

NIER PAPER:

# Multi-Objective Black-box Test Case Prioritization based on WordNet Distances

IMARA VAN DINTEN<sup>[1]</sup>, ANDY ZAIDMAN<sup>[1]</sup>, ANNIBALE PANICHELLA<sup>[1]</sup>

<sup>[1]</sup>: DELFT UNIVERSITY OF TECHNOLOGY



This work has been partially supported by the European Union's Horizon 2020 Research and Innovation Programme under grant agreement No. 957254, project COSMOS.



# Test Case Prioritization Methods

White-box Test Case Prioritization, e.g.:

- Based on Coverage
- Code Changes

Access to the system's code.

Black-box Test Case Prioritization:

- Test Case Diversity
  - Textual similarity of the tests
- Their run duration

No access to the system's code.

#	Test Case
1	EmployeeHoliday
2	GradesSemesterThree
3	AddStudentCourseLate
4	SendEmailAllFaculty
5	UpdateEmployeePhoto
...	...

Test Case Prioritization

# Black-box Test Case Prioritization with Different Test types

Different types of tests

- Unit test → invoke methods/classes
- System tests → invoke components
- Simulation tests → ?

Test Case Prioritization

- Which test type(s) should we run first?
- How can [we compute test diversity?](#)



Multi-level Regression testing

# Textual Similarity

## Test Suite (TS1)

```
← → ↻ 🏠 ⓘ File | C:/Prep/JavaProjects/CRUD_Cucumber/target/cucumber-reports/index.html
▼ Feature: Test CRUD methods in Sample Employee REST API testing
  ▼ Scenario: Add Employee record
    Given I Set POST employee service api endpoint
    When I Set request HEADER
    And Send a POST HTTP request
    Then I receive valid Response
  ▼ Scenario: Update Employee record
    Given I Set PUT employee service api endpoint
    When I Set Update request Body
    And Send PUT HTTP request
    Then I receive valid HTTP response code 200
  ▼ Scenario: Get Employee record
    Given I Set GET employee service api endpoint
    When I Set request HEADER
    And Send GET HTTP request
    Then I receive valid HTTP response code 200
  ▼ Scenario: DELETE Employee record
    Given I Set DELETE employee service api endpoint
    When I Send DELETE HTTP request
    Then I receive valid HTTP response code 200
```

## Unit Test (T1)

```
@test
void testSalary(){
    int salary = calculateSalary(2,010, "professor");
    assertTrue(salary < 100,000); // poor academic
}
```

## Unit Test (T2)

```
@test
void testPassingExam(){
    int grade = Exam.calculateGrade(15.0, 1.0, 2.0);
    assertTrue(grade > 9.0); // great result
}
```

# Textual Similarity

Textual Similarity

- $D(TS1, T1) = 0$
- $D(TS1, T2) = 0$

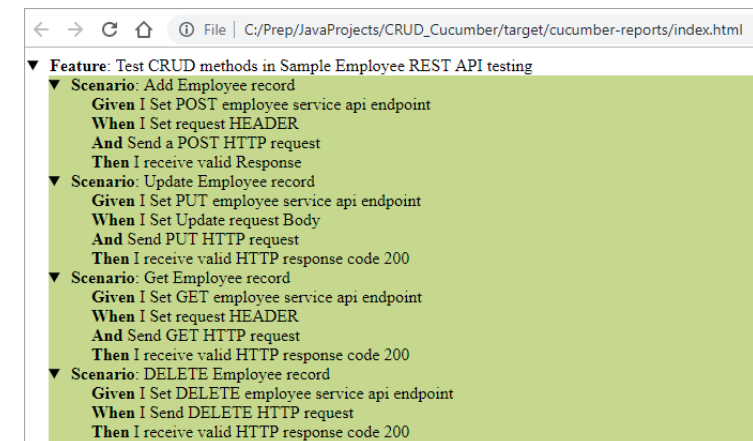
Semantics = Professor is a type of Employee

Semantics = Grade and Employee are distance concepts

Unit Test (T1)

```
@test
void testSalary(){
    int salary = calculateSalary(2,010, "professor");
    assertTrue(salary < 100,000); // poor academic
}
```

