Report from the WCET Tool Challenge 2006
Ideas for the WCET Tool Challenge 2008

Jan Gustafsson
Department of Computer Science and Electronics, Mälardalen University

jan.gustafsson@mdh.se

Abstract

The purpose of the WCET Tool Challenge is to be able to study, compare and discuss the properties of different WCET tools and approaches, to define common metrics, and to enhance the existing benchmarks. The WCET Tool Challenge has been designed to find a good balance between openness for a wide range of analysis approaches, and specific participation guidelines to provide a level playing field. This should make results transparent and facilitate friendly competition among the participants.

This short report presents conclusions from from the WCET Tool Challenge 2006 as well as some ideas for the WCET Tool Challenge 2006.

1 Goals

The goals of the WCET Tool Challenge are the following:

- To exhibit the wide range of timing analysis tools available today
  - using static program analysis, or
  - combining analysis and measurements,
  - for various target processors,
  - in various application domains,
  - supporting various programming languages and design tools,
  - academic, commercial; free or at a charge.
- To illuminate the features, abilities and intended uses of each tool
  - in finding the feasible execution paths in the SW,
  - in modelling complex processor and system HW,
  - in deriving useful WCET bounds or estimates,
  - in usability, scalability and adaptability,
  - in the range of supported targets (processors, compilers, ..)
- To collect and maintain a growing set of community standard benchmark programs and related test suites that
  - contain typical (both easy and hard) programming constructs,
  - can be analyzed by several tools with comparable results,
  - test enough of the actual behaviour of each benchmark to satisfy measurement-based tools and to validate results from static-analysis tools, and ideally, have known exact answers (paths and WCETs).

  For more details, consult the Challenge web page `http://www.idt.mdh.se/personal/jgn/challenge/`.

2 WCET Tool Challenge 2006

The first WCET Tool Challenge was performed during the autumn of 2006. It concentrated on three aspects of WCET analysis:

1. flow analysis,
2. required user interaction,
3. performance.

Two companies (developing the commercial tools aiT and Bound-T) and three research groups (developing the research tools SWEET, Chronos and MTime) participated. The actual work with the tools was made by an external user and the development teams. The evaluation was targeted on a set of benchmark programs.

The report was presented at ISoLA 2006 as two separate papers. The first [1] was the main report; the other [3] presented the experiences from the external user. There is also a technical report available [2], which contains more details.

ECRTS 2007
7th Intl. Workshop on Worst-Case Execution Time (WCET) Analysis
http://drops.dagstuhl.de/opus/volltexte/2007/1198

2.1 Results from WCET Tool Challenge 2006

This first Challenge was performed successfully in spite of some initial unclearness, debate and delay. We noted that there is no common format of the results from the different tests, and that different developers have made different tests sometimes. We draw the following conclusions:

- The tests have been a real challenge to the participating WCET tools. We have a success range from 0% to 100% in terms of how many of the benchmark programs that were analyzable by a certain tool.
- The tests have clearly pointed out problems existing in the tools as well as in the benchmarks and the used compilers.
- Most of the tools find more than half of the loop bounds automatically. Only one tool finds infeasible paths automatically.
- Several bugs in both the tools and the benchmarks have been corrected during the Challenge.
- Actual WCET estimates cannot be compared this time since the developers support different processors and compilers.
- The quality of WCET estimates is hard to judge for all tools but aiT, since aiT was the only tool to provide measurements for some of the benchmarks. Chronos provided simulated values that indicate the possible size of overestimation.

3 WCET Tool Challenge 2008

Since the work to perform the Challenge is extensive, both for the work group and the developers, we decided to make the event bi-annual. Another reason was to have some tool development time between tests, so real progress could be observed from one Challenge to the next. This means that the next Challenge is to be performed 2008.

A number of actions have to be made to enhance the next Challenge compared to the first. Some of them are:

- Update benchmarks and the setup of the Challenge according to the feedback from the WCET Tool Challenge 2006.
- Extend the active working group to 2 or 3 persons.
- Create a reference group which will act as a support to the working group.
- Try to extend the number of participants. Tools like OTAWA and Heptane, and measurement-based tools (e.g., Rapita, SYMTA/P) should be invited again. There are also some new tools entering the WCET analysis scene (like TimeBounder from Korea).
- The procedure of identifying test data (inputs) necessary for measurement-based based tools should be defined.
- The inputs for the worst-case behavior of the benchmark programs should be defined.
- Be more detailed concerning tests and format of result reports.
- Exclude benchmark programs that are redundant in terms of what is tested.
- Include industrial code, e.g. aero space and automotive code (e.g., Daimler-Chrysler code).
- If possible, include more automatically generated code.
- Try to focus on one common processor (ARM7?) to be able to compare WCET estimates.
- Define hardware setup in detail so comparative low-level analysis (e.g., cache and branch prediction analysis) can be performed.

References

[1] J. Gustafsson. The worst case execution time tool challenge 2006. In *Proc. 2nd International Symposium on Leveraging Applications of Formal Methods (ISOLA'06)*, Nov. 2006.

[2] J. Gustafsson. WCET challenge 2006 technical report. MRTC report 1209, 2007. Technical Report MRTC report 1209, 2007, Mälardalen University Real-Time Research Centre, Mälardalen University University, Västerås, Sweden, 2007.

[3] L. Tan. The worst case execution time tool challenge 2006: The external test. In *Proc. 2nd International Symposium on Leveraging Applications of Formal Methods (ISOLA'06)*, Nov. 2006.