



# EVALUATING CODE-BASED TEST INPUT GENERATOR TOOLS

Presented by Zoltan Micskei

## Code-based test generation

```
int fun1(int a, int b){  
    if (a == 0){  
1      printf(ERROR_MSG);  
2      return -1;  
    }  
    if (b > a)  
3      return b*a + 5;  
    else  
4      return (a+b) / 2;  
}
```

#	a	b	stm
1	0	*	1, 2
2	a!=0	b > a	3
3	a!=0	b < a	4

## Code-based test generator tools



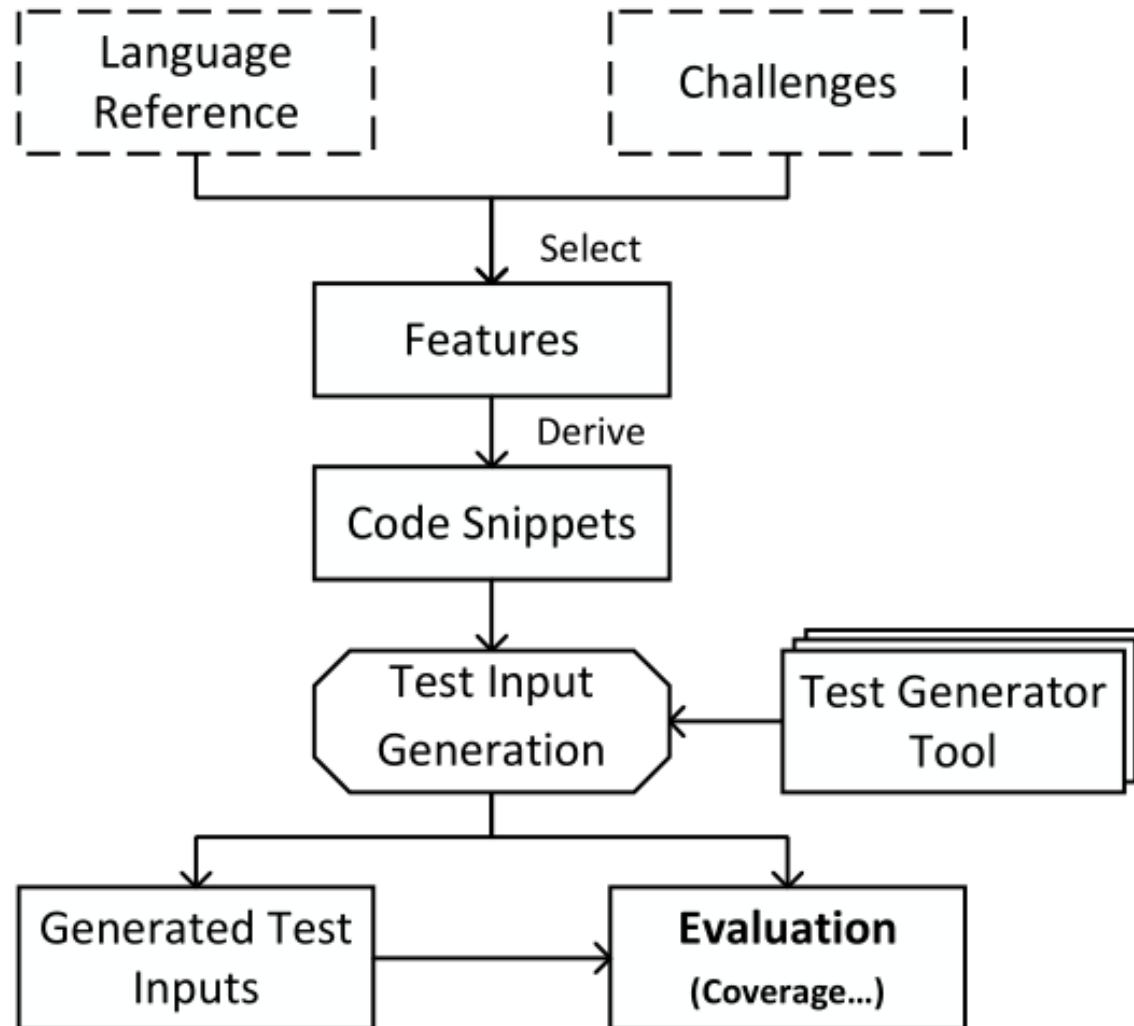
Detailed list of tools: <http://mit.bme.hu/~micskeiz/pages/cbtg.html>

## Motivation

How can the different  
test input generator tools  
**be compared and evaluated?**

# APPROACH

---



# Core features

- Types, Operators
- Conditionals, Loops
- Arrays, Functions

## Basic

- Structure usage
- Nested structures
- ...

