What Do We Know About Language Equations?

Michal Kunc*

Department of Mathematics and Statistics, Masaryk University,
Janáčkovo nám. 2a, 602 00 Brno, Czech Republic
kunc@math.muni.cz, http://www.math.muni.cz/~kunc/

In the talk we give an overview of recent developments in the area of language equations, with an emphasis on methods for dealing with non-classical types of equations whose theory has not been successfully developed already in the previous decades, and on results forming the current borderline of our knowledge. This abstract is in particular meant to provide the interested listener with references to the material discussed in the talk.

Motivations for studying equations over languages come from several sources (e.g. formal grammars, automata constructions, word equations, set constraints, games or natural computing) and most of the results on these equations are related to one of these topics.

Language equations were first applied in [14] to elegantly define semantics for context-free grammars by means of explicit systems of equations with the operations of union and concatenation. Some interesting examples of using these systems can be found in [52]. By allowing in these systems also intersection, one obtains the notion of conjunctive languages [36,37], which are more general than context-free ones even over a unary alphabet [17]. The special case of linear conjunctive languages was studied in [39].

The theory of explicit systems of language equations with concatenation and all Boolean operations was developed in [46], and even one-variable systems were proved computationally universal [43]. The appropriate restriction of these systems to define Boolean grammars was described in [38]. Several basic open problems about conjunctive and Boolean languages are proposed in [45]. The classes of languages obtained by allowing in explicit systems additionally to concatenation all possible clones of Boolean operations were also determined [47,44]. Explicit systems were further shown to naturally define arithmetical hierarchy [40]. Solutions of explicit systems with some language operations other than concatenation were also described, e.g. equations employing homomorphisms are related to ET0L languages [53].

Implicit language equations where concatenation is the only operation naturally appear as a generalization of equations over words to sets of words. Existence of solutions of word equations with constants was proved decidable by Makanin [34]. Currently best algorithms for solving word equations can be found in [49]. It is also well known that solvability of word equations is decidable even for infinite rational systems of equations [10,2,15].

* Supported by the Grant no. 201/06/0936 of the Grant Agency of the Czech Republic.

T. Harju, J. Karhumäki, and A. Lepistö (Eds.): DLT 2007, LNCS 4588, pp. 23–27, 2007.

For equations over languages, the situation is completely different. In [48] existence of arbitrary solutions was proved undecidable for equations with finite constants employing union and concatenation. When regular constant languages are allowed, the problem is undecidable already for one-variable systems using only the operation of concatenation [29]. But there is no such result about equations with only finite constants, and we also have virtually no knowledge about solvability of finite systems over finite or regular languages. On the other hand, it is known that already for a very simple rational system of equations with only concatenation we cannot algorithmically decide whether given finite languages form its solution [32,20,30].

Most of the results about implicit language equations and inequalities concern inequalities of particular forms, often related to important classes of formal grammars or basic automata constructions like those of minimal and universal automata (see [54]). Results of this kind were surveyed in [26]. General treatment of systems of implicit equations was initiated by Okhotin [41], who considered also strict inequalities [42].

General systems of equations and inequalities with constant right-hand sides were studied by Conway [9], and the exact complexity of determining their solvability was established in [6]. The study of such equations was also extended to the simplest equations with more general operations than concatenation based on shuffle and deletion along trajectories [22,23,12].

Some generalizations of standard systems of right-linear equations were considered by Leiss [31]. For general systems of right-linear inequalities, basic problems can be solved using Rabin's results on MSO logic on infinite trees [50]; the complexity of these problems has been determined in [1,8,4,3,5]. Regularity of largest solutions in the case of inequalities with non-regular left-hand sides was established in [28].

The method of proving regularity by means of well quasi-orders was developed by Ehrenfeucht et al. [13]; a number of results on regularity of languages based on well quasi-orders can be found in [11]. Well quasi-orders were used to show regularity of largest solutions of systems of inequalities with certain restrictions on constant languages [25].

The borderline between equations with algorithmically constructible regular largest solutions and those having universal expressive power appears to be formed by semi-commutation inequalities $XK \subseteq LX$. For any regular language L, the largest solution of such an inequality is always regular [25], but the only known proof of this fact is non-constructive, based on Kruskal's tree theorem [24], and we know how to algorithmically find the largest solution only in a very special case [33]. However, systems of two semi-commutation inequalities possess universal expressive power [27]. A prominent role among these systems is played by commutation equations, which were first considered by Conway [9] and later studied in many papers (see [21] for a survey). Basic results were achieved and conjectures formulated in [51]; regularity of the largest solutions was proved for three-element languages [19] (a more general result based on lexicographic ordering can be found in [35]) and regular codes [18]. The expressive universality of commutation equations over finite languages was established in [29] (a more intuitive incremental construction for this result is described in [16]). Some partial results were proved also for equations expressing conjugacy of languages [7]; an undecidability result for these equations can be found in [29].

