \cup \{<\alpha,\beta> : (\alpha, \beta) \in P\}$, $\overline{\Gamma} = \{\overline{a} : a \in \Gamma\}$ and $\hat{T} = \{\hat{a} : a \in T\}$. Let $\Sigma = \Gamma \cup \overline{\Gamma} \cup \hat{T} \cup \{\vdash, \$, 0, 2, 3\}$ and $\Delta = \Gamma \cup \overline{\Gamma} \cup \{0, 1, 2, 3, \vdash, \dashv\}$. Homomorphisms h_1, h_2 from Σ^*

to Δ^* are defined by the following table

ξ	$\vdash$	X	$\overline{X}$	$\langle\alpha, \beta\rangle$	$\overline{\langle\alpha, \beta\rangle}$	\$	$\hat{a}$	0	2	3
$h_1(\xi)$	$\vdash$	X	$\overline{X}$	α	$\overline{\alpha}$	#	$\overline{a}$	10	$\dashv$	123
$h_2(\xi)$	$\vdash$S#	$\overline{X}$	X	$\overline{\beta}$	β	$\dashv$1	01	ε	2	3

for all $X \in N \cup T \cup \{\#\}$, $(\alpha, \beta) \in P$ and $a \in T$. Let $P = \{p_1, p_2, \ldots, p_m\}$. Consider a derivation of grammar G

$$S = w_0 \Rightarrow w_1 \Rightarrow \ldots \Rightarrow w_n$$

where n is odd and $w_{i-1} \Rightarrow w_i$ is obtained using production $p_{k_i} = (\alpha_i, \beta_i)$ from $P = \{p_1, \ldots, p_m\}$ for $i = 0,\ldots,n-1$. Then there are u_i, v_i, such that $w_i = u_i\alpha_i v_i$ for $i = 0,\ldots,n-1$ and $w_i = u_{i-1}\alpha_{i-1}v_{i-1}$ for $i = 1,\ldots,n$. For $x \in T^*$ let $\hat{x}$ denote the string obtained from x by "hooding" each symbol. Consider the string γ in Σ^* of the form

$$(*)\quad \vdash\langle S, \beta_0\rangle\#\bar{u}_1\langle\alpha_1, \beta_1\rangle\bar{v}_1\#u_2\langle\alpha_2, \beta_2\rangle v_2\#\bar{u}_3\langle\alpha_1, \beta_1\rangle\bar{v}_3\ \#\ldots$$
$$\ldots\#u_{n-1}\langle\alpha_{n-1},\beta_{n-1}\rangle v_{n-1}\$\hat{w}_n 21(01)^{|w_n|}3$$

where $|w_n|$ is the length of w_n. We have
$h_1(u_i\langle\alpha_i, \beta_i\rangle v_i) = u_i\alpha_i v_i$, $h_2(u_i\langle\alpha_i, \beta_i\rangle v_i) = \bar{u}_i\bar{\beta}_i\bar{v}_i$, for $i = 0,2,\ldots,n-1$ and $h_1(\bar{u}_i\langle\alpha_i, \beta_i\rangle\bar{v}_i) = \bar{u}_i\bar{\alpha}_i\bar{v}_i$, $h_2(\bar{u}_i\langle\alpha_i, \beta_i\rangle\bar{v}_i) = u_i\beta_i v_i$, for $i = 1,3,\ldots,n-2$. Therefore

$h_1(\vdash\langle S, \beta_0\rangle\#) = h_2(\vdash) = \vdash S\#$,

$h_1(u_i\langle\alpha_i, \beta_i\rangle v_i) = h_2(\bar{u}_{i-1}\langle\alpha_{i-1}, \beta_{i-1}\rangle\bar{v}_{i-1}) = w_i$ for $i = 2,4,\ldots,n-1$

$h_1(\overline{u}_i\overline{<\alpha_i, \beta_i>}\overline{v}_i) = h_2(u_{i-1}<\alpha_{i-1}, \beta_{i-1}>v_{i-1}) = \overline{w}_i$ for $i = 1,3,\ldots,n-2$,

$h_1(\hat{w}_n) = h_2(u_{n-1}<\alpha_{n-1}, \beta_{n-1}>v_{n-1}) = \overline{w}_n$ and finally

$h_1(20^{|w_n|}3) = h_2(\$\hat{w}_n 20^{|w_n|}3) = \dashv 1(01)^{|w_n|}23$. When combining the above with the images of markers $\#, \overline{\#}$ we get $h_1(\gamma) = h_2(\gamma)$ as shown in Figure 1, we also see that $h_1(\gamma') \neq h_2(\gamma')$ for each proper prefix γ' of γ .

