

All Strings Attached: String and Sequence Constraints in Z3

Nikolaj Bjorner

Microsoft Research
nbjorner@microsoft.com

Abstract. In this talk I describe recent efforts for theory support for strings, sequences and regular expressions in Z3. One of the main conveniences of Satisfiability Modulo Theories (SMT) solvers is their support for theories that are commonly used in program verification and analysis. The theories of arithmetic and bit-vectors have shown to be ubiquitous in almost all applications of SMT, while other theories, such as algebraic data-types, seem to be essential to cover only more sophisticated applications. The theory of strings, sequences and regular expressions have been an occasional but persistent ask from users of SMT tools. Efforts on supporting strings and sequences, have however only been catching on relatively recently, first with tools that solve string constraints under assumptions of bounded lengths (Kaluza, Hampi and the string solver in Pex), followed by a proposal for sequences in the exchange format for SMT solvers and implementations for strings in CVC4, Princess, S3, and Z3Str. A separate line of work has considered Symbolic Automata, where transitions are labeled by formulas over a background theory, and automata operations are performed directly by solving satisfiability constraints. Our approach in Z3 combines some of the approaches taken in string solvers and integrate methods from symbolic automata. I will describe some of our experiences so far, propose new opportunities with using SMT solvers with sequences, and summarise some of the technical challenges ahead.