## Structures

- Object usage
- Inheritance, interfaces
- ...

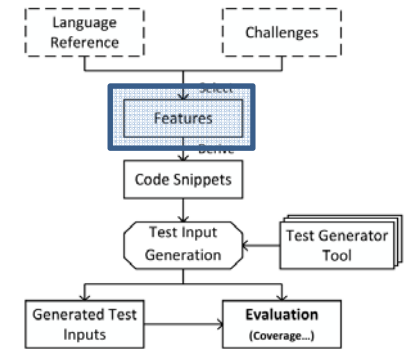
## Objects

- Generic functions
- Generic objects
- ...

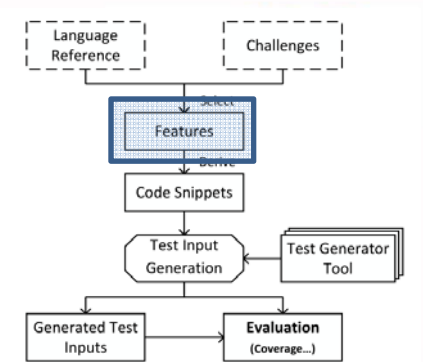
## Generics

- Arithmetic, Strings
- Collections
- ...

## Library



## Extra features



- Stdin, Properties
- Files, Sockets

Environment

- Threads, Locks
- Indeterminacy

Multi-threading

- Classes, Methods

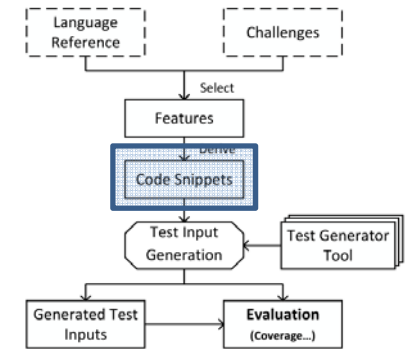
Reflection

- Native functions

Native code



# Snippets



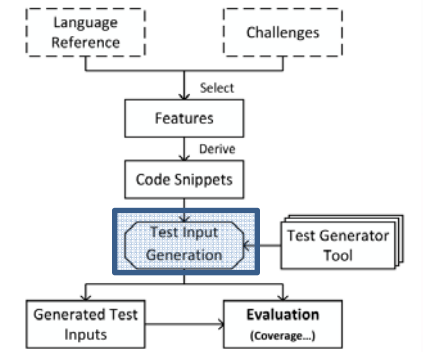
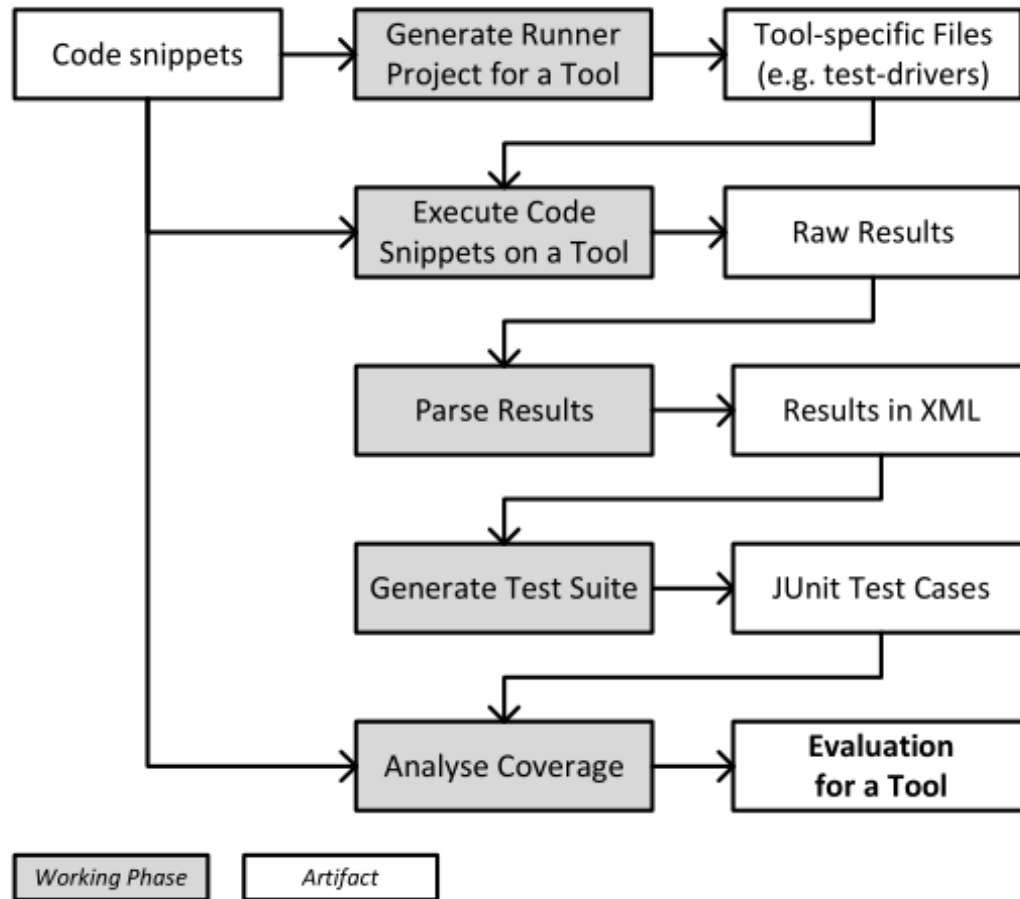
```
public static int guessObject (SimpleObject obj) {  
    if (obj == null) {  
        return -1;  
    }  
  
    if (obj.getOperationCount() == 2 && obj.getResult() == 3) {  
        return 1;  
    } else {  
        return 0;  
    }  
}
```

