



SOFTWARE ENGINEERING

Elite Graduate Program

TUM

TECHNISCHE
UNIVERSITÄT
MÜNCHEN

UNA

Universität
Augsburg
University



 **TU Delft**

Carolin E. Brandt



TECHNISCHE
UNIVERSITÄT
MÜNCHEN



Universität
Augsburg
University



How to Analyze Build Logs



SOFTWARE ENGINEERING

Elite Graduate Program

TUM

TECHNISCHE
UNIVERSITÄT
MÜNCHEN

UNA

Universität
Augsburg
University

LMU

 **TU Delft**

How to Analyze Build Logs

-

A Comparative Study of Chunk Retrieval Techniques

Carolyn E. Brandt



Travis CI

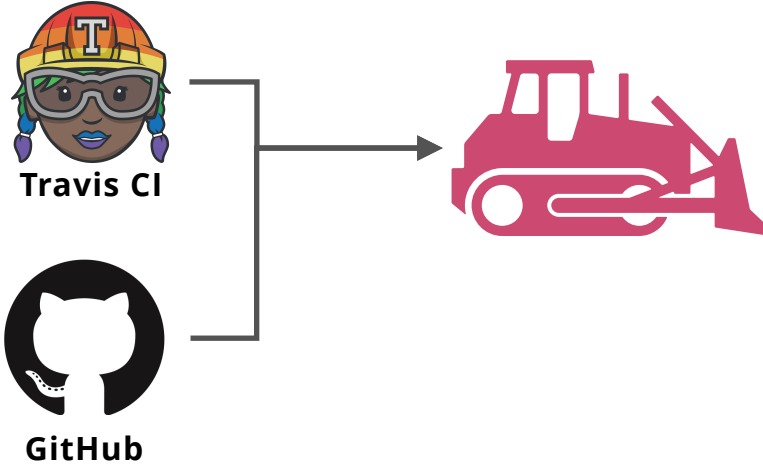


Travis CI



GitHub

Build Logs



Build Logs



Travis CI

```

name: 'Build system'
language: 'shell'
os: 'linux'
dist: 'bionic'
before_install:
  - sudo apt-get update
  - sudo apt-get install -y curl
  - curl -sS https://api.github.com/repos/travis-ci/docker | xargs git clone --depth 1 -q https://github.com
  - cd docker
  - ./bootstrap.sh
  - cd ..
install:
  - curl -sS https://api.github.com/repos/travis-ci/docker | xargs git clone --depth 1 -q https://github.com
  - cd docker
  - ./bootstrap.sh
  - cd ..
script:
  - echo "Build system is ready"

```

```

name: 'Build system'
language: 'shell'
os: 'linux'
dist: 'bionic'
before_install:
  - sudo apt-get update
  - sudo apt-get install -y curl
  - curl -sS https://api.github.com/repos/travis-ci/docker | xargs git clone --depth 1 -q https://github.com
  - cd docker
  - ./bootstrap.sh
  - cd ..
install:
  - curl -sS https://api.github.com/repos/travis-ci/docker | xargs git clone --depth 1 -q https://github.com
  - cd docker
  - ./bootstrap.sh
  - cd ..
script:
  - echo "Build system is ready"

```

```

name: 'Build system'
language: 'shell'
os: 'linux'
dist: 'bionic'
before_install:
  - sudo apt-get update
  - sudo apt-get install -y curl
  - curl -sS https://api.github.com/repos/travis-ci/docker | xargs git clone --depth 1 -q https://github.com
  - cd docker
  - ./bootstrap.sh
  - cd ..
install:
  - curl -sS https://api.github.com/repos/travis-ci/docker | xargs git clone --depth 1 -q https://github.com
  - cd docker
  - ./bootstrap.sh
  - cd ..
script:
  - echo "Build system is ready"

```


- **Log structure changes from project to project**
- **Developing custom parsers is a lot of work**

- Log structure changes from project to project
- Developing custom parsers is a lot of work



**Chunk Retrieval
Techniques**

- Log structure changes from project to project
- Developing custom parsers is a lot of work



**Chunk Retrieval
Techniques**

Program Synthesis by Example - PBE

- Log structure changes from project to project
- Developing custom parsers is a lot of work



**Chunk Retrieval
Techniques**

Program Synthesis by Example - PBE

Common Text Similarity - CTS

- Log structure changes from project to project
- Developing custom parsers is a lot of work



**Chunk Retrieval
Techniques**

Program Synthesis by Example - PBE

Common Text Similarity - CTS

Keyword Search - KWS

- Log structure changes from project to project
- Developing custom parsers is a lot of work



**Chunk Retrieval
Techniques**

Program Synthesis by Example (PBE)
Common Text Similarity (CTS)
Keyword Search - KWS

**When should I use which
technique?**

- Log structure changes from project to project
- Developing custom parsers is a lot of work



**Chunk Retrieval
Techniques**

Problem

**When should I use which
technique?**

Program Synthesis by Example (PBE)
Common Text Similarity (CTS)
Keyword Search - KWS

- Log structure changes from project to project
- Developing custom parsers is a lot of work



Chunk Retrieval
Techniques

Solution

Problem

Program Synthesis by Example (PBE)
Common Text Similarity (CTS)
Keyword Search - KWS

When should I use which technique?

- Log structure changes from project to project
- Developing custom parsers is a lot of work



Chunk Retrieval
Techniques

Solution

Problem

Contribution

Program Synthesis by Example (PBE)
Common Text Similarity (CTS)
Keyword Search - KWS

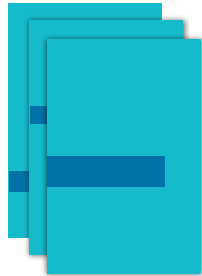
When should I use which technique?



