Independent Systems of Word Equations: From Ehrenfeucht to Eighteen

Aleksi Saarela[0000−0002−6636−2317]

Department of Mathematics and Statistics, University of Turku, 20014 Turku, Finland, amsaar@utu.fi

Abstract. A system of equations is called independent if it is not equivalent to any of its proper subsystems. We consider the following decades-old question: If we fix the number of variables, then what is the maximal size of an independent system of constant-free word equations? This can be easily answered in the trivial cases of one and two variables, but all other cases remain open, even the three-variable case, where the conjectured answer is as small as three. We survey some historical as well as more recent results related to this question, starting with the one known as Ehrenfeucht's compactness property: Every infinite system is equivalent to a finite subsystem, and consequently an independent system cannot be infinite. We also discuss several variations and related questions on word equations. Finally, we pay special attention to the following result from 2018: The maximal size of an independent system of three-variable equations is at most 18. This is the first such finite upper bound, but hopefully it will not be the last.

Keywords: Combinatorics on words · word equation · independent system

1 Introduction

The following two questions are among the most important open problems in combinatorics on words: Is the satisfiability problem of word equations (that is, the problem of deciding whether a given word equation has a solution) NP-complete? What is the maximal size of an independent system of constant-free n-variable word equations? The satisfiability problem is known to be NP-hard, it was proved to be decidable by Makanin [20] and in PSPACE by Plandowski [25], and a simpler PSPACE algorithm was given by Jeż [16]. The question about the maximal sizes of independent systems, on the other hand, is the topic of this article.

We start in Section 2 by giving formal definitions and simple examples of word equations and independent systems. We continue in Section 3 by describing Ehrenfeucht's conjecture and the idea of its proof. In Sections 4 and 5, we introduce the main open questions and survey some results related to them. Rather than just presenting the current state-of-the-art, we try to give a rough picture of how the research has progressed from 1983 to 2018. We conclude in Section 6 by describing some variations of the main questions.

2 Preliminaries

Let Ξ be an alphabet of n variables and Σ an alphabet of constants. An *n-variable word equation* is a pair $(u, v) \in (\Xi \cup \Sigma)^* \times (\Xi \cup \Sigma)^*$. The equation (u, v) is *constant-free* if $u, v \in \Xi^*$, and it is *trivial* if $u = v$. The *length* of (u, v) is $|uv|$ and it is denoted by $|(u, v)|$.

A constant-preserving morphism $h : (\Xi \cup \Sigma)^* \to \Sigma^*$ is a *solution* of an equation (u, v) if $h(u) = h(v)$. The set of all solutions of (u, v) is denoted by $\mathrm{Sol}((u, v))$. The morphism h is *periodic* if $h(\Xi) \subseteq w^*$ for some $w \in \Sigma^*$.

Example 1. First, let $\Xi = \{x\}$ and $\Sigma = \{a, b\}$. The equation $(xaxbab, abaxbx)$ has two solutions f and g defined by $f(x) = \varepsilon$ and $g(x) = ab$:

$$f(xaxbab) = \varepsilon \cdot a \cdot \varepsilon \cdot bab = abab = aba \cdot \varepsilon \cdot b \cdot \varepsilon = f(abaxbx),$$
$$g(xaxbab) = ab \cdot a \cdot ab \cdot bab = abaabbab = aba \cdot ab \cdot b \cdot ab = g(abaxbx).$$

Then, let $\Xi = \{x, y, z\}$ and $\Sigma = \{a, b\}$. The constant-free equation (xyz, zyx) has infinitely many solutions. For example, for all $p, q \in \Sigma^*$ and $i, j, k \geq 0$, the morphism h defined by $h(x) = (pq)^i p$, $h(y) = (qp)^j q$, $h(z) = (pq)^k p$ is a solution of this equation because

$$h(xyz) = (pq)^i p \cdot (qp)^j q \cdot (pq)^k p = (pq)^k p \cdot (qp)^j q \cdot (pq)^i p = h(zyx).$$

A *system of equations* is a set of equations. A morphism is a solution of a system if it is a solution of every equation in the system. The set of solutions of a system S is denoted by $\mathrm{Sol}(S)$, so $\mathrm{Sol}(S) = \bigcap_{E \in S} \mathrm{Sol}(E)$. Two equations or systems are *equivalent* if they have the same set of solutions. A subset (proper subset) of a system S is called a *subsystem* (*proper subsystem*, respectively) of S. A system is *independent* if it is not equivalent to any of its proper subsystems. Clearly, a system S is independent if and only if for every $E \in S$, there exists a morphism h such that $h \notin \mathrm{Sol}(E)$ but $h \in \mathrm{Sol}(E')$ for all $E' \in S \setminus \{E\}$.