Test Suite (TS1)



File | C:/Prep/JavaProjects/CRUD\_Cucumber/target/cucumber-reports/index.html

- ▼ Feature: Test CRUD methods in Sample Employee REST API testing
  - ▼ Scenario: Add Employee record
    - Given I Set POST employee service api endpoint
    - When I Set request HEADER
    - And Send a POST HTTP request
    - Then I receive valid Response
  - ▼ Scenario: Update Employee record
    - Given I Set PUT employee service api endpoint
    - When I Set Update request Body
    - And Send PUT HTTP request
    - Then I receive valid HTTP response code 200
  - ▼ Scenario: Get Employee record
    - Given I Set GET employee service api endpoint
    - When I Set request HEADER
    - And Send GET HTTP request
    - Then I receive valid HTTP response code 200
  - ▼ Scenario: DELETE Employee record
    - Given I Set DELETE employee service api endpoint
    - When I Send DELETE HTTP request
    - Then I receive valid HTTP response code 200

# Textual Similarity and WordNet similarity

Textual (COSINE) similarity

- Each word is considered separate, not using relationships between words.

Professor and Employee are **NOT** similar.

WordNet similarity

- Semantic similarity of words.

Professor is **a type of** Employee

# WordNet Similarity Metrics

Metrics using:

1. a thesaurus (e.g., WordNet)
  - Path Similarity
  - Leacock-Chodorow Similarity
  - **Wu-Palmer Similarity**
- a thesaurus and probabilistic information
  - Resnik Similarity
  - Lin Similarity