**Implement
Techniques**



**Implement
Techniques**

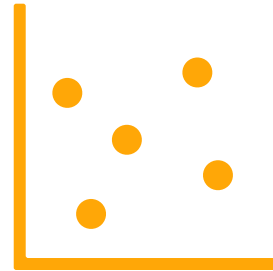


**Create Data Set:
*LogChunks***

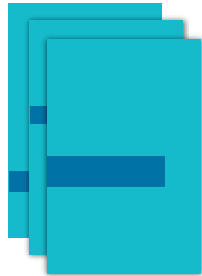
Approach



**Implement
Techniques**



**Empirical
Comparison
Study**



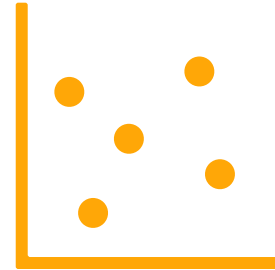
**Create Data Set:
*LogChunks***



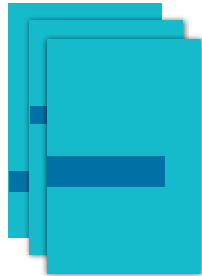
Approach



Implement
Techniques



Empirical
Comparison
Study



Create Data Set:
LogChunks



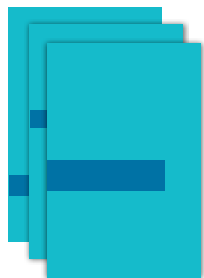
| PBE | | | |
|-----|--|--|--|
| CTS | | | |
| KWS | | | |

Recommendations
For Each
Technique

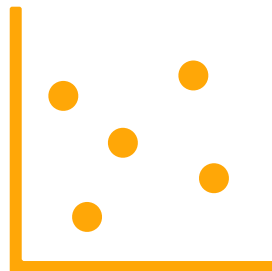
Approach



Implement
Techniques



Create Data Set:
LogChunks

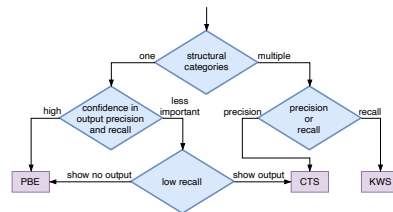


Empirical
Comparison
Study



| PBE | | | |
|-----|--|--|--|
| CTS | | | |
| KWS | | | |

Recommendations
For Each
Technique

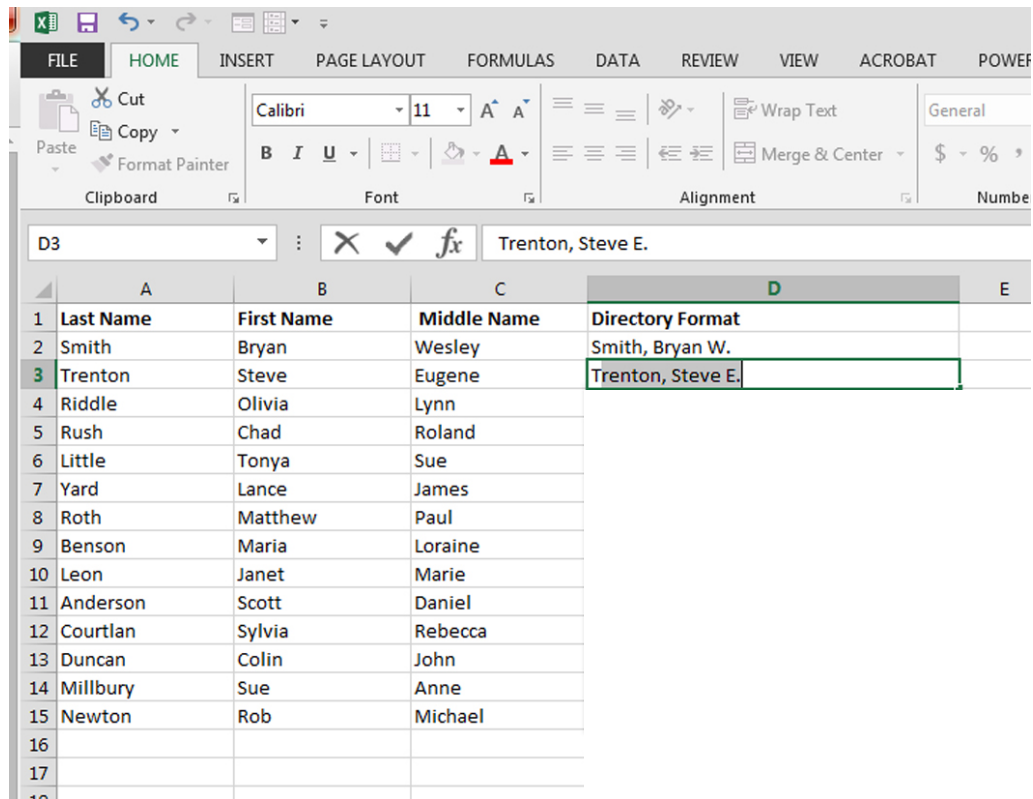


When to Use
Which Technique

**Microsoft
PROgram
Synthesis using
Examples SDK**

Program Synthesis By Example

Microsoft PROgram Synthesis using Examples SDK

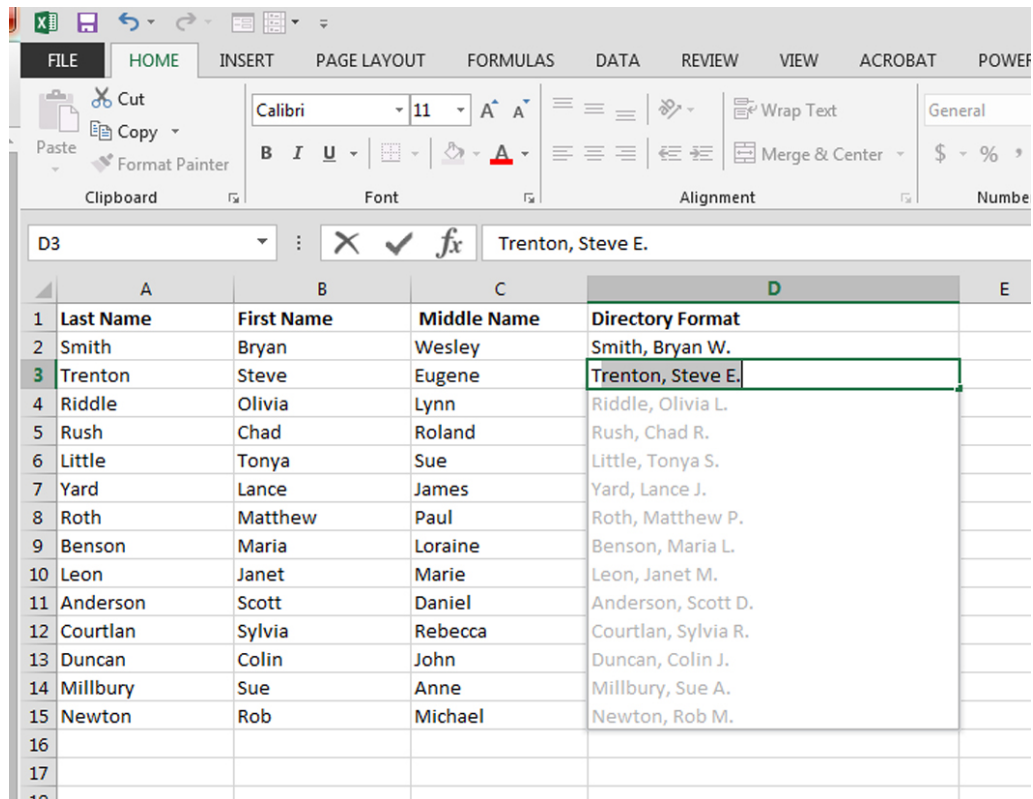


The screenshot shows the Microsoft Excel interface. The ribbon is set to 'HOME'. The active cell is D3, containing the text 'Trenton, Steve E.'. The formula bar shows the same text. The spreadsheet data is as follows:

| | A | B | C | D | E |
|----|-----------|------------|-------------|-------------------|---|
| 1 | Last Name | First Name | Middle Name | Directory Format | |
| 2 | Smith | Bryan | Wesley | Smith, Bryan W. | |
| 3 | Trenton | Steve | Eugene | Trenton, Steve E. | |
| 4 | Riddle | Olivia | Lynn | | |
| 5 | Rush | Chad | Roland | | |
| 6 | Little | Tonya | Sue | | |
| 7 | Yard | Lance | James | | |
| 8 | Roth | Matthew | Paul | | |
| 9 | Benson | Maria | Loraine | | |
| 10 | Leon | Janet | Marie | | |
| 11 | Anderson | Scott | Daniel | | |
| 12 | Courtlan | Sylvia | Rebecca | | |
| 13 | Duncan | Colin | John | | |
| 14 | Millbury | Sue | Anne | | |
| 15 | Newton | Rob | Michael | | |
| 16 | | | | | |
| 17 | | | | | |
| 18 | | | | | |

Program Synthesis By Example

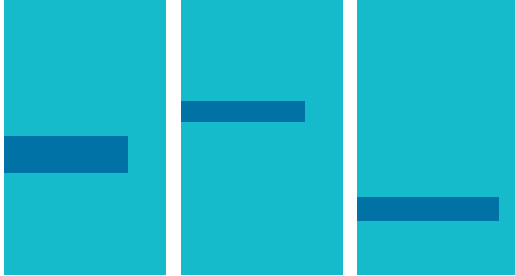
Microsoft PROgram Synthesis using Examples SDK



The screenshot shows the Microsoft Excel interface. The ribbon is set to 'HOME'. The active cell is D3, containing the text 'Trenton, Steve E.'. A dropdown menu is open, showing a list of names in 'Directory Format' based on the data in the table below. The names are: Smith, Bryan W.; Trenton, Steve E. (highlighted); Riddle, Olivia L.; Rush, Chad R.; Little, Tonya S.; Yard, Lance J.; Roth, Matthew P.; Benson, Maria L.; Leon, Janet M.; Anderson, Scott D.; Courtlan, Sylvia R.; Duncan, Colin J.; Millbury, Sue A.; and Newton, Rob M.

| | A | B | C | D | E |
|----|-----------|------------|-------------|---------------------|---|
| 1 | Last Name | First Name | Middle Name | Directory Format | |
| 2 | Smith | Bryan | Wesley | Smith, Bryan W. | |
| 3 | Trenton | Steve | Eugene | Trenton, Steve E. | |
| 4 | Riddle | Olivia | Lynn | Riddle, Olivia L. | |
| 5 | Rush | Chad | Roland | Rush, Chad R. | |
| 6 | Little | Tonya | Sue | Little, Tonya S. | |
| 7 | Yard | Lance | James | Yard, Lance J. | |
| 8 | Roth | Matthew | Paul | Roth, Matthew P. | |
| 9 | Benson | Maria | Loraine | Benson, Maria L. | |
| 10 | Leon | Janet | Marie | Leon, Janet M. | |
| 11 | Anderson | Scott | Daniel | Anderson, Scott D. | |
| 12 | Courtlan | Sylvia | Rebecca | Courtlan, Sylvia R. | |
| 13 | Duncan | Colin | John | Duncan, Colin J. | |
| 14 | Millbury | Sue | Anne | Millbury, Sue A. | |
| 15 | Newton | Rob | Michael | Newton, Rob M. | |
| 16 | | | | | |
| 17 | | | | | |
| 18 | | | | | |