Example 2. Let $\Xi = \{x, y, z\}$ and $\Sigma = \{a, b\}$. The system of equations $S = \{(xyz, zyx), (xyyz, zyyx)\}$ is independent and has a nonperiodic solution h defined by $h(x) = a$, $h(y) = b$, $h(z) = a$. To see independence, note that S is not equivalent to (xyz, zyx) because the morphism h defined by $h(x) = a$, $h(y) = b$, $h(z) = aba$ is a solution of (xyz, zyx) but not of S, and S is not equivalent to $(xyyz, zyyx)$ because the morphism h defined by $h(x) = a$, $h(y) = b$, $h(z) = abba$ is a solution of $(xyyz, zyyx)$ but not of S.

3 Ehrenfeucht's conjecture

A subset K of a language L is a *test set* of L if there does not exist morphisms f, g such that $f(w) = g(w)$ for all $w \in K$ but $f(w) \neq g(w)$ for some $w \in L$. Ehrenfeucht conjectured at the beginning of the 1970s that every language has a finite test set. In 1983, Culik and Karhumäki [3] proved that this conjecture can

be equivalently formulated as follows: Every system of word equations is equivalent to a finite subsystem. The conjecture was proved in 1985 by Albert and Lawrence [1] (an independent proof was given by Guba [9]), giving the following theorem, still known as *Ehrenfeucht's conjecture* or *Ehrenfeucht's compactness property*.

Theorem 1. *Every system of word equations is equivalent to a finite subsystem.*

The idea of the proof can be described as follows: Words can be turned into numbers by interpreting them as representations of integers in a suitable k-ary number system. A word equation can then be turned into a multivariate polynomial that has roots corresponding to all solutions of the word equation. A system of word equations then corresponds to a polynomial ideal. If an infinite system of word equations is not equivalent to any of its finite subsystems, then we get an infinite chain $I_1 \subsetneq I_2 \subsetneq I_3 \subsetneq \cdots$ of polynomial ideals, which contradicts a result from commutative algebra called *Hilbert's basis theorem*.

4 Size of independent systems

In 1983, Culik and Karhumäki [3] asked the following question, which is still open.

Question 1. Is it true that every independent system of three constant-free three-variable equations has only periodic solutions?

The following even more difficult question is also well-known.

Question 2. Let $\mathrm{IS}(n)$ be the maximal size of an independent system of constant-free n-variable word equations (or $\mathrm{IS}(n) = \infty$ if there is no maximum). How large is $\mathrm{IS}(n)$?

It is easy to prove that $\mathrm{IS}(1) = 1$, $\mathrm{IS}(2) = 2$, and $\mathrm{IS}(n) \geq n$ for all n. A positive answer to Question 1 would imply $\mathrm{IS}(3) = 3$. It follows from Theorem 1 that an independent system of word equations cannot be infinite. However, this does not prove even $\mathrm{IS}(n) < \infty$, because in principle it might be possible that there are arbitrarily large finite independent systems.

Some lower bounds better than the trivial $\mathrm{IS}(n) \geq n$ are known. In 1994, Karhumäki and Plandowski [17] gave examples of independent systems showing that $\mathrm{IS}(n) = \Omega(n^4)$. The hidden constant was improved by Karhumäki and Saarela [18], but no examples larger than $\Theta(n^4)$ have been found. The constructions can be said to be based on the fact that $(ababa)^k = (ab)^k a (ba)^k$ for all $k \leq 2$ but not for any $k \geq 3$. Plandowski [24] pointed out that if we could find words $u_0, \ldots, u_m$ such that $u_0^k = u_1^k \ldots u_m^k$ for all $k \leq 3$ but not for any $k \geq 4$, then it would follow that $\mathrm{IS}(n) = \Omega(n^5)$. However, these kinds of equalities had been studied before, for example in [10] and [13], and it was suspected that such words $u_0, \ldots, u_m$ do not exist. This was proved later, as will be discussed in Section 5.

