ISSN 0081-5438, Proceedings of the Steklov Institute of Mathematics, 2011, Vol. 274, pp. 116–123.

On Maximal Chains of Systems of Word Equations

Juhani Karhumäki [a] **and Aleksi Saarela** [a]

Received November 2010

Abstract—We consider systems of word equations and their solution sets. We discuss some fascinating properties of those, namely the size of a maximal independent set of word equations, and proper chains of solution sets of those. We recall the basic results, extend some known results and formulate several fundamental problems of the topic.

DOI: 10.1134/S0081543811060083

1. INTRODUCTION

Theory of word equations is a fundamental part of combinatorics on words. It is a challenging topic of its own which has a number of connections and applications, e.g., in pattern unification and group representations. There have also been several fundamental achievements in the theory over the last few decades.

Decidability of the existence of a solution of a given word equation is one fundamental result due to Makanin [14]. This is in contrast to the same problem on Diophantine equations, which is undecidable [15]. Although the complexity of the above *satisfiability problem* for word equations is not known, a nontrivial upper bound has been proved: it is in PSPACE [17].

Another fundamental property of word equations is the so-called *Ehrenfeucht compactness property*. It guarantees that any system of word equations is equivalent to some of its finite subsystems. The proofs (see [1] and [6]) are based on a transformation of word equations into Diophantine equations and then an application of Hilbert's basis theorem. Although we have this finiteness property, we do not know any upper bound, if it exists, for the size of an equivalent subsystem in terms of the number of unknowns. And this holds even in the case of three unknown systems of equations. In free monoids an equivalent formulation of the compactness property is that each *independent* system of word equations is finite, independent meaning that the system is not equivalent to any of its proper subsystems. We analyze in this paper the size of the maximal independent systems of word equations.

As a related problem we define the notion of *decreasing chains* of word equations. Intuitively, this asks how long chains of word equations exist such that the set of solutions always properly diminishes when a new element of the chain is taken into the system. Or more intuitively, how many proper constraints we can define such that each constraint reduces the set of words satisfying these constraints. It is essentially the above compactness property which guarantees that these chains are finite.

Another fundamental property of word equations is the result of Hmelevskii [10] stating that for each word equation with three unknowns its solution set is *finitely parameterizable*. This intriguing result is not directly related to our considerations, but it gives, we believe, a strong explanation and support to our view that our open problems, even the simplest looking ones, are not trivial. Hmelevskii's argumentation is simplified in the extended abstract [13], and used in [18] to show

[a] Department of Mathematics and Turku Centre for Computer Science TUCS, University of Turku, FI-20014 Turku, Finland.

E-mail addresses: karhumak@utu.fi (J. Karhumäki), amsaar@utu.fi (A. Saarela).

that the satisfiability problem for three unknown equations is in NP. A full version of these two conference articles is in preparation.[1]

The goal of this note is to analyze the above maximal independent systems of equations and maximal decreasing chains of word equations, as well as search for their relations. An essential part is to propose open problems on this area. The most fundamental problem asks whether the maximal independent system of word equations with n unknowns is bounded by some function of n. Amazingly, the same problem is open for three unknown equations, although we do not know larger than three equation systems in this case.

2. SYSTEMS AND CHAINS OF WORD EQUATIONS

The topics of this paper are independent systems and chains of equations in semigroups. We are mostly interested in free monoids; in this case the equations are constant-free word equations. We present some questions about the sizes of such systems and chains, state existing results, give some new ones, and list open problems.

Let S be a semigroup and Ξ be an alphabet of variables. We consider equations $U = V$, where $U, V \in \Xi^+$. A morphism $h\colon \Xi^+ \to S$ is a *solution* of this equation if $h(U) = h(V)$. (If S is a monoid, we can use Ξ^* instead of Ξ^+.)

A system of equations is *independent* if it is not equivalent to any of its proper subsystems. In other words, equations E_i form an independent system of equations if for every i there is a morphism h_i which is not a solution of E_i but which is a solution of all the other equations. This definition works for both finite and infinite systems of equations.