Program Synthesis By Example

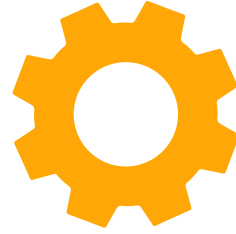


In/Output Examples

Program Synthesis By Example



In/Output Examples

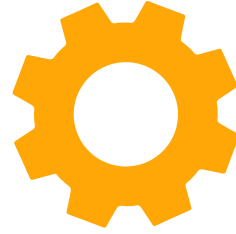


**PROSE
Synthesis**

Program Synthesis By Example



In/Output Examples



**PROSE
Synthesis**

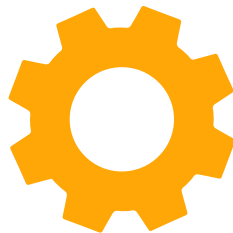


RegEx Program

Program Synthesis By Example



In/Output Examples



PROSE
Synthesis



RegEx Program

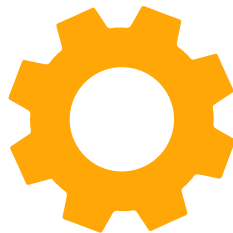


Divide & Conquer

Program Synthesis By Example



In/Output Examples



PROSE
Synthesis



RegEx Program

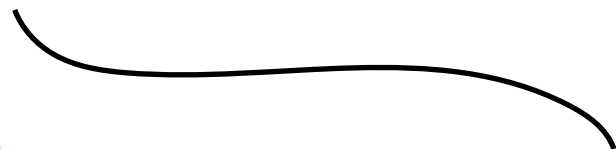


Divide & Conquer

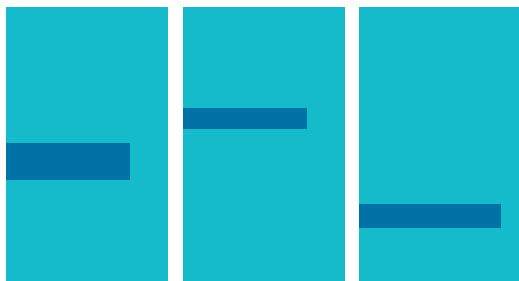


- `\n.*[A-Z]{5}`
- `\s[^\\-\\.]Error(s)?:`
- `\-{10} FAILURES \-{10}`

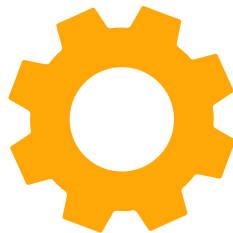
Enumerate



Program Synthesis By Example



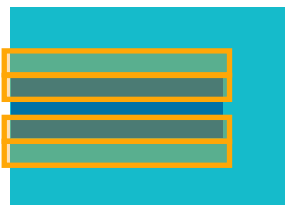
In/Output Examples



PROSE
Synthesis



RegEx Program



Divide & Conquer



- `\n.*[A-Z]{5}`
- `\s{^\-\.}Error(s)?:`
- `\-{10} FAILURES \-{10}`

Enumerate



1. `\-{10} FAILURES \-{10}`
2. `\n.*[A-Z]{5}`
3. `\s{^\-\.}Error(s)?:`

Ranking

PBE

Example

PBE

Example

```
FAILURE: Build failed with an exception.
```

```
* What went wrong:
```

```
Execution failed for task ':app:connectedAndroidTest'.
```

```
> com.android.builder.testing.api.DeviceException: java.lang.RuntimeException: ...
```

```
* Try:
```

```
FAILURE: Build failed with an exception.
```

```
* What went wrong:
```

```
Could not determine the dependencies of task ':app:jacocoTestDebugReport'.
```

```
> Task with path 'testDebug' not found in project ':app'.
```

```
* Try:
```

PBE

Example

FAILURE: Build failed with an **exception**.

* What went wrong:

Execution failed for task ':app:connectedAndroidTest'.

> `com.android.builder.testing.api.DeviceException: java.lang.RuntimeException: ...`

* Try:

FAILURE: Build failed with an **exception**.

* What went wrong:

Could not determine the dependencies of task ':app:jacocoTestDebugReport'.

> Task with path '`testDebug`' not found in project ':app'.

* Try:

PBE

Example

```
FAILURE: Build failed with an exception.
```

```
* What went wrong:
```

```
Execution failed for task ':app:connectedAndroidTest'.
```

```
> com.android.builder.testing.api.DeviceException: java.lang.RuntimeException: ...
```

```
* Try:
```

```
FAILURE: Build failed with an exception.
```

```
* What went wrong:
```

```
Could not determine the dependencies of task ':app:jacocoTestDebugReport'.
```

```
> Task with path 'testDebug' not found in project ':app'.
```

```
* Try:
```

```
StartExtraction(RegexPair("Colon°Line Separator", "ALL CAPS")) in  
EndExtraction(RegexPair("ε", "Line Separator°Line Separator"))
```

PBE

Example

```
FAILURE: Build failed with an exception.
```

```
* What went wrong:
```

```
Execution failed for task ':app:connectedAndroidTest'.
```

```
> com.android.builder.testing.api.DeviceException: java.lang.RuntimeException: ...
```

```
* Try:
```

```
FAILURE: Build failed with an exception.
```

```
* What went wrong:
```

```
Could not determine the dependencies of task ':app:jacocoTestDebugReport'.
```

```
> Task with path 'testDebug' not found in project ':app'.
```

```
* Try:
```

```
StartExtraction(RegexPair("Colon°Line Separator", "ALL CAPS")) in  
EndExtraction(RegexPair("ε", "Line Separator°Line Separator"))
```

```
=== RUN    TestSeparator
```

```
--- FAIL: TestSeparator (0.03s)
```

```
    repo_test.go:75: expected entry to be in form of `* [link] - description.` , got `* ...`
```

```
=== RUN    TestGenerateHTML
```

PBE

Example

```
FAILURE: Build failed with an exception.
```

```
* What went wrong:
```

```
Execution failed for task ':app:connectedAndroidTest'.
```

```
> com.android.builder.testing.api.DeviceException: java.lang.RuntimeException: ...
```

```
* Try:
```

```
FAILURE: Build failed with an exception.
```

```
* What went wrong:
```

```
Could not determine the dependencies of task ':app:jacocoTestDebugReport'.
```

```
> Task with path 'testDebug' not found in project ':app'.
```

```
* Try:
```

```
StartExtraction(RegexPair("Colon°Line Separator", "ALL CAPS")) in  
EndExtraction(RegexPair("ε", "Line Separator°Line Separator"))
```

```
=== RUN TestSeparator
```

```
--- FAIL: TestSeparator (0.03s)
```

```
repo_test.go:75: expected entry to be in form of `* [link] - description.` , got `* ...`
```

```
=== RUN TestGenerateHTML
```

```
no program found
```

PBE

Example

```
FAILURE: Build failed with an exception.
```

```
* What went wrong:
```

```
Execution failed for task ':app:connectedAndroidTest'.
```

```
> com.android.builder.testing.api.DeviceException: java.lang.RuntimeException: ...
```

```
* Try:
```

```
FAILURE: Build failed with an exception.
```

```
* What went wrong:
```

```
Could not determine the dependencies of task ':app:jacocoTestDebugReport'.
```

```
> Task with path 'testDebug' not found in project ':app'.
```

```
* Try:
```

```
StartExtraction(RegexPair("Colon°Line Separator", "ALL CAPS")) in  
EndExtraction(RegexPair("ε", "Line Separator°Line Separator"))
```

```
=== RUN TestSeparator
```

```
--- FAIL: TestSeparator (0.03s)
```

```
repo_test.go:75: expected entry to be in form of `* [link] - description.` , got `* ...`
```

```
=== RUN TestGenerateHTML
```

```
no program found
```

PBE

Example

Structural Category

```
FAILURE: Build failed with an exception.
```

```
* What went wrong:
```

```
Execution failed for task ':app:connectedAndroidTest'.
```

```
> com.android.builder.testing.api.DeviceException: java.lang.RuntimeException: ...
```

```
* Try:
```

```
FAILURE: Build failed with an exception.
```

```
* What went wrong:
```

```
Could not determine the dependencies of task ':app:jacocoTestDebugReport'.
```

```
> Task with path 'testDebug' not found in project ':app'.
```

```
* Try:
```

```
StartExtraction(RegexPair("Colon°Line Separator", "ALL CAPS")) in  
EndExtraction(RegexPair("ε", "Line Separator°Line Separator"))
```

```
=== RUN TestSeparator
```

```
--- FAIL: TestSeparator (0.03s)
```

```
repo_test.go:75: expected entry to be in form of `* [link] - description.` , got `* ...`
```

```
=== RUN TestGenerateHTML
```

```
no program found
```