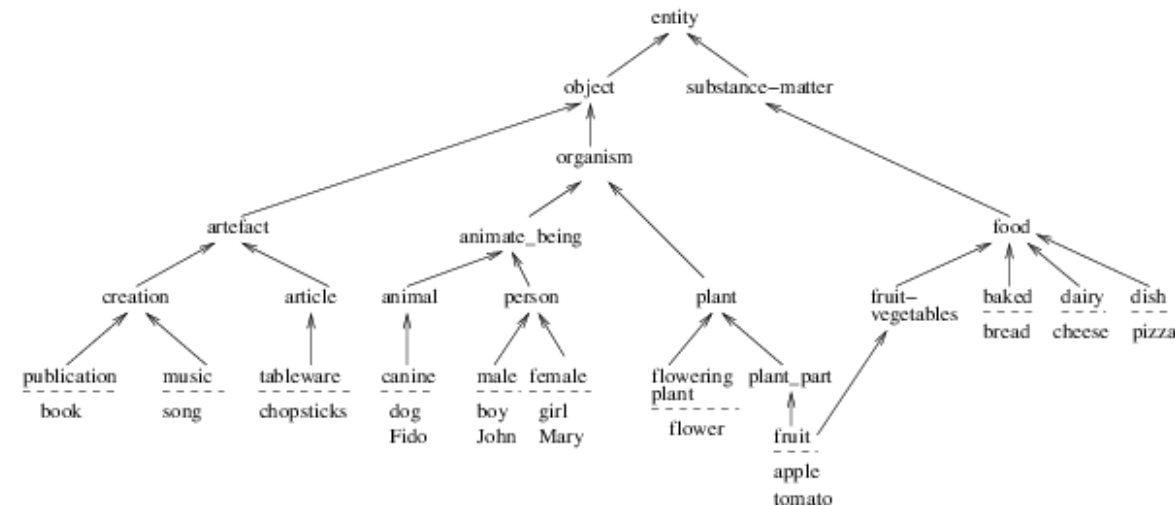


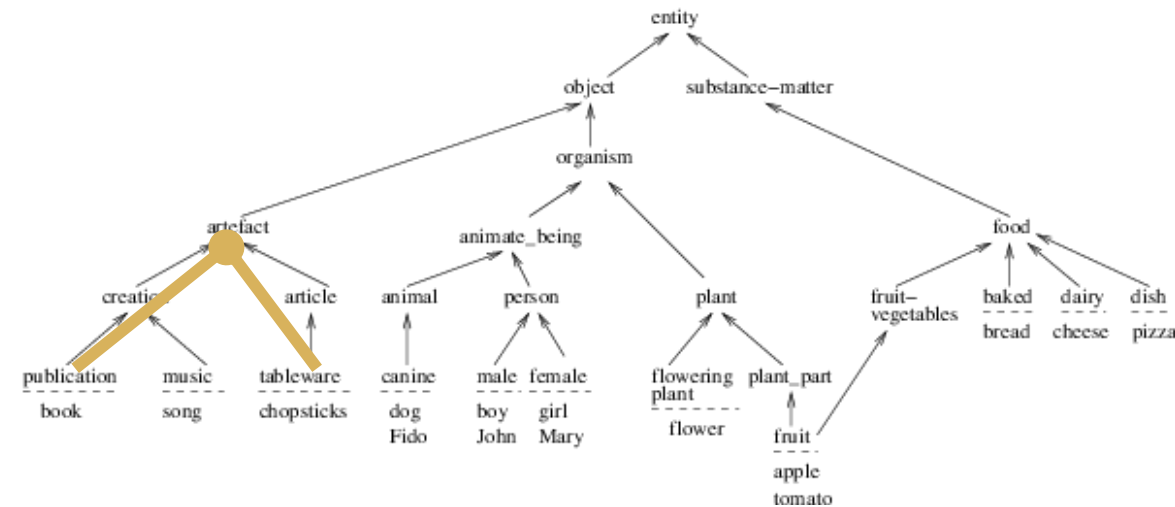
Image from:

Kazakov, Dimitar, and Simon Dobnik. "Inductive learning of lexical semantics with typed unification grammars." *Topics in Phonetics and Computational Linguistics* 8 (2003): 113-133.

# WordNet Similarity Metrics

Metrics using WordNet:

- Path Similarity
- Leacock-Chodorow Similarity
- Wu-Palmer Similarity



$$\text{Sim} = \text{pathlen}(w1, w2)$$

Image from:

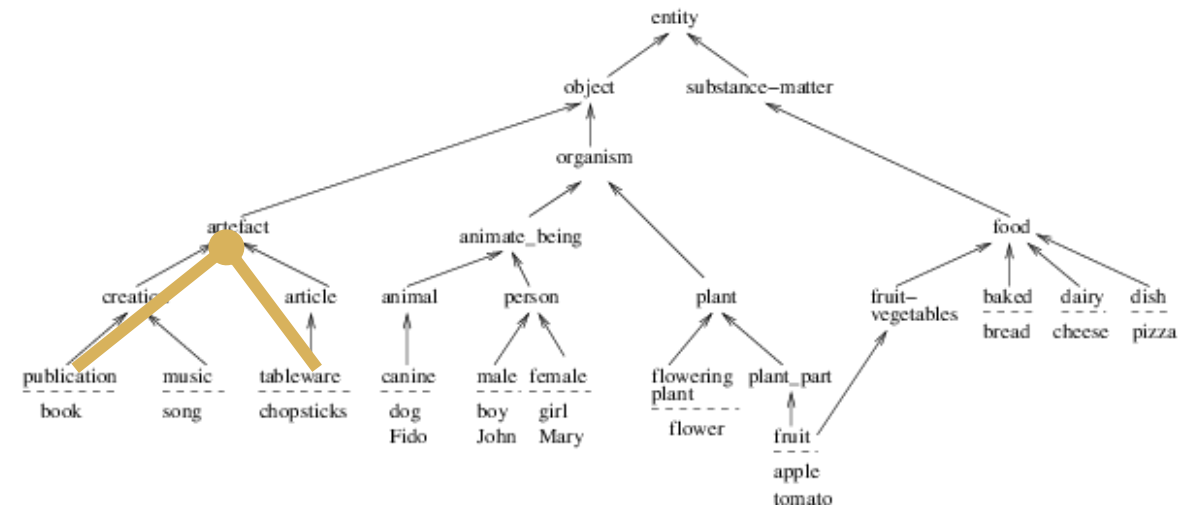
Kazakov, Dimitar, and Simon Dobnik. "Inductive learning of lexical semantics with typed unification grammars." Topics in Phonetics and Computational Linguistics 8 (2003): 113-133.