We define *decreasing chains* of equations. A finite sequence of equations $E_1, \ldots, E_m$ is a decreasing chain if for every $i \in \{0, \ldots, m-1\}$ the system $E_1, \ldots, E_i$ is inequivalent to the system $E_1, \ldots, E_{i+1}$. An infinite sequence of equations $E_1, E_2, \ldots$ is a decreasing chain if for every $i \geq 0$ the system $E_1, \ldots, E_i$ is inequivalent to the system $E_1, \ldots, E_{i+1}$.

Similarly we define *increasing chains* of equations. A sequence of equations $E_1, \ldots, E_m$ is an increasing chain if for every $i \in \{1, \ldots, m\}$ the system $E_i, \ldots, E_m$ is inequivalent to the system $E_{i+1}, \ldots, E_m$. An infinite sequence of equations $E_1, E_2, \ldots$ is an increasing chain if for every $i \geq 1$ the system $E_i, E_{i+1}, \ldots$ is inequivalent to the system $E_{i+1}, E_{i+2}, \ldots$.

Now $E_1, \ldots, E_m$ is an increasing chain if and only if $E_m, \ldots, E_1$ is a decreasing chain. However, for infinite chains these concepts are essentially different. Note that a chain can be both decreasing and increasing, for example, if the equations form an independent system.

We will consider the *maximal* sizes of independent systems of equations and chains of equations. If the number of unknowns is n, then the maximal size of an independent system is denoted by $\mathrm{IS}(n)$. We use two special symbols ub and ∞ for the infinite cases: if there are infinite independent systems, then $\mathrm{IS}(n) = \infty$, and if there are only finite but unboundedly large independent systems, then $\mathrm{IS}(n) = \mathrm{ub}$. We extend the order relation of numbers to these symbols: $k < \mathrm{ub} < \infty$ for every integer k. Similarly the maximal size of a decreasing chain is denoted by $\mathrm{DC}(n)$, and the maximal size of an increasing chain by $\mathrm{IC}(n)$.

Often we are interested in the finiteness of $\mathrm{DC}(n)$, or its asymptotic behaviour when n grows. However, if we are interested in the exact value of $\mathrm{DC}(n)$, then some technical remarks about the definition are in order. First, the case $i = 0$ means that there is a solution which is not a solution of the first equation E_1; that is, E_1 cannot be a trivial equation like $U = U$. If this condition was removed, then we could always add a trivial equation in the beginning, and $\mathrm{DC}(n)$ would be increased by one. Second, we could add the requirement that there must be a solution which is a solution of all the equations $E_1, \ldots, E_m$, and the definition would remain the same in the case of free

[1] J. Karhumäki and A. Saarela, "A Reproof of Hmelevskii's Theorem and the Satisfiability of Three Unknown Equations" (in preparation).

monoids. However, if we consider free semigroups, then this addition would change the definition, because then E_m could not be an equation with no solutions, like $xx = x$ in free semigroups. This would decrease $\mathrm{DC}(n)$ by one.

3. RELATIONS BETWEEN SYSTEMS AND CHAINS

Independent systems of equations are a well-known topic (see, e.g., [8]). Chains of equations have been studied less, so we prove here some elementary results about them. The following theorem states the most basic relations between IS, DC and IC.

Theorem 3.1. *For every* n, $\mathrm{IS}(n) \le \mathrm{DC}(n), \mathrm{IC}(n)$. *If* $\mathrm{DC}(n) < \mathrm{ub}$ *or* $\mathrm{IC}(n) < \mathrm{ub}$, *then* $\mathrm{DC}(n) = \mathrm{IC}(n)$.

Proof. Every independent system of equations is also a decreasing and increasing chain of equations, regardless of the order of the equations. This means that $\mathrm{IS}(n) \le \mathrm{DC}(n), \mathrm{IC}(n)$.