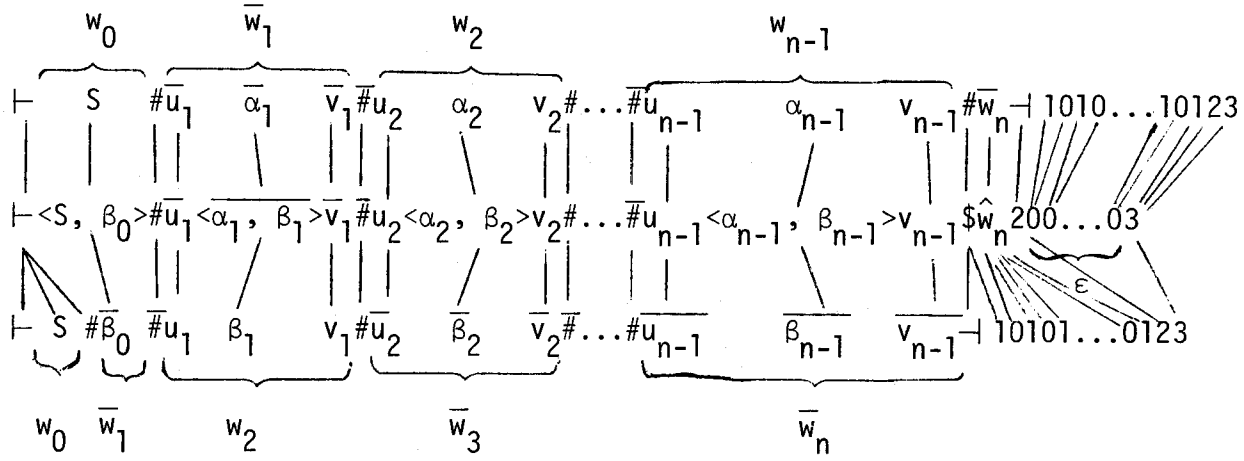

Figure 1

Now, we define homomorphism h_0 by $h_0(\hat{a}) = a$ for $a \in T$ and $h(b) = \varepsilon$ for $b \in \Sigma - \hat{T}$. Clearly, we have shown that $L(G) \subseteq h_0(e(h_1, h_2))$.

Consider now any string w in Σ^+ for which $h_1(w) = h_2(w)$ and $h_1(u) \neq h_2(u)$ for each proper prefix of w . We will analyze the form of w and its images. We start from the right end.

(0) First, we note that neither the first nor the last symbol of w is "0" since in $h_1(w)$ every occurrence of zero must be followed by "1" and in $h_2(w)$ preceded by "1".

(1) The only symbol $\xi \in \Sigma - \{0\}$ for which $h_1(\xi)$ is a suffix of $h_2(\xi)$ or vice versa is "3" which thus must be the last symbol of w. Since "3" in both homomorphisms cannot be obtained from any other symbol, the occurrences of "3" produced from the same "3" must match. Since w is a minimal solution we conclude that it has exactly one occurrence of 3 as its last symbol.

(2) By (1) $h_1(w)$ has a suffix 123. With each occurrence of "0" in $h_1(w)$ there must be a left-adjacent occurrence of "1" and in $h_2(w)$ with each occurrence of "0" there must be a right-adjacent occurrence of "1". Since there is exactly one occurrence of "1" in $h_1(w)$ not parred with "0" coming from 3 we conclude that $h_1(w)$ has the form $t(01)^k123$ for some $k \geq 0$, and $t \in \Delta^*$.

Now, it is easy to deduce step by step more information about the form of w and $h_1(w)$. We get

(3) w has a suffix 20^k3 for some $k \geq 0$.

(4) $h_1(w)$ has a suffix $\dashv(10)^k123$ for some $k \geq 0$.

(5) w has a suffix $\$\hat{x}20^k3$ for some $x \in T^*$, $|x| = k$.

(6) $h_1(w)$ has a suffix $\#\bar{x}\dashv(10)^k123$ for some $x \in T^*$, $|x| = k$.

Now, we will analyze the form of w and $h_1(w)$ from the left end:

(7) The only symbol $\xi \in \Sigma - \{0\}$ for which $h_1(\xi)$ is a prefix of $h_2(\xi)$, or vice versa, is $\vdash$. Therefore the first symbol of w must be $\vdash$. Thus $h_2(w)$ has a prefix $\vdash S\#$.

(8) By (7) we must have either prefix $\dashv\langle S, \beta_0\rangle\#$ or prefix $\vdash S\#$. In either case $h_1(w)$ has a prefix $\vdash S\#\bar{y}_1\bar{\#}$ where $S \overset{\Rightarrow}{G}{}^* y_1$.