# WordNet Similarity Metrics

Metrics using WordNet:

- Path Similarity
- Leacock-Chodorow Similarity
- Wu-Palmer Similarity



$$\text{Sim} = -\log(\text{pathlen}(w_1, w_2))$$

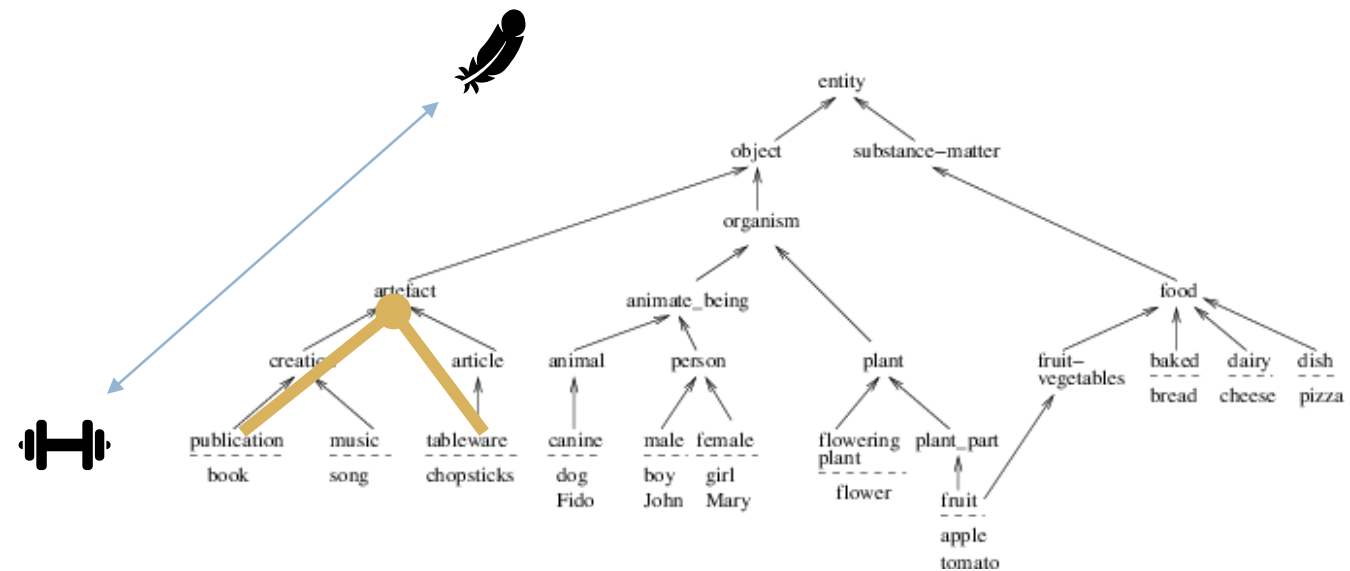
Image from:

Kazakov, Dimitar, and Simon Dobnik. "Inductive learning of lexical semantics with typed unification grammars." Topics in Phonetics and Computational Linguistics 8 (2003): 113-133.

# WordNet Similarity Metrics

Metrics using WordNet:

- Path Similarity
- Leacock-Chodorow Similarity
- Wu-Palmer Similarity



$$\text{Sim} = -\log(\text{pathlen}(\text{w1}, \text{w2}))$$

Image from:

Kazakov, Dimitar, and Simon Dobnik. "Inductive learning of lexical semantics with typed unification grammars." Topics in Phonetics and Computational Linguistics 8 (2003): 113-133.