A finite sequence of equations is a decreasing chain if and only if the reverse of this sequence is an increasing chain. Thus $\mathrm{DC}(n) = \mathrm{IC}(n)$ if $\mathrm{DC}(n) < \mathrm{ub}$ or $\mathrm{IC}(n) < \mathrm{ub}$. □

A semigroup has the *compactness property* if every system of equations has an equivalent finite subsystem. Many results on the compactness property are collected in [8]. In terms of chains, the compactness property turns out to be equivalent to the property that every decreasing chain is finite.

Theorem 3.2. *A semigroup has the compactness property if and only if* $\mathrm{DC}(n) \le \mathrm{ub}$ *for every* n.

Proof. Assume first that the compactness property holds. Let $E_1, E_2, \dots$ be an infinite decreasing chain of equations. As a system of equations, it is equivalent to some finite subsystem $E_{i_1}, \dots, E_{i_k}$, where $i_1 < \dots < i_k$. But now $E_1, \dots, E_{i_k}$ is equivalent to $E_1, \dots, E_{i_k+1}$. This is a contradiction.

Assume then that $\mathrm{DC}(n) \le \mathrm{ub}$. Let $E_1, E_2, \dots$ be an infinite system of equations. If there is an index N such that $E_1, \dots, E_i$ is equivalent to $E_1, \dots, E_{i+1}$ for all $i \ge N$, then the whole system is equivalent to $E_1, \dots, E_N$. If there is no such index, then let $i_1 < i_2 < \dots$ be all indexes such that $E_1, \dots, E_{i_k}$ is not equivalent to $E_1, \dots, E_{i_k+1}$. But then $E_{i_1}, E_{i_2}, \dots$ is an infinite decreasing chain, which is a contradiction. □

The next example shows that the values of IS, DC and IC can differ significantly.

Example 3.3. We give an example of a monoid where $\mathrm{IS}(1) = 1$, $\mathrm{DC}(1) = \mathrm{ub}$ and $\mathrm{IC}(1) = \infty$. The monoid is

$$\langle a_1, a_2, \dots \mid a_i a_j = a_j a_i,\ a_i^{i+1} = a_i^i \rangle.$$

Now every equation on one unknown is of the form $x^i = x^j$. If $i < j$, then this is equivalent to $x^i = x^{i+1}$. So all nontrivial equations are, up to equivalence,

$$x = 1, \quad x^2 = x, \quad x^3 = x^2, \quad \dots,$$

and these have strictly increasing solution sets. Thus $\mathrm{IC}(1) = \infty$, $\mathrm{DC}(1) = \mathrm{ub}$ and $\mathrm{IS}(1) = 1$.

4. FREE MONOIDS

From now on we will consider free monoids and semigroups. The bounds related to free monoids are denoted by IS, DC and IC, and the bounds related to free semigroups, by IS_+, DC_+ and IC_+.

We give some definitions related to word equations and make some easy observations about the relations between maximal sizes of independent systems and chains, assuming these are finite.

A solution h is *periodic* if there exists a $t \in S$ such that every $h(x)$, where $x \in \Xi$, is a power of t. Otherwise h is *nonperiodic*. An equation $U = V$ is *balanced* if every variable occurs as many times in U as in V.

The maximal size of an independent system in a free monoid having a nonperiodic solution is denoted by $\mathrm{IS}'(n)$. The maximal size of a decreasing chain having a nonperiodic solution is denoted by $\mathrm{DC}'(n)$. Similar notation can be used for free semigroups.

Every independent system of equations $E_1, \dots, E_m$ is also a chain of equations, regardless of the order of the equations. If the system has a nonperiodic solution, then we can add an equation that forces the variables to commute. If the equations in the system are also balanced, then we can add equations $x_i = 1$ for all variables $x_1, \dots, x_n$ and thus get a chain of length $m + n + 1$. If they are not balanced, then we can add at least one of these equations.