When continuing the analysis of w and its images in the way outlined we can verify that $h_1(w)$ $(= h_2(w))$ must have the form $\vdash y_0\#\overline{y}_1\overline{\#}y_2\#\overline{y}_3\overline{\#}\ldots\#\overline{y}_r\dashv (10)^{|y_r|}123$ where $r \geq 1$, $y_0 = S$ and $y_{i-1} \overset{}{\underset{\overline{G}}{\Rightarrow}}{}^* y_i$ for $i = 1,2,\ldots,r$. Hence, w must be of the form

$$(**) \qquad \vdash z\$\hat{y}20^{|y|}3$$

for some z in $(\Gamma \cup \overline{\Gamma})^*$ and y in T^* such that $S \overset{}{\underset{\overline{G}}{\Rightarrow}}{}^* y$. Since each w in $e(h_1, h_2)$ has the form $(**)$ we conclude that $h_0(e(h_1, h_2)) \subseteq L(G)$.

□

Corollary 1 A language over alphabet Σ is recursively enumerably iff there exist an alphabet Δ and two homomorphisms h_1, h_2 on $(\Sigma \cup \Delta)^*$ so that

$$L = h(e(h_1, h_2))$$

where h is the homomorphism mapping each symbol of Σ to itself and each symbol of Δ to ε.

Proof We can suitably rename the symbols and extend the homomorphisms h_1, h_2 from the proof of Theorem 1 to a common alphabet so that we do not get any additional strings on which h_1 and h_2 are equal. It is, of course, obvious that each language expressed it this way is recursively enumerable.

□

From Theorem 1 immediately follows the solution of an open problem from [5]:

Corollary 2 For every recursively enumerable language L, there effectively exist homomorphisms h_0, h_1, h_2, where h_0 maps each symbol to itself or to ε, so that $L^* = h_0(E(h_1, h_2))$.

We show that by modifying the above construction we can show that every recursively enumerable language can be expressed as a gsm mapping of an equality set. This result has been shown in [5].

If we would like to use $E(h_1, h_2)$ instead of $e(h_1, h_2)$ in the above construction than we have to modify the mapping h_0 to erase everything after the first occurrence of the "endmarker" 3 since, clearly, $E(h_1, h_2) = (e(h_1, h_2))^*$. This cannot be done anymore by any homomorphism but can be done easily by a deterministic gsm mapping.

Theorem 2 For each recursively enumerable set L, there effectively exist homomorphisms h_1, h_2 and deterministic gsm mapping g so that $L = g(E(h_1, h_2))$.

Proof Consider h_0, h_1, h_2 and G from the proof of Theorem 1. Using the notation of [4], let gsm mapping g be defined by gsm machine $M = (\{q_0, q_f\}, \Sigma, \Delta, \delta, q_0, \{q_f\})$, where $\delta(q_0, x) = (q_0, h_0(x))$, for $x \in \Sigma - \{3\}$, $\delta(q_0, 3) = (q_1, \varepsilon)$ and $\delta(q_1, x) = (q_1, \varepsilon)$ for each $x \in \Sigma$. Clearly, $g(E(h_1, h_2)) = L(G)$.

□

We can further exploit the construction from the proof of Theorem 1 to show that each recursively enumerable set can be obtained from a minimal set by operations of left and right quotient with regular sets (see [4]) or from an equality set using the same operations plus intersection with a regular set.

Theorem 3 For each recursively enumerable set L there exist homomorphisms h_1, h_2 and regular sets R_1, R_2, R_3 so that

(a) $L = R_1 \backslash e(h_1, h_2)/R_2$

(b) $L = (R_1 \backslash E(h_1, h_2)/R_2) \cap R_3$.

Proof Consider L, h_1, h_2 from the proof of Theorem 1, where the role of T and $\hat{T}$ is interchanged. Let $R_1 = \Sigma^*\{\$\}$, $R_2 = \{2\}\Sigma^*$ and $R_3 = T^*$. Clearly, (a) and (b) are satisfied for this choice.

□

We conclude with a purely homomorphic characterization of regular sets. For that purpose we need the notion of balance originally introduced in [2].

Definition Consider two fixed homomorphisms h_1 and h_2 from Σ^* to Δ^* and a word w in Σ^* . The balance of w is defined by

$$B(w) = |h_1(w)| - |h_2(w)| .$$

We say that the pair (h_1, h_2) has k-bounded balance on a given language L for some $k \geq 0$ if $|B(u)| \leq k$ holds for each prefix u of every word in L .

We say that the pair (h_1, h_2) has bounded balance on L if it has k-bounded balance on L for some $k \geq 0$.

For $k \geq 0$, we denote by $E_k(h_1, h_2)$, $e_k(h_1, h_2)$ the largest subsets of $E(h_1, h_2)$, $e(h_1, h_2)$ such that the pair (h_1, h_2) has k-bounded balance on $E_k(h_1, h_2)$, $e_k(h_1, h_2)$, respectively. In [6], and essentially already in [2], it was shown that $E_k(h_1, h_2)$ is regular. Using almost the same argument we have the following:

Lemma 1 For homomorphisms h_1, h_2 and $k \geq 0$, $e_k(h_1, h_2)$ is regular.

