Program Analysis and Test Data Generation Through Constraint Solving

JIAN ZHANG
CHINESE ACADEMY OF SCIENCES

Introduction

Main Research Topics

Automated Reasoning and Constraint Solving
satisfiability in the first-order logic, SAT, SMT, ..., model/solution counting

Black-box testing –
combinatorial testing
EFSM-based testing
stress testing

Given a C program, find

- a small set of test cases to meet some criterion like
  - statement coverage
  - branch coverage
  - basis path
- bugs in the program
- general bugs (e.g., memory leak and infinite looping)
- application-specific bugs (violation of user-specified assertions)
- hot paths in the program

A popular approach –
Symbolic Execution + Constraint Solving

[Zhang VSTTE 2005 (LNCS 4171)]

The approach can be used for

- Verification / bug finding
- Unit testing: model-based testing
- Combination with classical static analysis

Combinatorial Testing

Black-box testing technique, used in AT&T, Motorola, IBM, ...
The system-under-test (SUT) has a set of parameters/components, each of which can take some values.

Example:

- Browser: IE, Netscape, Firefox, ...
- Operating system: Linux, Windows NT, ...
- Manufacturer: HP, Dell, Lenovo, ...

Finding Smallest Combinatorial Test Suite

- Backtracking search + heuristics
- Tool: EXACT for finding Covering Arrays
  Charles Colbourn: “The CA(24,4,12,2) yields a *lot* of improvements!”
  Jun Yan and Jian Zhang, J. Systems and Software 2008
- Tool: BOAS for finding Orthogonal Arrays
  Feifei Ma and Jian Zhang, PRICAI 2008.

Static Analysis and White-box Test Generation

- An approach to path feasibility analysis: PAT / ePAT
  [Zhang-Wang 2001]
- A tool for generating small test suites for unit testing
  [Xu-Zhang 2006]
- A method for finding executable/feasible basis paths
  [Yan-Zhang 2008]
- A sufficient condition for the detection of infinite looping
  [Zhang 2001]
- Inter-procedural, path-sensitive memory leak detection

Unit Testing

- GNU coreutils [Xu-Zhang 2006]
  - remove_suffix() in basename.c
  - cat() in cat.c
  - cut_bytes() in cut.c
  - parse_line() in dircolors.c
  - set_prefix() in fmt.c
  - attach() in ls.c

Memory Leak

- Whole programs
  - which
  - wget
  - gcc
  - zebra

Further Information

http://lex.ios.ac.cn/~zj/
zj@ios.ac.cn, jian_zhang@acm.org