

Towards Craig Interpolation for String Constraints

Parosh Aziz Abdulla¹, Mohamed Faouzi Atig¹, Yu-Fang Chen², Lukáš Holík³,
Ahmed Rezzine⁴, Philipp Rümmer¹, and Jari Stenman¹

¹ Department of Information Technology, Uppsala University, Sweden

² Institute of Information Science, Academia Sinica, Taiwan

³ Faculty of Information Technology, Brno University of Technology, Czech Republic

⁴ Department of Computer and Information Science, Linköping University, Sweden

The problem of reasoning automatically about *string formulae*, containing constraints such as string/word equations, regular language membership, or length restrictions, has increasingly received attention over the past years. For program analysis, for instance in the database and web area, the ability to precisely represent and reason about string data types is often essential. Much of the existing work has concentrated on checking satisfiability of string formulae, by means of translation to bit-vector formulae or analysis of finite automata representations, among others.

For verifying correctness of programs operating on strings, in addition functionality to extract Craig interpolants for string constraints is beneficial. We report about ongoing research on Craig interpolation for string formulae, mainly done in the context of work presented in our CAV'14 paper. Our workshop presentation covers the following topics:

- we outline a new algorithm for solving string constraints in an SMT setting. The algorithm is sound and complete for formulae in a rich language providing equations, regular expression constraints, and length constraints in Presburger arithmetic.
- applying this decision procedure, we propose a constraint-based method for computing bounded string interpolants in form of regular expression constraints. The procedure is incomplete, in the sense that interpolants cannot be computed for all unsatisfiable conjunctions of string formulae, but covers a range of practically relevant cases. Boundedness is essential for preventing divergence of model checking algorithms that use the computed Craig interpolants as building blocks for invariants.
- we give an experimental evaluation using a set of (small) string-processing programs that can be verified automatically using our techniques.