LogChunks Data Set



GitHub



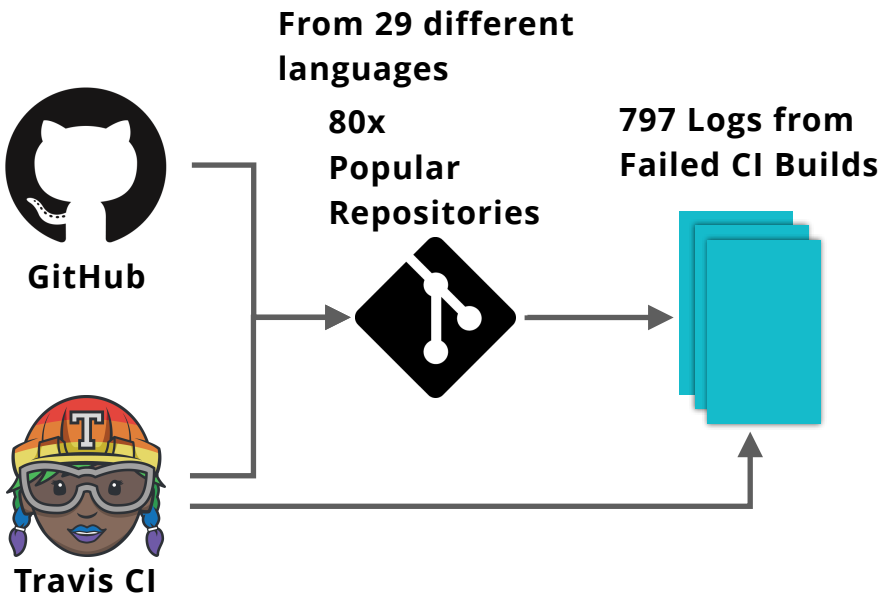
Travis CI

From 29 different
languages

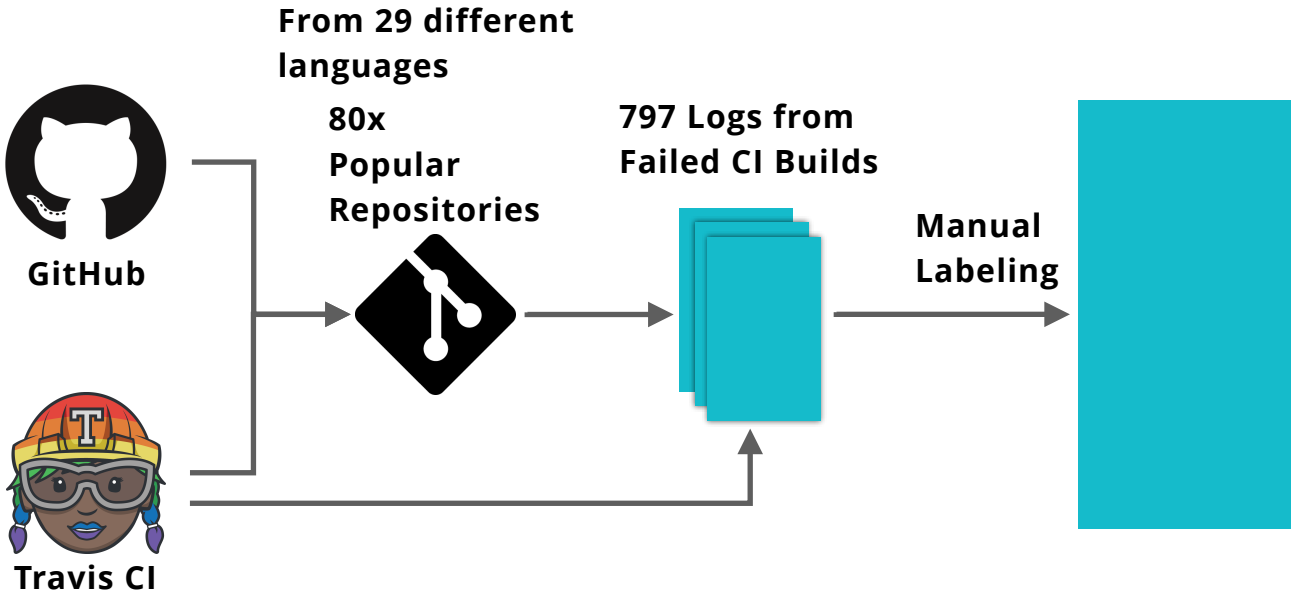
80x
Popular
Repositories



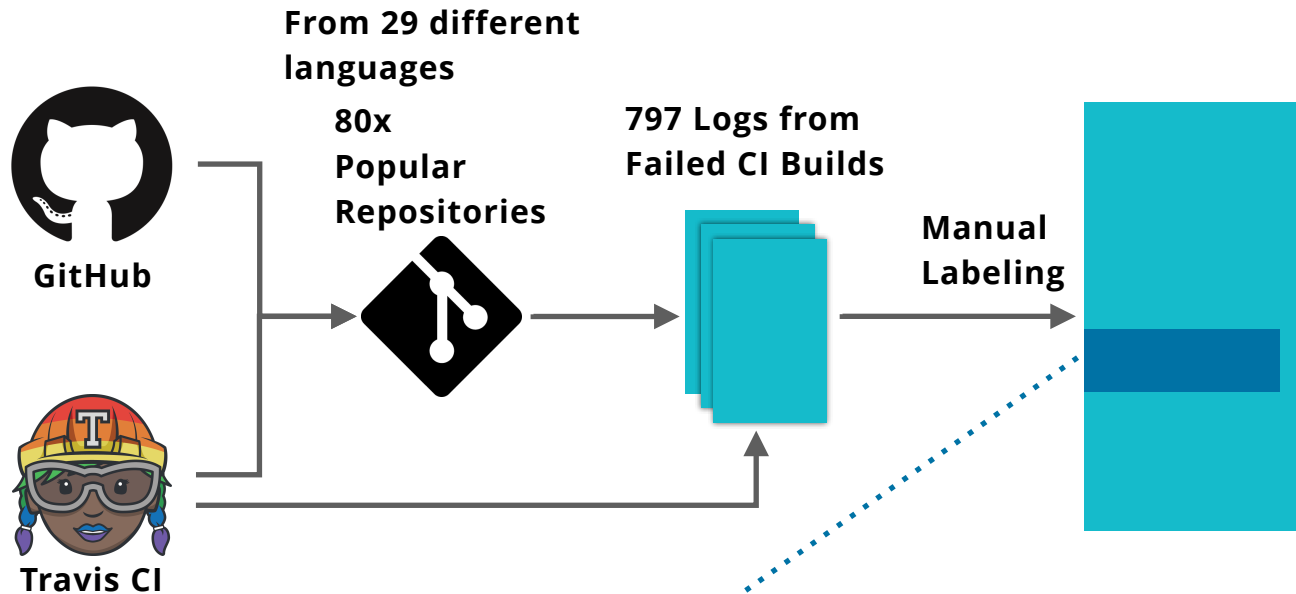
LogChunks Data Set



LogChunks Data Set



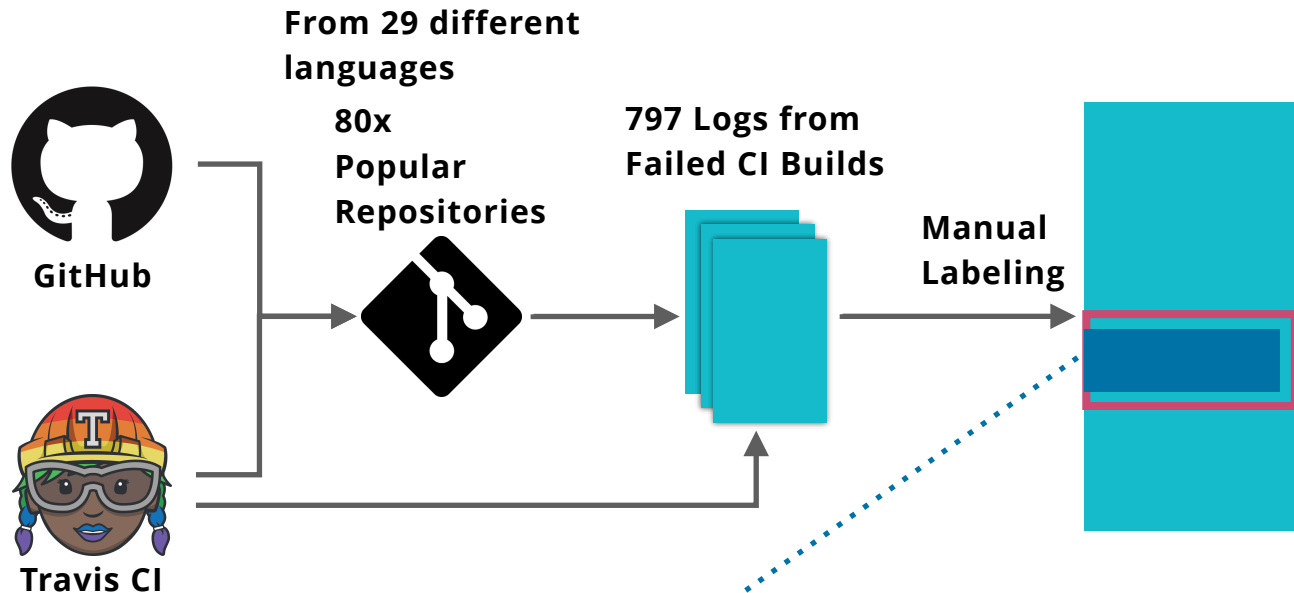
LogChunks Data Set



Log Chunk:
Why the Build Failed

FAIL: TestClientNotificationStorm (2.33s)
client_test.go:378: didn't get expected error

LogChunks Data Set