In all cases we obtain the inequalities $\mathrm{IS}'(n) \le \mathrm{IS}(n) \le \mathrm{IS}'(n) + 1$ and $\mathrm{DC}'(n) + 2 \le \mathrm{DC}(n) \le \mathrm{DC}'(n) + n + 1$, as well as $\mathrm{IS}(n) + 1 \le \mathrm{DC}(n)$ and $\mathrm{IS}'(n) \le \mathrm{DC}'(n)$. In the case of free semigroups we derive similar inequalities. Thus IS' and DC' are basically the same as IS and DC, if we are only interested in their finiteness or asymptotic growth.

It was conjectured by Ehrenfeucht in a language theoretic setting that the compactness property holds for free monoids. This conjecture was reformulated in terms of equations in [2], and it was proved independently by Albert and Lawrence [1] and by Guba [6].

Theorem 4.1. $\mathrm{DC}(n) \le \mathrm{ub}$, *and hence also* $\mathrm{IS}(n) \le \mathrm{ub}$.

The proofs are based on Hilbert's basis theorem. The compactness property means that $\mathrm{DC}(n) \le \mathrm{ub}$ for every n. No better upper bounds are known when $n > 2$. Even the seemingly simple question about the size of $\mathrm{IS}'(3)$ is still completely open; the only thing that is known is that $2 \le \mathrm{IS}'(3) \le \mathrm{ub}$. The lower bound is given by the example $xyz = zyx$, $xyyz = zyyx$.

5. THREE AND FOUR UNKNOWNS

The cases of three and four variables have been studied in [3]. The article gives examples showing that $\mathrm{IS}'_+(3) \ge 2$, $\mathrm{DC}_+(3) \ge 6$, $\mathrm{IS}'_+(4) \ge 3$ and $\mathrm{DC}_+(4) \ge 9$. We are able to give better bounds for $\mathrm{DC}_+(3)$ and $\mathrm{DC}(4)$.

First we assume that there are three unknowns x, y, z. There are trivial examples of independent systems of three equations, for example, $x^2 = y$, $y^2 = z$, $z^2 = x$, so $\mathrm{IS}_+(3) \ge 3$. There are also easy examples of independent pairs of equations having a nonperiodic solution, like $xyz = zyx$, $xyyz = zyyx$, so $\mathrm{IS}'_+(3) \ge 2$. Amazingly, no other bounds are known for $\mathrm{IS}_+(3)$, $\mathrm{IS}'_+(3)$, $\mathrm{IS}(3)$ or $\mathrm{IS}'(3)$.

The following chain of equations shows that $\mathrm{DC}(3) \ge 7$:

$$\begin{aligned}
xyz &= zxy, & \qquad & x = a,\ y = b,\ z = abab,\\
xyxzyz &= zxzyxy, && x = a,\ y = b,\ z = ab,\\
xz &= zx, && x = a,\ y = b,\ z = 1,\\
xy &= yx, && x = a,\ y = a,\ z = a,\\
x &= 1, && x = 1,\ y = b,\ z = a,\\
y &= 1, && x = 1,\ y = 1,\ z = a,\\
z &= 1, && x = 1,\ y = 1,\ z = 1.
\end{aligned}$$

Here the second column gives a solution which is not a solution of the equation on the next row but is a solution of all the preceding equations. Also $\mathrm{DC}_+(3) \ge 7$, as shown by the chain

$$\begin{aligned}
xxyz &= zxyx, & \qquad & x = a,\ y = b,\ z = aabaaba,\\
xxyxzyz &= zzyxxyx, && x = a,\ y = b,\ z = aaba,
\end{aligned}$$

$$\begin{aligned} xz &= zx, && x = a,\ y = b,\ z = a,\\ xy &= yx, && x = a,\ y = aa,\ z = a,\\ x &= y, && x = a,\ y = a,\ z = aa,\\ x &= z, && x = a,\ y = a,\ z = a,\\ xx &= x, && \text{no solutions.} \end{aligned}$$

If there are three variables, then every independent pair of equations having a nonperiodic solution consists of balanced equations (see [9]). It follows that $\mathrm{IS}'(3) + 4 \le \mathrm{DC}(3)$. There are also some other results about the structure of equations in independent systems in three unknowns (see [4] and [5]).