300 core snippets

63 extra snippets

Manual sample inputs

# Execution (SETTE framework)



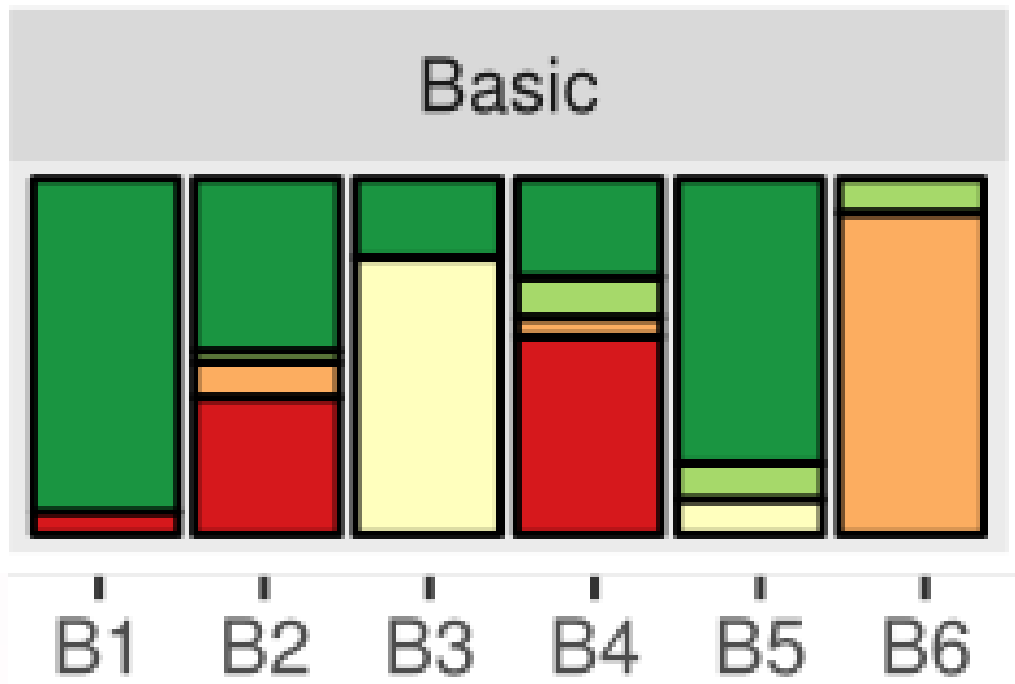
# RESULTS

---

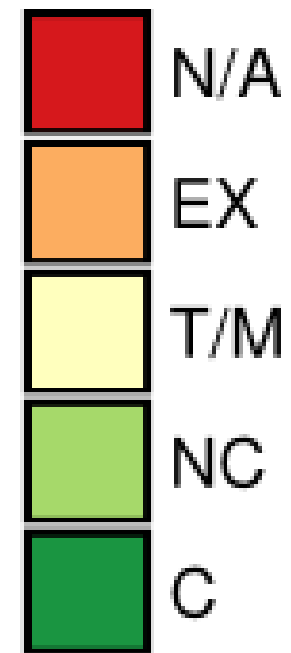
## Experiments

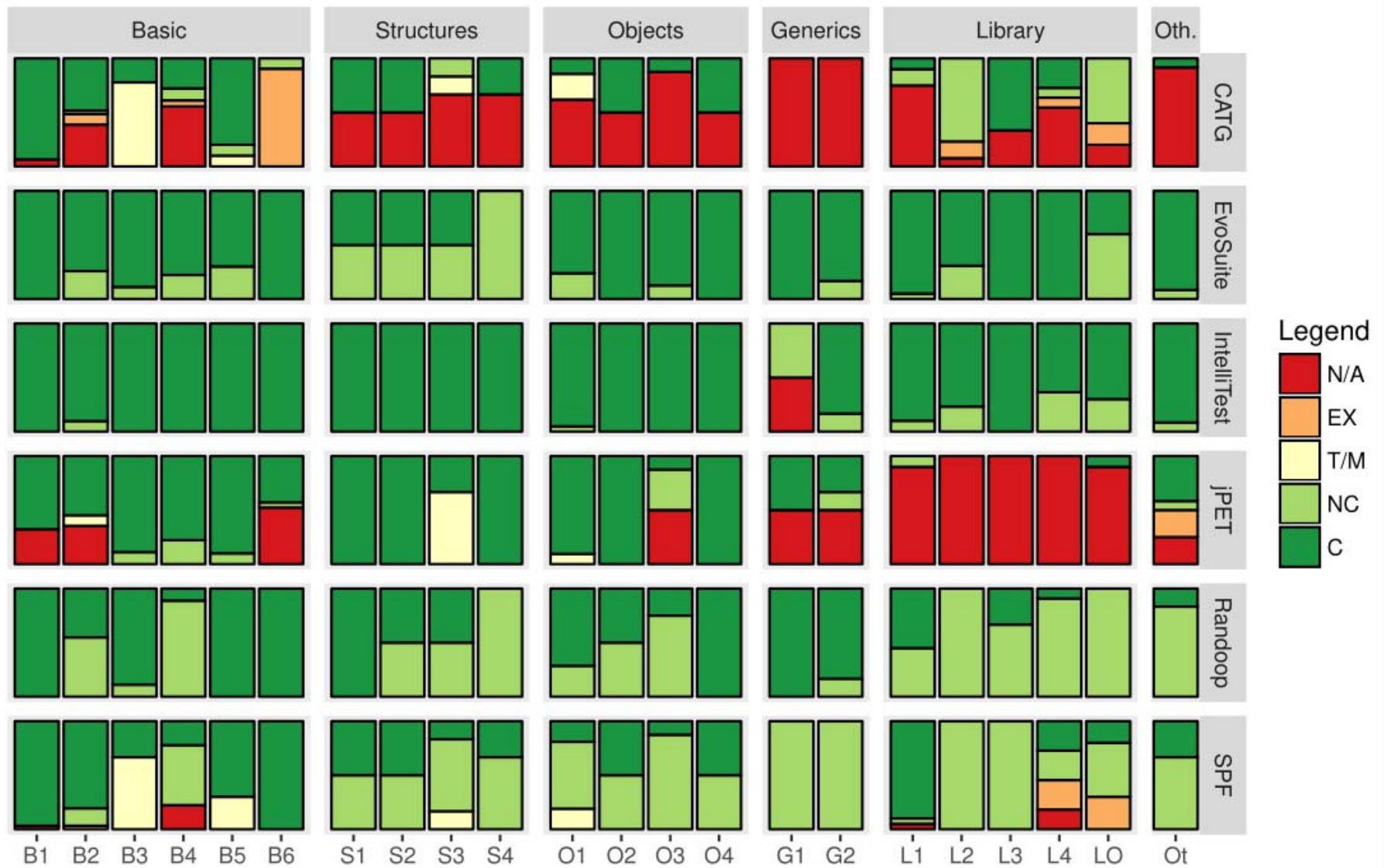
Tools	Status	Measures	Setup
<ul style="list-style-type: none"><li>• CATG</li><li>• EvoSuite</li><li>• IntelliTest</li><li>• jPET</li><li>• SPF</li><li>• Randoop</li></ul>	<ul style="list-style-type: none"><li>• N/A</li><li>• EX</li><li>• T/M</li><li>• NC</li><li>• C</li></ul>	<ul style="list-style-type: none"><li>• Coverage</li><li>• Size</li><li>• Duration</li><li>• Mutation score</li></ul>	<ul style="list-style-type: none"><li>• 30s limit</li><li>• Repeat 10x</li><li>• Variable time</li></ul>

