Informatik 11
Embedded Software

RWTH AACHEN UNIVERSITY

Master’s Thesis

Analysis of Structured Text by translation to LLVM Bytecode

Problem Statement

Programmable Logic Controllers (PLCs) are industrial computers used in a variety of factory automation tasks. As factory automation is an especially safety critical environment, it is imperative that the software running on PLCs is bug-free and conforms to its specification. While there is a wide ecosystem of analysis tools for commonly used languages such as Java, C/C++ or Python, PLCs distinguish themselves by using specialised programming languages standardised by IEC 61131-3. An immediate consequence of this is that in the past, known algorithms for program analysis had to be reimplemented for the IEC languages separately, which is unnecessary busywork.

LLVM COMPILER INFRASTRUCTURE

Task

In this thesis, leveraging existing analysis tools for PLC software should be enabled by translating the PLC programming language Structured Text into LLVM bytecode, and then using the LLVM-based verification tools CRAB and SeaHorn to perform verification tasks. A parser for Structured Text already exists as part of the PLC code analysis tool ARCADE developed by Informatik 11, which transforms the string representation of the code into an abstract syntax tree (AST). This AST representation should be used as input for the translation algorithm.

Vorkenntnisse

ARCADE is implemented in C++, therefore knowledge of C++ is required in order to be able to implement the translation. Knowledge of LLVM, Compilers and/or Structured Text and PLCs in general is helpful, but should be acquirable during the thesis.

Ansprechpartner

Marcus Völker, M.Sc. RWTH
voelker@embedded.rwth-aachen.de

09.10.2014