If we add a fourth unknown t, then we can trivially extend any independent system by adding the equation $t = x$. This gives $\mathrm{IS}_+(4) \ge 4$ and $\mathrm{IS}'_+(4) \ge 3$. For chains the improvements are nontrivial. The following chain of equations shows that $\mathrm{DC}(4) \ge 12$:

$$\begin{aligned} xyz &= zxy, && x = a,\ y = b,\ z = abab,\ t = a,\\ xyt &= txy, && x = a,\ y = b,\ z = abab,\ t = abab,\\ xyxzyz &= zxzyxy, && x = a,\ y = b,\ z = ab,\ t = abab,\\ xyxtyt &= txtyxy, && x = a,\ y = b,\ z = ab,\ t = ab,\\ xyxztyzt &= ztxztyxy, && x = a,\ y = b,\ z = ab,\ t = 1,\\ xz &= zx, && x = a,\ y = b,\ z = 1,\ t = ab,\\ xt &= tx, && x = a,\ y = b,\ z = 1,\ t = 1,\\ xy &= yx, && x = a,\ y = a,\ z = a,\ t = a,\\ x &= 1, && x = 1,\ y = a,\ z = a,\ t = a,\\ y &= 1, && x = 1,\ y = 1,\ z = a,\ t = a,\\ z &= 1, && x = 1,\ y = 1,\ z = 1,\ t = a,\\ t &= 1, && x = 1,\ y = 1,\ z = 1,\ t = 1. \end{aligned}$$

The next theorem sums up the new bounds given in this section.

Theorem 5.1. $\mathrm{DC}_+(3) \ge 7$ *and* $\mathrm{DC}(4) \ge 12$.

6. LOWER BOUNDS

In [12] it is proved that $\mathrm{IS}(n) = \Omega(n^4)$ and $\mathrm{IS}_+(n) = \Omega(n^3)$. The former is proved by a construction that uses $n = 10m$ variables and gives a system of m^4 equations. Thus $\mathrm{IS}(n)$ is asymptotically at least $n^4/10\,000$. We present here a slightly modified version of this construction. By "reusing" some of the unknowns we get a bound that is asymptotically $n^4/1536$.

Theorem 6.1. *If* $n = 4m$, *then* $\mathrm{IS}'(n) \ge m^2(m-1)(m-2)/6$.

Proof. We use unknowns x_i, y_i, z_i, t_i, where $1 \le i \le m$. The equations in the system are

$$E(i,j,k,l)\colon \quad x_i x_j x_k y_i y_j y_k z_i z_j z_k t_l = t_l x_i x_j x_k y_i y_j y_k z_i z_j z_k,$$

where $i,j,k,l \in \{1,\dots,m\}$ and $i < j < k$. If $i,j,k,l \in \{1,\dots,m\}$ and $i < j < k$, then

$$x_r = \begin{cases} ab & \text{if } r \in \{i,j,k\},\\ 1 & \text{otherwise,} \end{cases} \qquad y_r = \begin{cases} a & \text{if } r \in \{i,j,k\},\\ 1 & \text{otherwise,} \end{cases}$$

$$z_r = \begin{cases} ba & \text{if } r \in \{i,j,k\}, \\ 1 & \text{otherwise,} \end{cases} \qquad t_r = \begin{cases} ababa & \text{if } r = l, \\ 1 & \text{otherwise} \end{cases}$$

is not a solution of $E(i,j,k,l)$, but is a solution of all the other equations. Thus the system is independent. □

The idea behind this construction (both the original and the modified) is that $(ababa)^k = (ab)^k a^k (ba)^k$ holds for $k < 3$, but not for $k = 3$. It was noted in [16] that if we could find words u_i such that $(u_1 \dots u_m)^k = u_1^k \dots u_m^k$ holds for $k < K$, but not for $k = K$, then we could prove that $\mathrm{IS}(n) = \Omega(n^{K+1})$. However, it has been proved that such words do not exist for $K \ge 5$ (see [11]), and conjectured that such words do not exist for $K = 4$.