# Overview of capabilities (Status)



## Legend

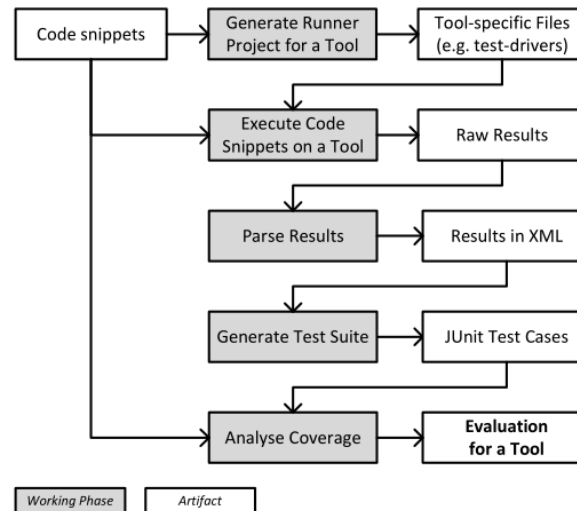
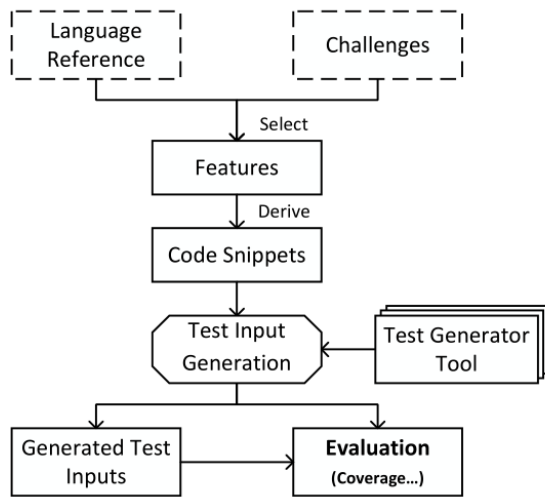




## Summary of insights

- **Duration:** 20 min – 200 min
  - Randoop, EvoSuite use all available time
- **Size:**
  - 700 (manual) < 1000 (IntelliTest) << 270 000 (Randoop)
- **Extra snippets**
  - EvoSuite: custom sandbox
  - Otherwise significant challenge

# Summary



All code and results are available:  
<http://sette-testing.github.io>