# Semantic Distance with WordNet

Unit Test (T1)

**WS4J Demo**  
 WS4J (WordNet Similarity for Java) measures semantic similarity/relatedness between words

Type in texts below, or use:

1. Input mode  Word  Sentence

2. Sentence 1  
 salary salary calculate salary professor salary poor academic

3. Sentence 2  
 employee

4. Submit

---

**WUP**

	salary /NN	salary /NN	calculate /JJ	salary /NN	professor /NN	salary /NN	poor /JJ	academic /JJ
employee/NN	0.2222	0.2222		-0.2222	0.6957	0.2222		

Semantic Similarity = 0.6957

Unit Test (T2)

**WS4J Demo**  
 WS4J (WordNet Similarity for Java) measures semantic similarity/relatedness between words

Requested the server to load WordNet on server-side memory.

WordNet loading status: Loading

Type in texts below, or use:

1. Input mode  Word  Sentence

2. Sentence 1  
 passing exam  
 grade exam grade  
 grade great result

3. Sentence 2  
 employee

4. Submit

---

**WUP**

	passing /VBG	exam /NN	grade /NN	exam /NN	grade /NN	grade /NN	great /JJ	result /NN
employee/NN		-0.2667	0.4828	0.2667	0.4828	0.4828		-0.4615

Semantic Similarity = 0.4828

# Textual Similarity and WordNet similarity: Improve Test Diversity

Textual (COSINE) similarity

- Each word is considered separate, not using relationships between words.

Professor and Employee are **NOT** similar.

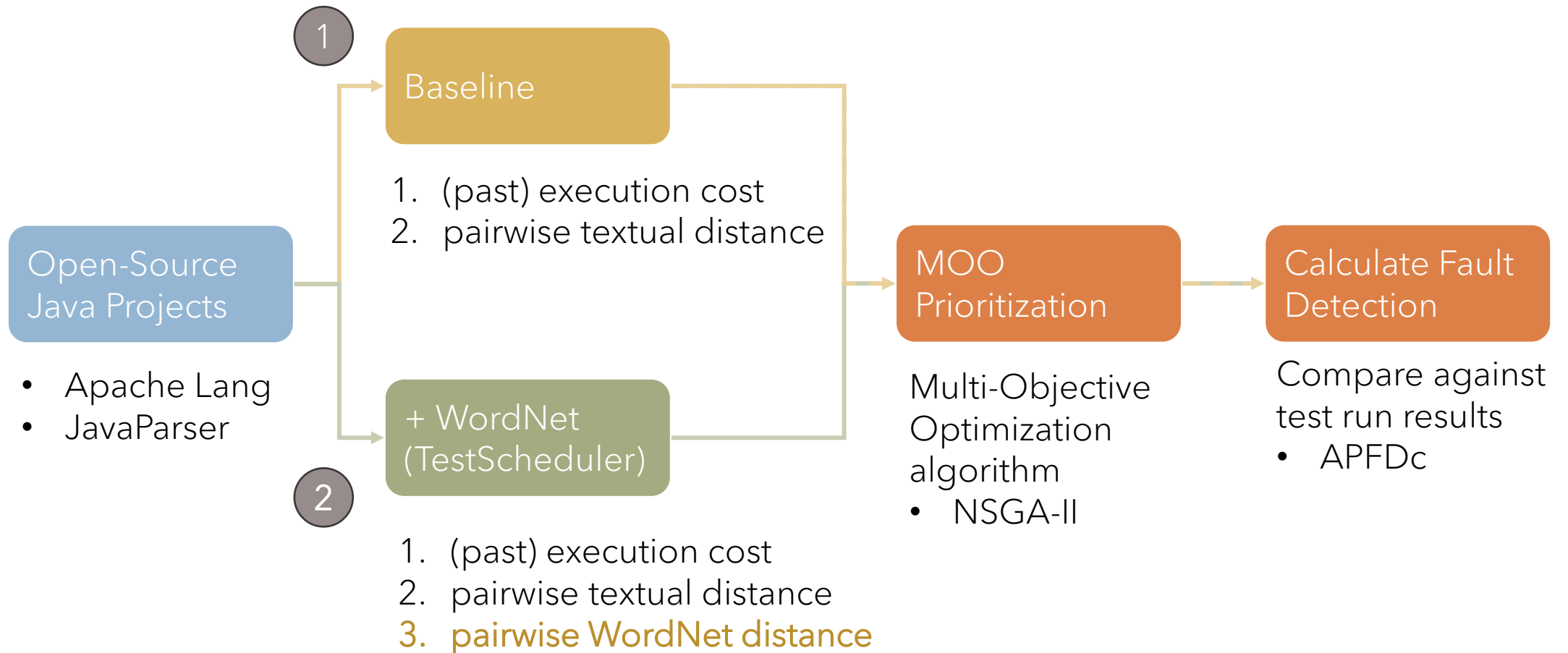
WordNet similarity

- Semantic similarity of words.

