ELSEVIER

Theoretical Computer Science 244 (2000) 267–270

Theoretical Computer Science

www.elsevier.com/locate/tcs

Note

A short solution for the HDT0L sequence equivalence problem

Juha Honkala[a,b,*,1]

[a] *Department of Mathematics, University of Turku, FIN-20014 Turku, Finland*
[b] *Turku Centre for Computer Science (TUCS), Lemminkäisenkatu 14, FIN-20520 Turku, Finland*

Received 1 October 1999; revised 1 December 1999; accepted 28 February 2000

Abstract

We give a solution for the HDT0L sequence equivalence problem which uses Hilbert's Basis Theorem but avoids the use of Makanin's algorithm or Hall's results about metabelian groups. © 2000 Elsevier Science B.V. All rights reserved.

Keywords: D0L system; DT0L system; Equivalence problem

An instance of the HDT0L sequence equivalence problem consists of finite alphabets Σ and Δ, two words $w_1, w_2 \in \Sigma^*$, morphisms $h_j, g_j : \Sigma^* \to \Sigma^*$, $1 \leqslant j \leqslant n$, and morphisms $h, g : \Sigma^* \to \Delta^*$. To solve the problem we have to determine whether or not

$$hh_{i_k} \ldots h_{i_1}(w_1) = gg_{i_k} \ldots g_{i_1}(w_2)$$

holds true for all $k \geqslant 0$, $1 \leqslant i_1, \ldots, i_k \leqslant n$. This problem is known to be decidable.

Theorem 1. *The HDT0L sequence equivalence problem is decidable.*

Theorem 1 is due to Culik II and Karhumäki [2]. The proof is based on Ehrenfeucht's Conjecture and Makanin's algorithm. A different proof was given in Ruohonen [4] by using the theory of metabelian groups. In this note we give a proof which uses ideas from both of these solutions but succeeds in avoiding the use of Makanin's algorithm and Hall's results about metabelian groups. Indeed, our proof below is essentially self-contained and can be presented in full in a few pages. The proof uses Hilbert's

[*] Correspondence address: Department of Mathematics, University of Turku, FIN-20014 Turku, Finland. *E-mail address:* jhonkala@utu.fi (J. Honkala).

[1] Research supported by the Academy of Finland.

Basis Theorem which can very easily be proved from first principles (see, e.g., [1] or [5]).

We proceed to prove Theorem 1. First we recall some facts from the proof of Ehrenfeucht's Conjecture. We use the approach in Harju and Karhumäki [3] where all details omitted below can be found.

Let $\mathbf{SL}(2,\mathbf{N})$ denote the special linear monoid consisting of the matrices $M\in\mathbf{N}^{2\times 2}$ such that $\det(M)=1$. By Lemma 6.1 in [3] the morphism $\mu:\{a,b\}^*\to\mathbf{SL}(2,\mathbf{N})$ defined by

$$\mu(a)=\begin{pmatrix}1&1\\0&1\end{pmatrix}\quad\text{and}\quad\mu(b)=\begin{pmatrix}1&0\\1&1\end{pmatrix}$$

is an isomorphism. Suppose $X=\{x_1,x_2,\ldots,x_k\}$ is a set of word variables. For each $x_i\in X$ introduce four new integer variables x_{ij}, $1\leqslant j\leqslant 4$, and denote $\bar{X}=\{x_{ij}\mid 1\leqslant i\leqslant k,\ 1\leqslant j\leqslant 4\}$. Furthermore, for each $i=1,2,\ldots,k$, define

$$X_i=\begin{pmatrix}x_{i1}&x_{i2}\\x_{i3}&x_{i4}\end{pmatrix}$$

and let $\mathbf{M}(\bar{X})$ be the submonoid of the matrix monoid $\mathbf{Z}[\bar{X}]^{2\times 2}$ generated by the matrices $X_1,X_2,\ldots,X_k$. By Lemma 6.2 in [3] the monoid morphism $\varphi:X^*\to\mathbf{M}(\bar{X})$ defined by

$$\varphi(x_i)=X_i,$$

$1\leqslant i\leqslant k$, is an isomorphism. Let

$$\mathscr{W}=\{W\mid W \text{ is a set of word equations over } X\}$$

and

$$\mathscr{I}=\{I\mid I \text{ is an ideal of } \mathbf{Z}[\bar{X}]\}.$$

Lemma 1. *There is a mapping which associates to each $W\in\mathscr{W}$ an ideal $I(W)\in\mathscr{I}$ such that if $I(W_1)=I(W_2)$ for $W_1,W_2\in\mathscr{W}$ then W_1 and W_2 have the same solutions.*

Proof. For a word $w\in X^*$ denote

$$\varphi(w)=\begin{pmatrix}P_{w1}&P_{w2}\\P_{w3}&P_{w4}\end{pmatrix},$$

where $P_{wj}\in\mathbf{Z}[\bar{X}]$, $1\leqslant j\leqslant 4$. If $u=v$ is a word equation over X, define the set $I(u,v)\subseteq\mathbf{Z}[\bar{X}]$ by

$$I(u,v)=\{P_{uj}-P_{vj}\mid 1\leqslant j\leqslant 4\}\cup\{x_{t1}x_{t4}-x_{t2}x_{t3}-1\mid 1\leqslant t\leqslant k\}.$$