Theorem 4 A language L is regular iff there are homomorphisms h_0, h_1, h_2 such that $L = h_0(e(h_1, h_2))$ and the pair (h_1, h_2) has bounded balance on $e(h_1, h_2)$.

Proof Assume that $L = h_0(e(h_1, h_2))$ and the pair (h_1, h_2) has k-bounded balance for some $k > 0$. Then $e(h_1, h_2) = e_k(h_1, h_2)$ and by Lemma 1 $e(h_1, h_2)$ is regular. Thus also its homomorphic image $h_0(h_1, h_2)$ is regular.

Assume that L is regular. Then there is a (nondeterministic) finite automaton with only one final state q_f (possibly $q_0 = q_f$) $A = (K, \Sigma, \delta, q_0, \{q_f\})$ such that $L = L(A)$. Let $\pi = \{\langle q, a, p\rangle : p \in \delta(q, a), q, p \in K, a \in \Sigma\}$, π is a finite alphabet. Let $\overline{\pi} = \{\overline{\alpha} : \alpha \in \pi\}$ and $\Gamma = \pi \cup \overline{\pi} \cup \{\vdash, \dashv, \$\}$.

Let homomorphisms h_0, h_1, h_2 be given by the table

ξ	$\vdash$	$\langle q, a, p\rangle$	$\overline{\langle q, a, p\rangle}$	$\dashv$	\$
$h_0(\xi)$	ε	a	a	ε	ε
$h_1(\xi)$	$\vdash$	q	$\overline{q}$	$q_f\dashv$	$\overline{q}_f\dashv$
$h_2(\xi)$	$\vdash q_0$	$\overline{p}$	p	$\dashv$	$\dashv$

for all $a \in \Sigma$, $p,q \in K$ such that $p \in \delta(q, a)$.

It is easy to verify that $L = h_0(e(h_1, h_2))$. This fact is illustrated in Figure 2 where n is even. For computations of add length the "end marker" would be used.

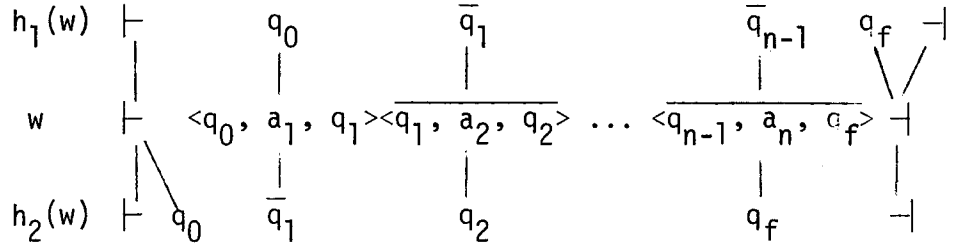

Figure 2

The condition of bounded balance in Theorem 4 is noneffective. However we can give the following stronger and effective conditions on the homomorphisms.

Theorem 5 A language L is regular iff there are homomorphisms h_0, h_1, h_2 such that $L = h_0(e(h_1, h_2))$ where h_0 is 1-limited on $e(h_1, h_2)$ (cf. [4]) and the pair (h_1, h_2) has 1-bounded balance on $e(h_1, h_2)$. If ε is not in L , then h_0 can be chosen to be a letter-to-letter homomorphism.

Proof The if-part follows by Theorem 4. Also the first statement of the only-if-part has been proved in the proof of Theorem 4 since h_0 erases only the first and the last of each string in $e(h_1, h_2)$ and, clearly, $B(w) = 1$ for all proper prefixes of strings in $e(h_1, h_2)$. It is easy to see that if $\varepsilon \notin L$ we can modify the construction of automaton A by merging $\vdash$ with $\langle q_0, a, q\rangle$ and $\dashv$ or \$ with $\langle q, a, qf\rangle$ for each $a \in \Sigma$ and $q \in K$. We get larger but still a finite alphabet. The homomorphism h_0 is then defined only on triples $\langle q, a, p\rangle$ and therefore letter to letter.

References

[1] J.A. Brzozowski, "Roots of star events", J. Assoc. Comput. Mach. 14 (1967), 466-477.

[2] K. Culik and I. Fris, "The decidability of the equivalence problem for DOL-systems", Information and Control 35 (1977), 20-39.

[3] K. Culik and A. Salomaa, "On the decidability of homomorphism equivalence for languages", J. Comput. System Sci., to appear.

[4] J.E. Hopcroft and J.D. Ullman, Formal Languages and Their Relations to Automata, Addison Wesley, 1969.

[5] A. Salomaa, Equality Sets for Homomorphisms of Free Monoids, manuscript.

[6] A. Salomaa, "DOL equivalence: The problem of iterated homomorphisms", Bulletin of EATCS, to appear.