Professor is a **type** of Employee

RQ: To what extent does the use of WordNet improve the effectiveness of diversity-based test case prioritization?

# Study Setup



# Preliminary Results

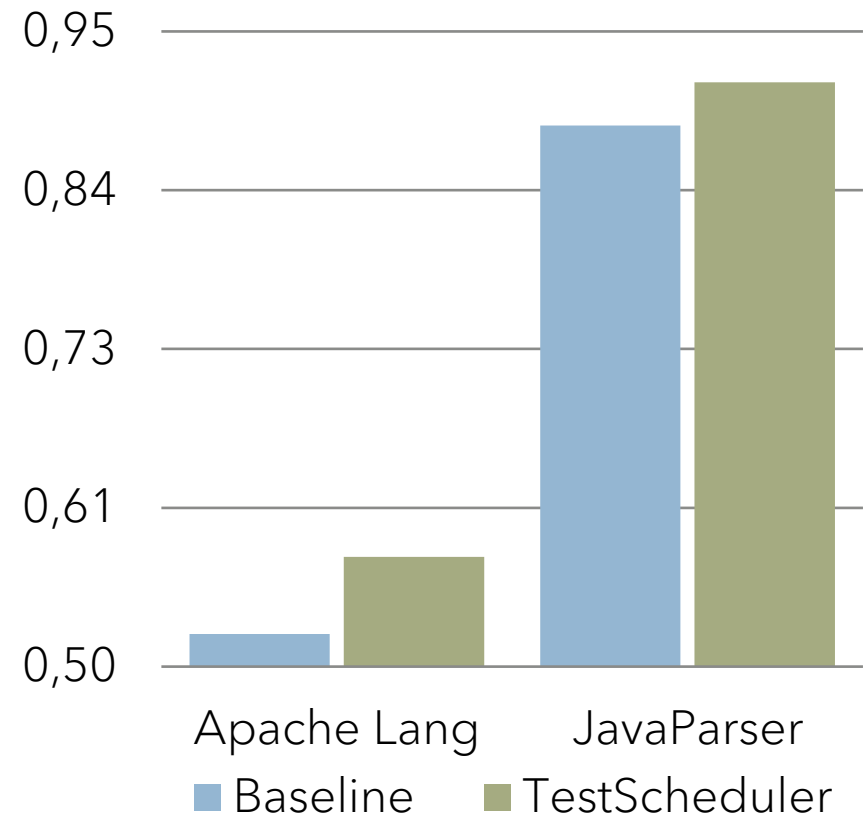
## TestScheduler:

- Multi-objective search for test case prioritization
  - ✓ Minimize (past) execution cost
  - ✓ Maximize pairwise WordNet distance
  - ✓ Maximize pairwise textual distance

## Baseline:

- Multi-objective search for test case prioritization
  - ✓ Minimize (past) execution cost
  - ✓ Maximize pairwise textual distance

Average Percentage of Fault Detection (APFDc)



# Summary and Future Work

Adding WordNet's Semantic Similarity improves the effectiveness of diversity-based test case prioritization.

For future work:

- + Expand project(s) selection, larger project(s)
- + Different WordNet Similarity calculations
- + Large Language Models (LLMs)



Multi-objective Black-Box  
Test Case Prioritization  
Based on Wordnet Distances



✉ i.vandinten@tudelft.nl  
🌐 imaravandinten.com



✉ a.e.zaidman@tudelft.nl  
🌐 azaidman.github.io



✉ a.panichella@tudelft.nl  
🌐 apanichella.github.io