Log Chunk:
Why the Build Failed

```
FAIL: TestClientNotificationStorm (2.33s)  
client_test.go:378: didn't get expected error
```

| Structural Category | |
|---------------------|--------------|
| Failure: | ===stderr=== |
| xxx | xxx |
| ---- | ===stdout=== |

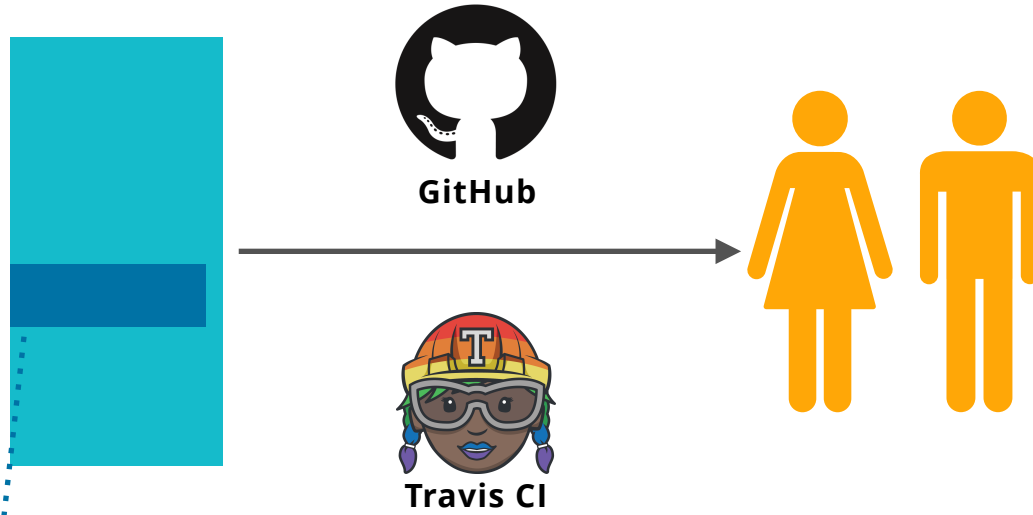
Validation of *LogChunks*



**Log Chunk:
Why the Build Failed**

FAIL: TestClientNotificationStorm (2.33s)
client_test.go:378: didn't get expected error