For small values of n it is better to use ideas from the constructions showing that $\mathrm{DC}(3) \ge 7$ and $\mathrm{DC}(4) \ge 12$. This gives $\mathrm{IS}'(n) \ge (n^2 - 5n + 6)/2$ and $\mathrm{DC}(n) \ge (n^2 + 3n - 4)/2$. The equations in the system are

$$xyxz_iz_jyz_iz_j = z_iz_jxz_iz_jyxy,$$

where $i, j \in \{1, \dots, n-2\}$ and $i < j$. The equations in the chain are

$$\begin{aligned} xyz_k &= z_kxy, \\ xyxz_kyz_k &= z_kxz_kyxy, \\ xyxz_iz_jyz_iz_j &= z_iz_jxz_iz_jyxy, \\ xz_k &= z_kx, \\ xy &= yx, \\ x &= 1, \\ y &= 1, \\ z_k &= 1, \end{aligned}$$

where $i, j \in \{1, \dots, n-2\}$, $i < j$ and $k \in \{1, \dots, n-2\}$. Here we should first take the equations on the first row in some order, then the equations on the second row in some order, and so on.

We conclude this section by mentioning a related question. It is well known that any nontrivial equation in n variables forces a defect effect; that is, the values of the variables in any solution can be expressed as products of $n-1$ words (see [7] for a survey on the defect effect). If a system has only periodic solutions, then the system can be said to force a maximal defect effect, so $\mathrm{IS}'(n)$ is the maximal size of an independent system not doing that. But how large can an independent system be if it forces only the minimal defect effect, that is, the system has a solution in which the variables cannot be expressed as products of $n-2$ words? In [12] it is proved that there are such systems of size $\Omega(n^3)$ in free monoids and of size $\Omega(n^2)$ in free semigroups. Again, no upper bounds are known.

7. CONCLUDING REMARKS AND OPEN PROBLEMS

To summarize, we list a few fundamental open problems about systems and chains of equations in free monoids.

Question 1. Is IS(3) finite?

Question 2. Is DC(3) finite?

Question 3. Is IS(n) finite for every n?

Question 4. Is DC(n) finite for every n?

A few remarks on these questions are in order. First we know that each of these values is at most ub. Second, if the answer to any of the questions is "yes," a natural further question is: What is an upper bound for this value, or more sharply, what is the best upper bound, that is, the exact value? For the lower bounds the best what is known, according to our knowledge, is the following:

1. $\mathrm{IS}(3) \geq 3$.
2. $\mathrm{DC}(3) \geq 7$.
3. $\mathrm{IS}(n) = \Omega(n^4)$.
4. $\mathrm{DC}(n) = \Omega(n^4)$.

A natural sharpening of Question 3 (and 4) asks whether these values are exponentially bounded.

A related question to Question 1 is the following amazing open problem from [2] (see, e.g., [3] and [4] for an extensive study of it):

Question 5. Does there exist an independent system of three equations with three unknowns having a nonperiodic solution?

As a summary we make the following remarks. As we see it, Question 3 is a really fundamental question on word equations or even on combinatorics on words as a whole. It is very intriguing because we do not know the answer even in the case of three unknowns (Question 1). This becomes really amazing when we recall that still the best known lower bound is only 3!

To conclude, we have considered equations over word monoids and semigroups. All of the questions can be stated in any semigroup, and the results would be different. For example, in commutative monoids the compactness property (Theorem 4.1) holds, but in this case the value of the maximal independent system of equations is ub (see [12]).

ACKNOWLEDGMENTS

This work was supported by the Academy of Finland under grant 121419.

