A Global View of Formal Methods

Techniques for Improving Software Quality

Dr. James J. Hunt
aicas GmbH and aicas incorporated
Realtime and Embedded Systems Forum
26th of July 2007, Austin, TX
Definition of Formal Methods

analysis of software (and hardware) using rigorous mathematical methods such as calculi, logic, automata, or graph theory
Why Formal Methods?

• Errors can not be tolerated in safety critical applications.
• Security is not possible without safety.
• System complexity is increasing dramatically.
• Increasingly critical decisions are being made automatically in software.
• Testing is not good enough.
Strengths and Weaknesses of Testing

- **Strengths**
  - Well understood
  - Mostly language independent
  - Includes execution environment

- **Weaknesses**
  - Hard to cover all execution paths
  - Hard to cover all possible parallel paths
  - Internal states are not visible
Dynamic Analysis

- Testing
- Profiling & Monitoring
  - Path Tracing
  - Call tracing
  - Time tracing
  - Bounds tracing
- Simulation
- Instrumentation
  - Race detection
  - Assertion checking
  - Aliasing detection
  - Memory analysis
  - Invariant inference

Coverage must be determined!
Static Analysis (Formal)

- Type Analysis
- Control Flow Analysis
- Data Flow Analysis
- Abstract Interpretation
- Symbolic Execution
- Model Checking
- Deductive verification
Comparison of Analysis Techniques

Static analysis

- Abstract domain: slow but precise
- Conservative: due to abstraction
- Sound: due to conservatism

Dynamic analysis

- Concrete execution: slow but exhaustive
- Precise: no approximation
- Unsound: does not generalize
Type Checking

* Most common formal method
* Attributes used to ensure consistency
  - Ensure that a given variable or field is always used as intended
  - Limits what can be assigned to a given variable or field
* Base type can be augmented with refinement types
* Checking can be done modularly
Examples of Type Checking

- Unit consistency
  ```java
  @unit("meters") int a = 4;
  @unit("feet")   int b;
  b = a;  /* Assignment Error */
  ```

- Null pointer detection
- Invariance checking
- Tool examples
  - Most modern compilers
  - JavaCop
Control Flow Analysis

- Exhaustive search of all paths through a graph representing program execution
- Code divided into basic block and links
  - Basic block is a sequence of statements or instructions that do not change control flow
  - Links can be method or function calls, branches, and links to next basic block
- Features of graph can be identified more easily
Uses of Control Flow Analysis

- Worst case execution time analysis
- Stack analysis
- Test coverage analysis
- Reachability analysis
- Numerous tools on the market
Data Flow Analysis

- Extension of control flow analysis
- Data values are propagated as well
- Fixed point algorithm
- Necessary extension for OO Languages
  - Method dispatch is data dependent
  - More precise than considering all possible subclasses at each call point
Uses of Data Flow Analysis

- all uses of control flow analysis with more precision
- Exception checking
- Memory usage
- Shared object detection
- Synchronization (deadlocks)
- Tools emerging
Example DFA Tool from aicas

```
package com.aicas.jamaica;

public class Deadlock{
    public static void main(String[] args) {
        Thread firstThread = new Thread()
            new Runnable() {
                @Override
                public void run() {
                    synchronized (A) {
                        synchronized (B) {
                            // code...
                        }
                    }
                }
            }
        firstThread.start();

        Thread secondThread = new Thread()
            new Runnable() {
                @Override
                public void run() {
                    synchronized (B) {
                        synchronized (A) {
                            // code...
                        }
                    }
                }
            }
        secondThread.start();
    }
}
```
Abstract Interpretation

• A theory of sound approximation of the semantics of a program

• Concrete state and operations mapped to abstract state and operations

• Based on monotonic functions over ordered sets, especially lattices

• Can be viewed as a partial execution of a program to gain semantic information without performing all calculations
Uses of Abstract Interpretation

• Liveliness
• Race conditions
• Simultaneous access
• Tools

  Academic
  – ASTRÉE (CNRS)
  – Airac (SNU)

  Commercial
  – CodeHawk (KT)
  – PAG (AbsInt)
  – PolySpace
Symbolic Execution

- Also known as Symbolic Simulation
- Considers all possible execution paths
- Many possible executions of a system are considered simultaneously
- Models concrete semantics of all primitive operations (calculus)
- Set of values instead of concrete value
- Base for other techniques
Model Checking

• A variant of abstract interpretation
• Abstraction is a finite state machine
• Some aspect of program is modeled as states and transitions in state machine
• Both simulation and reachability analysis can be performed on state machine
• Error states are used to detect faults
Examples of Model Checking

• Model consistency (e.g., UML models)
• Checking parallel execution
• Some runtime errors (entry into a state)

• Numerous Tools
  Model Level
  - SPIN
  - PROSPER
  - Uppaal

  Java Programs
  - Pathfinder
Deductive Verification

• Uses formal specification language
  - Preconditions
  - Postconditions
  - invariants

• Checks program code against specification

• Based on theorem proving, Hoare Logic, and Liskov Substitution Principle
Formal Specification Languages

- Z notation (Specification)
- B method (Refinement)
- Object Constraint Language (OCL)
- Java Modeling Language (JML)
Examples of Deductive Verification

- Proving that a given Java method respects its post conditions given it preconditions
- Showing that invariants are respected
- Numerous Tools for Java (JML)
  - ESC/Java2 (Simplify)
  - JACK (B-Method, Simplify, PVS, Coq)
  - KeY (Dynamic Logic) (OCL & JML)
Which to use?

- Depends on what is to be checked and when in the development process
- Each tool has its strength and weaknesses and point of application
- A combination of tools works best
Formal Code Generation

- Based on abstract models and compiler theory
- Specify what instead of how
- Correct by construction
- Methods
  - Rule based selection and composition
  - Graph transformations
  - Algebraic transformations
Code from Abstract State Machine

- System is described as a finite state machine
- State machine description is translated into program code
- Generation can be done completely automatically
- Language less powerful than touring machine
Code from Metalanguage

- Program described in terms of transformations (what)
- Translator chooses between implementations (how)
- Transformations are composable
- Correctness must be proven for each implementation of each transformation
- User interaction often required
Formal Code Generation Tools

- State Machine
  - SCADE (Esterel)

- UML & Metalanguage
  - Perfect Developer (Escher Technologies)

- Meta Language
  - Specware (Kestrel Technology)
  - ACL2 (University of Texas)
Caveats

- Analysis is limited to information at hand
  - Check internal consistency
  - Compare alternative descriptions, e.g., code against specification
  - State explosion must be managed

- Code generation
  - may not produce efficient programs
  - Individual transitions must be checked

- Neither can fully replace testing
Conclusion

- Formal methods can drastically improve program quality.
- Can validate code against requirements.
- Different techniques for different aspects of interest.
- Can be combined to be more complete.
- More effective than test alone.
- Some testing will always be necessary.
Contact Information
aicas GmbH
Haid-und-Neu-Straße 18
D-76139 Karlsruhe
+49 721 663 968 22

aicas incorporated
69 West Rock Ave.
New Haven, CT 06515
+1 203 676 9807
jjh@aicas.com