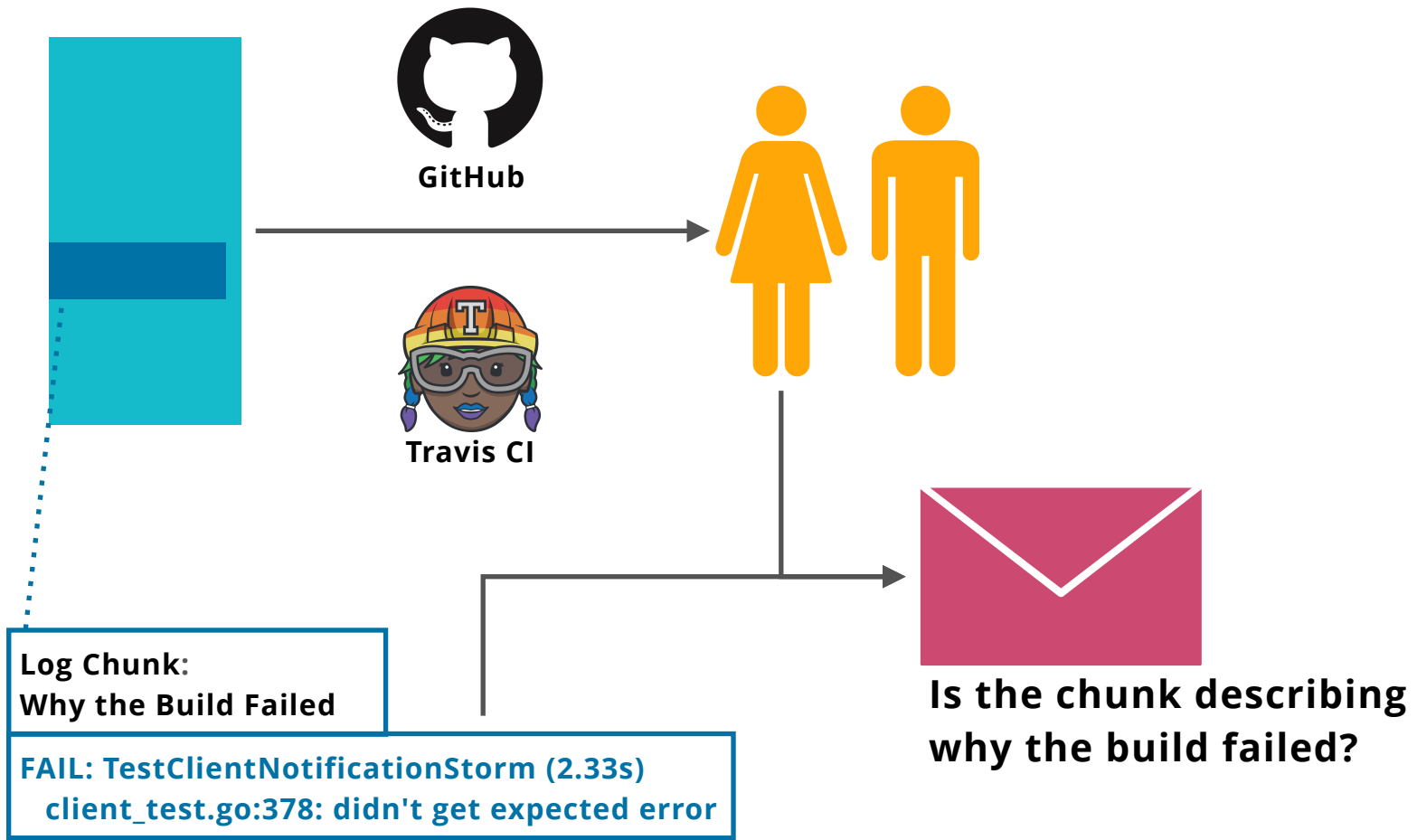
Validation of LogChunks



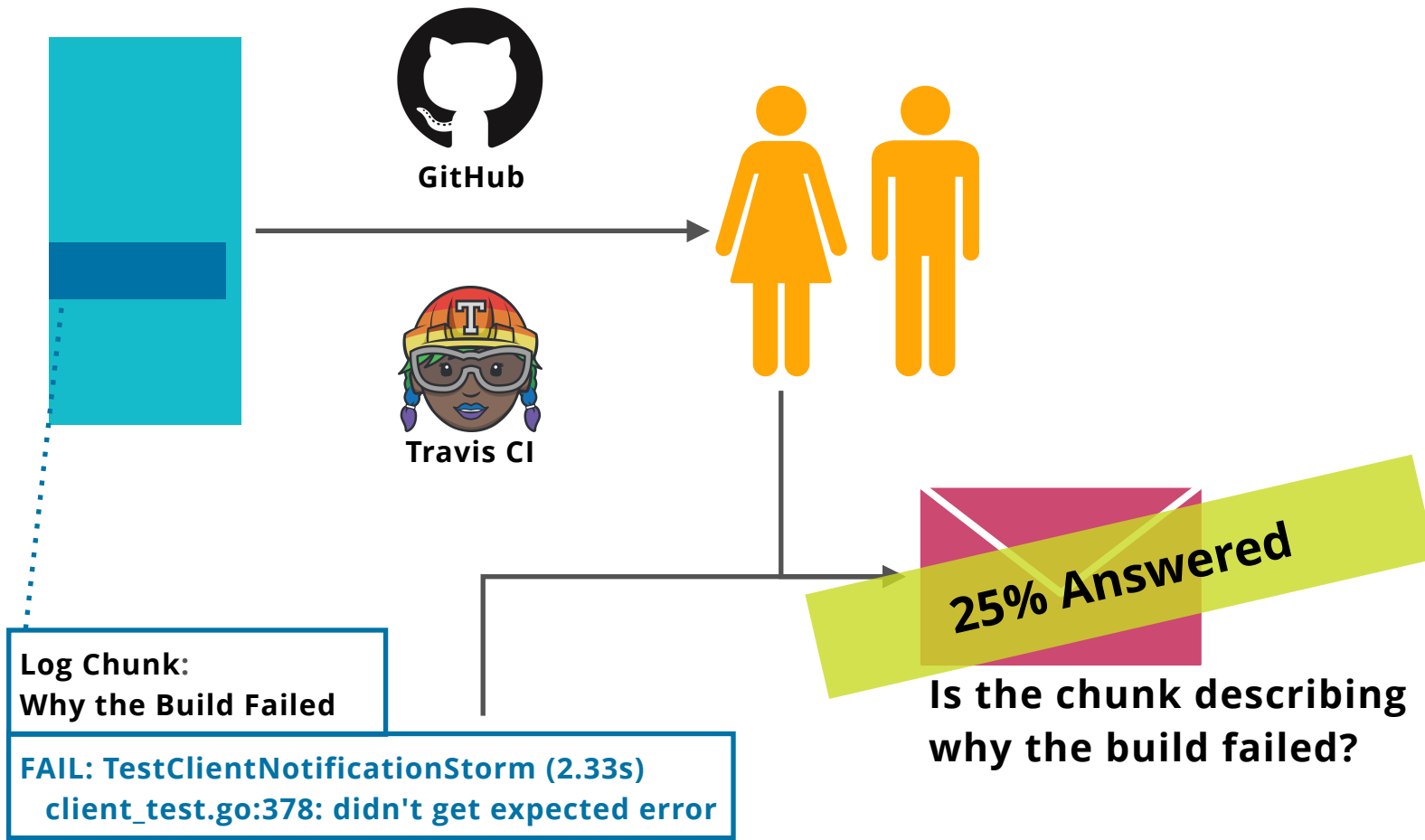
Log Chunk:
Why the Build Failed

FAIL: TestClientNotificationStorm (2.33s)
client_test.go:378: didn't get expected error