REFERENCES

1. M. H. Albert and J. Lawrence, "A Proof of Ehrenfeucht's Conjecture," Theor. Comput. Sci. **41** (1), 121–123 (1985).
2. K. Culik II and J. Karhumäki, "Systems of Equations over a Free Monoid and Ehrenfeucht's Conjecture," Discrete Math. **43** (2–3), 139–153 (1983).
3. E. Czeizler, "Multiple Constraints on Three and Four Words," Theor. Comput. Sci. **391** (1–2), 14–19 (2008).
4. E. Czeizler and J. Karhumäki, "On Non-periodic Solutions of Independent Systems of Word Equations over Three Unknowns," Int. J. Found. Comput. Sci. **18** (4), 873–897 (2007).
5. E. Czeizler and W. Plandowski, "On Systems of Word Equations over Three Unknowns with at Most Six Occurrences of One of the Unknowns," Theor. Comput. Sci. **410** (30–32), 2889–2909 (2009).
6. V. S. Guba, "Equivalence of Infinite Systems of Equations in Free Groups and Semigroups to Finite Subsystems," Mat. Zametki **40** (3), 321–324 (1986) [Math. Notes **40**, 688–690 (1986)].
7. T. Harju and J. Karhumäki, "Many Aspects of Defect Theorems," Theor. Comput. Sci. **324** (1), 35–54 (2004).
8. T. Harju, J. Karhumäki, and W. Plandowski, "Independent Systems of Equations," in M. Lothaire, *Algebraic Combinatorics on Words* (Cambridge Univ. Press, Cambridge, 2002), Ch. 13, pp. 443–472.
9. T. Harju and D. Nowotka, "On the Independence of Equations in Three Variables," Theor. Comput. Sci. **307** (1), 139–172 (2003).
10. Ju. I. Hmelevskiĭ, *Equations in Free Semigroups* (Nauka, Moscow, 1971), Tr. Mat. Inst. im. V.A. Steklova, Akad. Nauk SSSR **107** [Proc. Steklov Inst. Math. **107** (1976)].
11. Š. Holub, "Local and Global Cyclicity in Free Semigroups," Theor. Comput. Sci. **262** (1–2), 25–36 (2001).
12. J. Karhumäki and W. Plandowski, "On the Size of Independent Systems of Equations in Semigroups," Theor. Comput. Sci. **168** (1), 105–119 (1996).
13. J. Karhumäki and A. Saarela, "An Analysis and a Reproof of Hmelevskii's Theorem," in *Developments in Language Theory: Proc. 12th Int. Conf. DLT 2008, Kyoto, Sept. 16–19, 2008* (Springer, Berlin, 2008), Lect. Notes Comput. Sci. **5257**, pp. 467–478.

14. G. S. Makanin, “The Problem of Solvability of Equations in a Free Semigroup,” Mat. Sb. **103** (2), 147–236 (1977) [Math. USSR, Sb. **32**, 129–198 (1977)].
15. Yu. V. Matiyasevich, “Enumerable Sets Are Diophantine,” Dokl. Akad. Nauk SSSR **191** (2), 279–282 (1970) [Sov. Math., Dokl. **11**, 354–358 (1970)].
16. W. Plandowski, “Test Sets for Large Families of Languages,” in *Developments in Language Theory: Proc. 7th Int. Conf. DLT 2003, Szeged, Hungary, July 7–11, 2003* (Springer, Berlin, 2003), Lect. Notes Comput. Sci. **2710**, pp. 75–94.
17. W. Plandowski, “Satisfiability of Word Equations with Constants Is in PSPACE,” J. ACM **51** (3), 483–496 (2004).
18. A. Saarela, “On the Complexity of Hmelevskii’s Theorem and Satisfiability of Three Unknown Equations,” in *Developments in Language Theory: Proc. 13th Int. Conf. DLT 2009, Stuttgart, June 30–July 3, 2009* (Springer, Berlin, 2009), Lect. Notes Comput. Sci. **5583**, pp. 443–453.

This article was submitted by the authors in English