Sequence Disunification and one Application to Collaborative Schema Construction

Jorge Coelho[1], Mário Florido[2], and Temur Kutsia[3]

[1] Instituto Superior de Engenharia do Porto & LIACC Porto, Portugal jcoelho@ncc.up.pt
[2] University of Porto, DCC-FC & LIACC Porto, Portugal amf@ncc.up.pt
[3] RISC, Johannes Kepler University Linz, Austria tkutsia@risc.uni-linz.ac.at

Introduction

In this paper we use an algebra of terms with function symbols of flexible arity (suitable to represent XML) and variables that can be substituted by sequences of terms, called sequence variables. Unification in theories with sequence variables and flexible arity function symbols (sequence unification or SU in short) is a recent topic of research with applications in several areas in computer science. In [7, 8] decidability of sequence unification was shown, and a minimal complete solving procedure was introduced. Recently, unification of terms with flexible arity functions symbols has been successfully applied to XML-processing and website content verification [9, 3, 2, 4]. Sequence unification is infinitary, that means that for some problems the minimal complete set of solutions may be infinite, which implies that any complete procedure is in general nonterminating. Several finitary fragments of sequence unification have been identified in [8].

We define a terminating procedure that covers larger fragments than those from [8]. Moreover, we incorporate negative information in the specification, for instance, specifying that two sequences do not share a common element. It leads to sequence disunification that is a new development. It extends sequence unification with disjointness equations, that for terms without sequence variables coincides with standard disunification [5, 6, 1].

Thus, the main contributions of this work are: 1. A minimal complete procedure for sequence disunification (dealing with disjointness equations). 2. A minimal complete algorithm for the case where each sequence variable occurs maximum twice in the disunification problem. 3. Applying these techniques to Collaborative Schema Construction in XML where disunification of schemas is used to create more general ones and a parser for the general schema is generated.

References

1. W. L. Buntine and H.-J. Bürckert. On solving equations and disequations. *J. ACM*, 41(4):591–629, 1994.
2. J. Coelho and M. Florido. Unification with flexible arity symbols: a typed approach. In *Proc. 20th Int. Workshop on Unification (UNIF'06)*, Seattle, USA, 2006.
3. J. Coelho and M. Florido. VeriFLog: Constraint Logic Programming Applied to Verification of Website Content. In *Int. Workshop XML Research and Applications (XRA'06)*, volume 3842 of *LNCS*. Springer-Verlag, 2006.
4. J. Coelho and M. Florido. XCentric: A logic-programming language for XML processing. In *PLAN-X*, 2007.
5. H. Comon. Disunification: A survey. In *Computational Logic - Essays in Honor of Alan Robinson*, pages 322–359, 1991.
6. H. Comon and P. Lescanne. Equational problems and disunification. *J. Symb. Comput.*, 7(3/4):371–425, 1989.
7. T. Kutsia. Unification with sequence variables and flexible arity symbols and its extension with pattern-terms. In *Proc. AISC'2002 - Calculemus'2002*, Volume 2385 of LNAI, 2002.
8. T. Kutsia. Solving equations with sequence variables and sequence functions. *J. Symb. Comp.*, 42:352–388, 2007.
9. T. Kutsia and M. Marin. Can context sequence matching be used for querying XML? In *Proc. UNIF'05*, pages 77–92, 2005.