If $W\in\mathscr{W}$, let $I(W)$ be the ideal of $\mathbf{Z}[\bar{X}]$ generated by the set

$$\bigcup_{u=v\in W} I(u,v).$$

Let now $h : X^* \to \Delta^*$ be a morphism where $\Delta = \{a, b\}$ and let $\bar{h} : \mathbf{M}(\bar{X}) \to \mathbf{SL}(2, \mathbf{N})$ be a monoid morphism such that $\mu h = \bar{h}\varphi$. Define the mapping $h' : \bar{X} \to \mathbf{N}$ by

$$\bar{h}(X_i) = \begin{pmatrix} h'(x_{i1}) & h'(x_{i2}) \\ h'(x_{i3}) & h'(x_{i4}) \end{pmatrix},$$

$1 \leqslant i \leqslant k$. The mapping h' extends in a unique way to a ring morphism $h' : \mathbf{Z}[\bar{X}] \to \mathbf{Z}$.

Now, suppose $u, v \in X^*$. Then we have $h(u) = h(v)$ if and only if $h'(P) = 0$ for all $P \in I(u, v)$ (see [3]). Consequently, h is a solution of $W \in \mathscr{W}$ if and only if $h'(P) = 0$ for all $P \in I(W)$. Suppose that $W_1, W_2 \in \mathscr{W}$ and $I(W_1) = I(W_2)$. Then h is a solution of W_1 if and only if h is a solution of W_2. This implies that W_1 and W_2 have the same solutions. □

The following lemma makes it possible to avoid the use of Makanin's algorithm.

Lemma 2. *Suppose W_j, $j \geqslant 1$, are finite sets of word equations over X such that*

$$W_1 \subseteq W_2 \subseteq \cdots \subseteq W_j \subseteq W_{j+1} \subseteq \cdots .$$

Then one can effectively find an integer n such that W_n and W_{n+1} are equivalent.

Proof. We have

$$I(W_1) \subseteq I(W_2) \subseteq \cdots \subseteq I(W_j) \subseteq I(W_{j+1}) \subseteq \cdots .$$

By Hilbert's Basis Theorem we have

$$I(W_m) = I(W_{m+1}) = \cdots$$

for some $m \geqslant 1$. Now if

$$I(W_{j+1}) \subseteq I(W_j),$$

we can verify this by showing that all generators of $I(W_{j+1})$ belong to $I(W_j)$. (Here it is essential that we have explicitly the generators of both ideals.) Therefore we proceed as follows. We try to show that $I(W_{j+1}) \subseteq I(W_j)$ for all consecutive integers $j \geqslant 1$ until we succeed in finding an integer n such that $I(W_{n+1}) \subseteq I(W_n)$. Then $I(W_{n+1}) = I(W_n)$ and Lemma 1 implies that W_n and W_{n+1} are equivalent. □

A reader familiar with computational algebraic geometry and commutative algebra will notice that to obtain a practical algorithm for Lemma 2, it is advisable to regard $I(W)$ as ideals of $\mathbf{Q}[\bar{X}]$ and apply Gröbner bases. However, Gröbner bases are not needed to prove the decidability stated in Lemma 2.

To conclude the proof of Theorem 1 it suffices to show that for any word $w \in \Sigma^*$ and any set $H = \{h_1, \ldots, h_t\}$ of endomorphisms of Σ^*, a test set of $H^*(w)$ can be effectively found. For an integer $s \geqslant 0$, denote

$$H^{\leqslant s}(w) = \{h_{i_k} \ldots h_{i_1}(w) \mid 0 \leqslant k \leqslant s, 1 \leqslant i_1, \ldots, i_k \leqslant t\}.$$

Now, let $\bar{\Sigma} = \{\bar{\sigma} \mid \sigma \in \Sigma\}$ be a new alphabet. For $j \geqslant 1$, denote

$$W_j = \{u = \bar{u} \mid u \in H^{\leqslant j}(w)\}.$$

Then Lemma 2 implies the effective existence of an integer n such that W_n and W_{n+1} are equivalent. It follows that $H^{\leqslant n}(w)$ is a test set of $H^{\leqslant n+1}(w)$. But then $H^{\leqslant n}(w)$ is a test set of $H^{*}(w)$. This concludes the proof of Theorem 1.

As a special case of Theorem 1 we obtain also a solution of the D0L sequence equivalence problem which uses Hilbert's Basis Theorem but avoids Makanin's algorithm.

References

[1] D. Cox, J. Little, D. O'Shea, Ideals, Varieties, and Algorithms, Springer, Berlin, 1997.

[2] K. Culik II, J. Karhumäki, A new proof for the D0L sequence equivalence problem and its implications, in: G. Rozenberg, A. Salomaa (Eds.), The Book of L, Springer, Berlin, 1986, pp. 63–74.

[3] T. Harju, J. Karhumäki, Morphisms, in: G. Rozenberg, A. Salomaa (Eds.), Handbook of Formal Languages, Vol. 1, Ch. 7, Springer, Berlin, 1997.

[4] K. Ruohonen, Equivalence problems for regular sets of word morphisms, in: G. Rozenberg, A. Salomaa (Eds.), The Book of L, Springer, Berlin, 1986, pp. 393–401.

[5] A. Salomaa, The Ehrenfeucht Conjecture: A proof for language theorists, Bull. EATCS 27 (1985) 71–82.