In the three-variable case, some results restricting the form of equations in an independent system have been proved. For example, an equation (u, v) is called *balanced* if every variable occurs in u as many times as in v, and Harju and Nowotka [12] proved that if a system of two constant-free three-variable equations is independent and has a nonperiodic solution, then both equations are balanced. For more results, see [5] and [6].

We can try to find upper bounds that depend on the length of the equations in the system. If $E_1, \ldots, E_k$ is an independent system, then trivially k is at most exponential with respect to $\max\{|E_1|, \ldots, |E_k|\}$, simply because the number of equations of a certain length is exponential. The first nontrivial bound of a similar type was proved by Saarela [26]: In the case of constant-free three-variable equations, k is at most quadratic with respect to $\min\{|E_1|, \ldots, |E_k|\}$. This bound was improved to a linear one by Holub and Žemlička [14]. These two results were proved with the help of polynomials and linear algebra, so there are some similarities to the proof of Theorem 1, but the way in which polynomials were used is very different. The linear bound has been improved since then, as we shall see in Section 5, but the articles [26] and [14] contain also some results about n-variable equations that are still the best ones known.

5 Recent results

In 2016, Nowotka and Saarela [21] found a new way to apply an old characterization of three-generator subsemigroups of a free semigroup by Budkina and Markov [2] (or alternatively a similar result by Spehner [29,30]) to analyze independent systems. Specifically, they found that an upper bound for the number of solutions a one-variable word equation with only finitely many solutions can have implies a (worse) upper bound for IS(3).

The question about the maximal finite number of solutions of one-variable equations had been considered before [8,7,19], but the above connnection made it even more important. It had been proved that the number of solutions is at most logarithmic with respect to the number of occurrences of the variable in the equation [19]. From this it now followed that the size of an independent system of constant-free three-variable equations is at most logarithmic with respect to the length of the shortest equation in the system, thus improving the previous linear bound.

It had been conjectured that a one-variable equation with only finitely many solutions has at most two solutions. In view of [21], this conjecture would have implied that $\mathrm{IS}(3) \leq 18$. In 2016, however, the attempts to prove the conjecture led to the discovery of the counterexample equation (first published in [22])

$$(xaxbxaabbabaxbabaabbab, abaabbabaxbabaabbxaxbx)$$

with exactly three solutions $h(x) = \varepsilon$, $h(x) = ab$ and $h(x) = abaabbab$, and to a weaker version of the conjecture: A one-variable equation with only finitely many solutions has at most three solutions. This weaker conjecture would have implied that $\mathrm{IS}(3) \leq 23$. In the journal version [22] of the conference article [21],

the conditional result was improved so that even the weaker conjecture would imply $\mathrm{IS}(3) \leq 18$.

In 2017, Saarela [27] proved that there does not exist words $u_0, \ldots, u_m$ such that $u_0^k = u_1^k \ldots u_m^k$ for all $k \leq 3$ but not for any $k \geq 4$ (a stronger result was proved later in the journal version [28]). This meant that one particular approach for improving the lower bound $\mathrm{IS}(n) = \Omega(n^4)$ was impossible, as was mentioned in Section 4. However, the more significant consequence of [27] was that Nowotka and Saarela found a way to apply a method similar to the one used in [27] to study one-variable equations. In 2018, the weaker conjecture about one-variable equations was finally proved, thus proving that $\mathrm{IS}(3) \leq 18$ [23].

Theorem 2. *A one-variable word equation has either infinitely many solutions or at most three.*

Theorem 3. $3 \leq \mathrm{IS}(3) \leq 18$.

The rough idea behind the proof of Theorem 2 is that given an equation (u, v) on the variable x, we can assume that the equation is in a certain kind of normal form, and then we find a morphism $\sigma : \Sigma^* \to \mathbb{Z}$ that satisfies certain properties, for example $\sigma(h(x)) = 0$ for all $h \in \mathrm{Sol}((u, v))$. Then we study the images of prefixes of $h(u)$ and $h(v)$ under σ, how they match and how they change when the solution h changes. This eventually leads to a proof of Theorem 2.

Theorem 3 follows from Theorem 2 and the result in [22]. The idea behind the proof of the result in [22] is that if $E_1, \ldots, E_k$ is an independent system of constant-free three-variable equations and $h_1, \ldots, h_k$ are morphisms such that $h_i \in \mathrm{Sol}(E_j)$ for all $i \neq j$ but not for $i = j$, then the morphisms h_i can be classified into a small number of families by the results in [2], and, given Theorem 2, not too many of them can be in the same family. Improving the result in [22] seems like a potential path towards a smaller upper bound than 18.