References

1. Aiken, A., Kozen, D., Vardi, M., Wimmers, E.: The complexity of set constraints. In: Meinke, K., Börger, E., Gurevich, Y. (eds.) CSL 1993. LNCS, vol. 832, pp. 1–17. Springer, Heidelberg (1994)
2. Albert, M.H., Lawrence, J.: A proof of Ehrenfeucht's conjecture. Theoret. Comput. Sci. 41(1), 121–123 (1985)
3. Baader, F., Küsters, R.: Unification in a description logic with transitive closure of roles. In: Nieuwenhuis, R., Voronkov, A. (eds.) LPAR 2001. LNCS (LNAI), vol. 2250, pp. 217–232. Springer, Heidelberg (2001)
4. Baader, F., Narendran, P.: Unification of concept terms in description logics. J. Symbolic Comput. 31(3), 277–305 (2001)
5. Baader, F., Okhotin, A.: Complexity of language equations with one-sided concatenation and all Boolean operations. In: Proc. UNIF'06, pp. 59–73 (2006)
6. Bala, S.: Complexity of regular language matching and other decidable cases of the satisfiability problem for constraints between regular open terms. Theory Comput. Syst. 39(1), 137–163 (2006)
7. Cassaigne, J., Karhumäki, J., Maňuch, J.: On conjugacy of languages. Theor. Inform. Appl. 35(6), 535–550 (2001)
8. Charatonik, W., Podelski, A.: Co-definite set constraints. In: Nipkow, T. (ed.) Rewriting Techniques and Applications. LNCS, vol. 1379, pp. 211–225. Springer, Heidelberg (1998)
9. Conway, J.H.: Regular Algebra and Finite Machines. Chapman and Hall, London (1971)
10. Culik II, K., Karhumäki, J.: Systems of equations over a free monoid and Ehrenfeucht's conjecture. Discrete Math. 43(2–3), 139–153 (1983)
11. de Luca, A., Varricchio, S.: Finiteness and Regularity in Semigroups and Formal Languages. Springer, Heidelberg (1999)
12. Domaratzki, M., Salomaa, K.: Decidability of trajectory-based equations. Theoret. Comput. Sci. 345(2–3), 304–330 (2005)
13. Ehrenfeucht, A., Haussler, D., Rozenberg, G.: On regularity of context-free languages. Theoret. Comput. Sci. 27(3), 311–332 (1983)
14. Ginsburg, S., Rice, H.G.: Two families of languages related to ALGOL. J. ACM 9(3), 350–371 (1962)
15. Guba, V.S.: Equivalence of infinite systems of equations in free groups and semigroups to finite subsystems. Mat. Zametki 40(3), 321–324 (1986)
16. Jeandel, E., Ollinger, N.: Playing with Conway's problem. Technical report ccsd-00013788, Laboratoire d'Informatique Fondamentale de Marseille (2005). Available at `http://hal.archives-ouvertes.fr/hal-00013788`
17. Jeż, A.: Conjunctive grammars can generate non-regular unary languages. In this volume
18. Karhumäki, J., Latteux, M., Petre, I.: Commutation with codes. Theoret. Comput. Sci. 340(2), 322–333 (2005)