6 Variations

If we consider equations in a free semigroup instead of a free monoid, that is, we require that a solution h cannot map a variable to the empty word, then how large can an independent system of constant-free n-variable equations be? The largest known examples have size $\Theta(n^3)$ [17]. Independent systems can be studied also in other semigroups, see the survey of Harju, Karhumäki and Plandowski [11].

A finite sequence of nontrivial equations $E_1, \ldots, E_n$ is a *decreasing chain* if

$$\mathrm{Sol}(E_1) \supsetneq \mathrm{Sol}(E_1, E_2) \supsetneq \cdots \supsetneq \mathrm{Sol}(E_1, \ldots, E_n).$$

and an *increasing chain* if

$$\mathrm{Sol}(E_1, \ldots, E_n) \subsetneq \mathrm{Sol}(E_2, \ldots, E_n) \subsetneq \cdots \subsetneq \mathrm{Sol}(E_n).$$

Similar definition can be given for infinite chains. Chains have been studied by Honkala [15] and Czeizler [4], for example. If the number of variables is fixed,

then how long can chains be? Clearly, $E_1, \ldots, E_n$ is a decreasing chain if and only if $E_n, \ldots, E_1$ is an increasing chain, so if there exists a finite maximal length, then it is the same for both types of chains. However, if there are arbitrarily long chains, then the situation is more complicated: Ehrenfeucht's conjecture is equivalent to the fact that decreasing chains cannot be infinite, but no such result is known for increasing chains. In the case of constant-free three-variable equations, there are chains of length 7 [18], and the results in [22] and [23] imply an upper bound of 24.

Finally, independent systems can be considered also in the case of equations with constants. Independent systems can be larger in this case, but Ehrenfeucht's conjecture still holds. Moreover, an independent n-variable system of equations with constants cannot be larger than an independent $(n+2)$-variable system of constant-free equations. This is because we can assume that the alphabet of constants is binary, and then replace the constant letters with new variables, and this preserves independence. Thus the question of the maximal size of independent systems does not become fundamentally different if we allow constants.

References

1. Albert, M.H., Lawrence, J.: A proof of Ehrenfeucht's conjecture. Theoret. Comput. Sci. **41**(1), 121–123 (1985). https://doi.org/10.1016/0304-3975(85)90066-0
2. Budkina, L.G., Markov, A.A.: F-semigroups with three generators. Mat. Zametki **14**, 267–277 (1973)
3. Culik, II, K., Karhumäki, J.: Systems of equations over a free monoid and Ehrenfeucht's conjecture. Discrete Math. **43**(2–3), 139–153 (1983). https://doi.org/10.1016/0012-365X(83)90152-8
4. Czeizler, E.: Multiple constraints on three and four words. Theoret. Comput. Sci. **391**(1–2), 14–19 (2008). https://doi.org/10.1016/j.tcs.2007.10.026
5. Czeizler, E., Karhumäki, J.: On non-periodic solutions of independent systems of word equations over three unknowns. Internat. J. Found. Comput. Sci. **18**(4), 873–897 (2007). https://doi.org/10.1142/S0129054107005030
6. Czeizler, E., Plandowski, W.: On systems of word equations over three unknowns with at most six occurrences of one of the unknowns. Theoret. Comput. Sci. **410**(30–32), 2889–2909 (2009). https://doi.org/10.1016/j.tcs.2009.01.023
7. Dąbrowski, R., Plandowski, W.: On word equations in one variable. Algorithmica **60**(4), 819–828 (2011). https://doi.org/10.1007/s00453-009-9375-3
8. Eyono Obono, S., Goralčík, P., Maksimenko, M.: Efficient solving of the word equations in one variable. In: Proceedings of the 19th MFCS. LNCS, vol. 841, pp. 336–341. Springer (1994). https://doi.org/10.1007/3-540-58338-6_80
9. Guba, V.S.: Equivalence of infinite systems of equations in free groups and semigroups to finite subsystems. Mat. Zametki **40**(3), 321–324 (1986). https://doi.org/10.1007/BF01142470
10. Hakala, I., Kortelainen, J.: On the system of word equations $x_1^i x_2^i \cdots x_m^i = y_1^i y_2^i \cdots y_n^i$ $(i = 1, 2, \cdots)$ in a free monoid. Acta Inform. **34**(3), 217–230 (1997). https://doi.org/10.1007/s002360050081
11. Harju, T., Karhumäki, J., Plandowski, W.: Independent systems of equations. In: Lothaire, M. (ed.) Algebraic Combinatorics on Words, pp. 443–472. Cambridge University Press (2002)

12. Harju, T., Nowotka, D.: On the independence of equations in three variables. Theoret. Comput. Sci. **307**(1), 139–172 (2003). https://doi.org/10.1016/S0304-3975(03)00098-7
13. Holub, Š.: Local and global cyclicity in free semigroups. Theoret. Comput. Sci. **262**(1–2), 25–36 (2001). https://doi.org/10.1016/S0304-3975(00)00156-0
14. Holub, Š., Žemlička, J.: Algebraic properties of word equations. J. Algebra **434**, 283–301 (2015). https://doi.org/10.1016/j.jalgebra.2015.03.021
15. Honkala, J.: On chains of word equations and test sets. Bull. Eur. Assoc. Theor. Comput. Sci. EATCS **68**, 157–160 (1999)
16. Jeż, A.: Recompression: a simple and powerful technique for word equations. J. ACM **63**(1), Art. 4, 51 (2016). https://doi.org/10.1145/2743014
17. Karhumäki, J., Plandowski, W.: On the defect effect of many identities in free semigroups. In: Paun, G. (ed.) Mathematical aspects of natural and formal languages, pp. 225–232. World Scientific (1994). https://doi.org/10.1142/9789814447133_0012
18. Karhumäki, J., Saarela, A.: On maximal chains of systems of word equations. Proc. Steklov Inst. Math. **274**, 116–123 (2011). https://doi.org/10.1134/S0081543811060083
19. Laine, M., Plandowski, W.: Word equations with one unknown. Internat. J. Found. Comput. Sci. **22**(2), 345–375 (2011). https://doi.org/10.1142/S0129054111008088
20. Makanin, G.S.: The problem of the solvability of equations in a free semigroup. Mat. Sb. (N.S.) **103**(2), 147–236 (1977), english translation in Math. USSR Sb. 32:129–198, 1977
21. Nowotka, D., Saarela, A.: A connection between one-unknown word equations and constant-free three-unknown word equations. In: Proceedings of the 20th DLT. LNCS, vol. 9840, pp. 332–343. Springer (2016). https://doi.org/10.1007/978-3-662-53132-7_27
22. Nowotka, D., Saarela, A.: One-variable word equations and three-variable constant-free word equations. Internat. J. Found. Comput. Sci. **29**(5), 935–950 (2018). https://doi.org/10.1142/S0129054118420121
23. Nowotka, D., Saarela, A.: An optimal bound on the solution sets of one-variable word equations and its consequences. In: Proceedings of the 45th ICALP. LIPIcs, vol. 107, pp. 136:1–136:13. Schloss Dagstuhl–Leibniz-Zentrum fuer Informatik (2018). https://doi.org/10.4230/LIPIcs.ICALP.2018.136
24. Plandowski, W.: Test sets for large families of languages. In: Proceedings of the 7th DLT. LNCS, vol. 2710, pp. 75–94. Springer (2003). https://doi.org/10.1007/3-540-45007-6_6
25. Plandowski, W.: Satisfiability of word equations with constants is in PSPACE. J. ACM **51**(3), 483–496 (2004)
26. Saarela, A.: Systems of word equations, polynomials and linear algebra: A new approach. European J. Combin. **47**, 1–14 (2015). https://doi.org/10.1016/j.ejc.2015.01.005
27. Saarela, A.: Word equations where a power equals a product of powers. In: Proceedings of the 34th STACS. LIPIcs, vol. 66, pp. 55:1–55:9. Schloss Dagstuhl–Leibniz-Zentrum fuer Informatik (2017). https://doi.org/10.4230/LIPIcs.STACS.2017.55
28. Saarela, A.: Word equations with kth powers of variables. J. Combin. Theory Ser. A **165**, 15–31 (2019). https://doi.org/10.1016/j.jcta.2019.01.004
29. Spehner, J.C.: Quelques problémes d'extension, de conjugaison et de présentation des sous-monoïdes d'un monoïde libre. Ph.D. thesis, Univ. Paris (1976)
30. Spehner, J.C.: Les systemes entiers d'équations sur un alphabet de 3 variables. In: Semigroups. pp. 342–357 (1986)