19. Karhumäki, J., Latteux, M., Petre, I.: Commutation with ternary sets of words. Theory Comput. Syst. 38(2), 161–169 (2005)
20. Karhumäki, J., Lisovik, L.P.: The equivalence problem of finite substitutions on ab^*c, with applications. Internat. J. Found. Comput. Sci. 14(4), 699–710 (2003)
21. Karhumäki, J., Petre, I.: Two problems on commutation of languages. In: Current Trends in Theoretical Computer Science, The Challenge of the New Century, vol. 2, pp. 477–494. World Scientific, Singapore (2004)
22. Kari, L.: On language equations with invertible operations. Theoret. Comput. Sci. 132(1–2), 129–150 (1994)
23. Kari, L., Sosík, P.: On language equations with deletion. Bull. EATCS 83, 173–180 (2004)
24. Kruskal, J.B.: Well-quasi-ordering, the tree theorem, and Vazsonyi's conjecture. Trans. Amer. Math. Soc. 95(2), 210–225 (1960)
25. Kunc, M.: Regular solutions of language inequalities and well quasi-orders. Theoret. Comput. Sci. 348(2–3), 277–293 (2005)
26. Kunc, M.: Simple language equations. Bull. EATCS 85, 81–102 (2005)
27. Kunc, M.: On language inequalities $XK \subseteq LX$. In: De Felice, C., Restivo, A. (eds.) DLT 2005. LNCS, vol. 3572, pp. 327–337. Springer, Heidelberg (2005)
28. Kunc, M.: Largest solutions of left-linear language inequalities. In: Proc. AFL 2005, University of Szeged, pp. 178–186 (2005), Also available at `http://www.math.muni.cz/~kunc/math/left_linear.ps`.
29. Kunc, M.: The power of commuting with finite sets of words. Theory Comput. Syst. 40(4), 521–551 (2007)
30. Kunc, M.: The simplest language where equivalence of finite substitutions is undecidable. Preprint available at `http://www.math.muni.cz/~kunc/math/finite_substitutions.pdf`.
31. Leiss, E.L.: Language Equations. Springer, Heidelberg (1999)
32. Lisovik, L.P.: The equivalence problem for finite substitutions on regular languages. Dokl. Akad. Nauk 357(3), 299–301 (1997)
33. Ly, O.: A constructive solution of the language inequation $XA \subseteq BX$. Preprint available at `http://www.labri.fr/perso/ly/publications/LanguageEquation.pdf`.
34. Makanin, G.S.: The problem of solvability of equations in a free semigroup. Mat. Sb. 103(2), 147–236 (1977)
35. Massazza, P., Salmela, P.: On the simplest centralizer of a language. Theor. Inform. Appl. 40(2), 295–301 (2006)
36. Okhotin, A.: Conjunctive grammars. J. Autom. Lang. Comb. 6(4), 519–535 (2001)
37. Okhotin, A.: Conjunctive grammars and systems of language equations. Program. Comput. Software 28(5), 243–249 (2002)
38. Okhotin, A.: Boolean grammars. Inform. and Comput. 194(1), 19–48 (2004)
39. Okhotin, A.: On the equivalence of linear conjunctive grammars and trellis automata. Theor. Inform. Appl. 38(1), 69–88 (2004)
40. Okhotin, A.: A characterization of the arithmetical hierarchy by language equations. Internat. J. Found. Comput. Sci. 16(5), 985–998 (2005)
41. Okhotin, A.: Unresolved systems of language equations: Expressive power and decision problems. Theoret. Comput. Sci. 349(3), 283–308 (2005)
42. Okhotin, A.: Strict language inequalities and their decision problems. In: Jedrzejowicz, J., Szepietowski, A. (eds.) MFCS 2005. LNCS, vol. 3618, pp. 708–719. Springer, Heidelberg (2005)
43. Okhotin, A.: Computational universality in one-variable language equations. Fund. Inform. 74(4), 563–578 (2006)

44. Okhotin, A.: Language equations with symmetric difference. In: Grigoriev, D., Harrison, J., Hirsch, E.A. (eds.) CSR 2006. LNCS, vol. 3967, pp. 292–303. Springer, Heidelberg (2006)
45. Okhotin, A.: Nine open problems for conjunctive and Boolean grammars. Bull. EATCS 91 (to appear, 2007)
46. Okhotin, A.: Decision problems for language equations. Submitted for publication. Preliminary version. In: Baeten, J.C.M., Lenstra, J.K., Parrow, J., Woeginger, G.J. (eds.) ICALP 2003. LNCS, vol. 2719, pp. 239–251. Springer, Heidelberg (2003)
47. Okhotin, A., Yakimova, O.: Language equations with complementation: Decision problems. Theoret. Comput. Sci. (to appear)
48. Parikh, R., Chandra, A., Halpern, J., Meyer, A.: Equations between regular terms and an application to process logic. SIAM J. Comput. 14(4), 935–942 (1985)
49. Plandowski, W.: An efficient algorithm for solving word equations. In: Proc. STOC'06, pp. 467–476. ACM, New York (2006)
50. Rabin, M.O.: Decidability of second-order theories and automata on infinite trees. Trans. Amer. Math. Soc. 141, 1–35 (1969)
51. Ratoandromanana, B.: Codes et motifs. RAIRO Inform. Théor. Appl. 23(4), 425–444 (1989)
52. Rozenberg, G., Salomaa, A. (eds.): Handbook of Formal Languages. Springer, Heidelberg (1997)
53. Ruohonen, K.: A note on language equations involving morphisms. Inform. Process. Lett. 7(5), 209–212 (1978)
54. Sakarovitch, J.: Elements of Automata Theory. Cambridge University Press, Cambridge (2007)