Validation of LogChunks



Validation of LogChunks



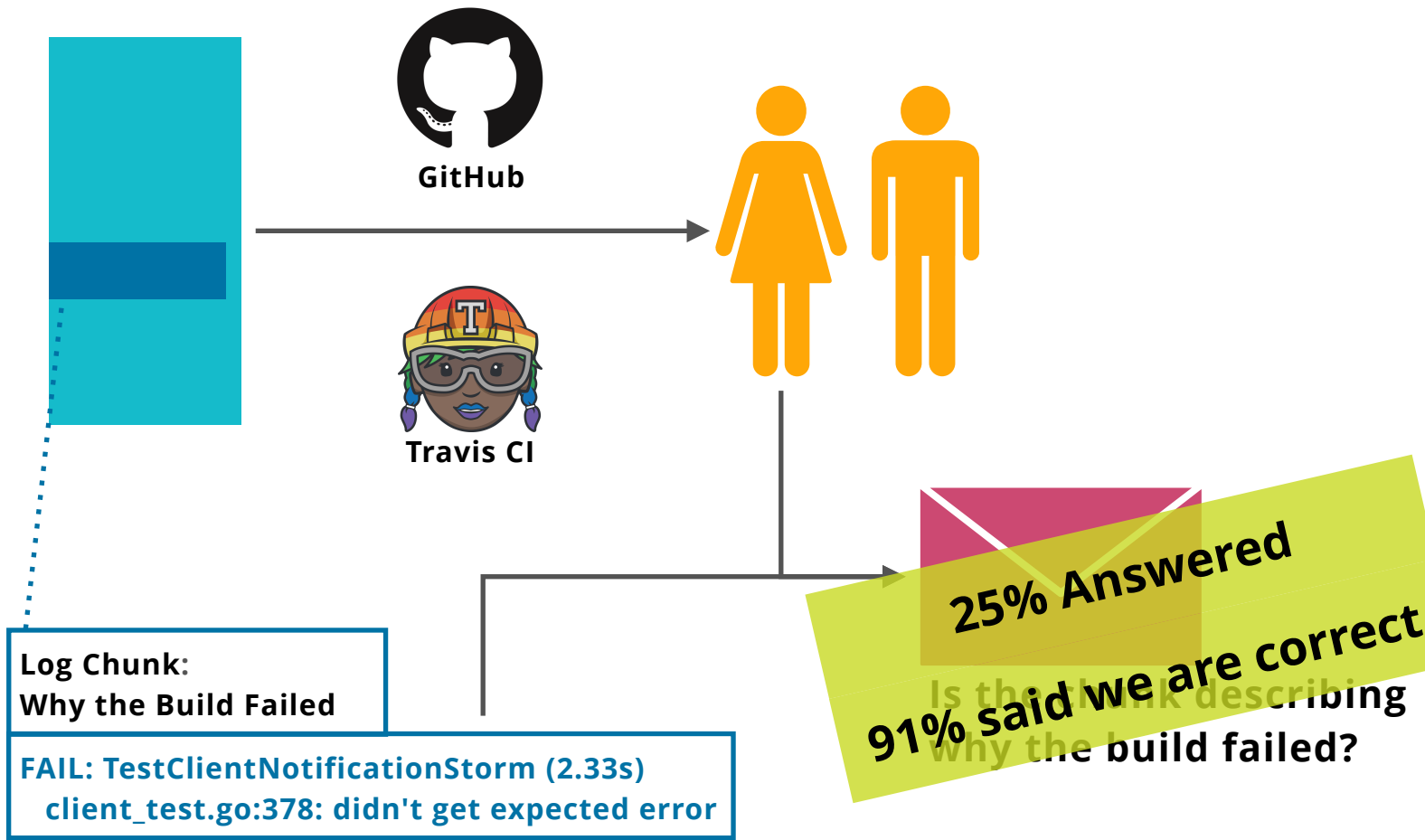
Log Chunk:
Why the Build Failed

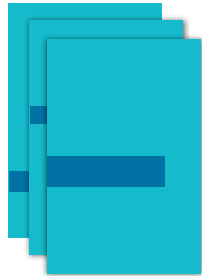
FAIL: TestClientNotificationStorm (2.33s)
client_test.go:378: didn't get expected error

25% Answered

Is the chunk describing why the build failed?

Validation of LogChunks





Create Data Set:
LogChunks



**Implement
Techniques**



**Empirical
Comparison
Study**



**Create Data Set:
*LogChunks***

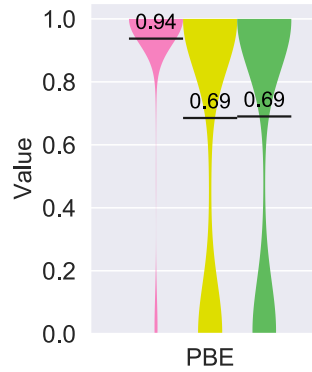


Impact of Structural Categories

One structural category in training & test set:

Failure:
xxx

Failure:
xxx



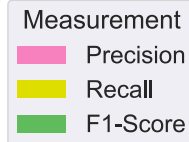
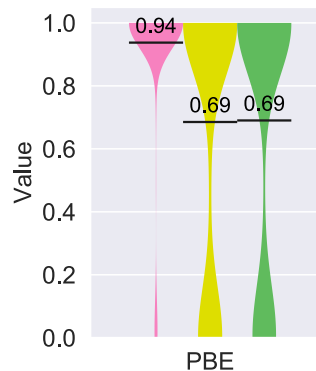
Measurement

- Precision
- Recall
- F1-Score

One structural category in training & test set:

Failure:
xxx

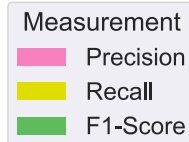
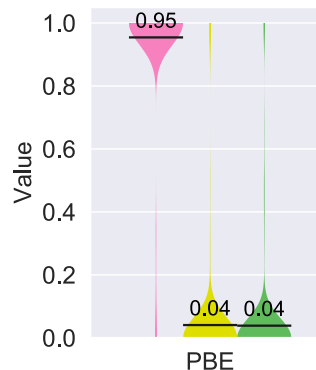
Failure:
xxx



Multiple structural categories in training & test set:

Failure:
xxx

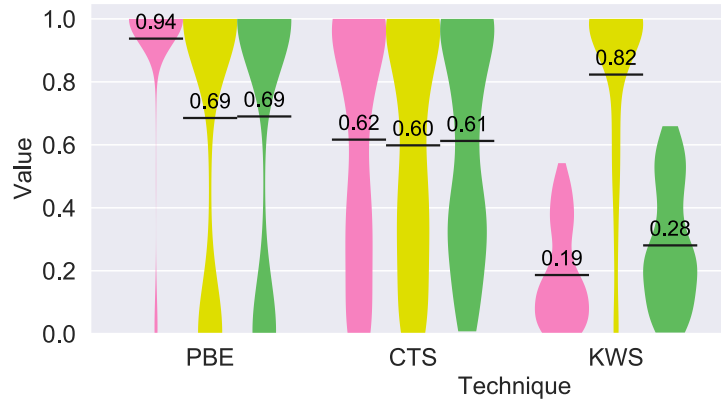
====stderr====
xxx
====stdout====



One structural category in training & test set:

Failure:
XXX

Failure:
XXX



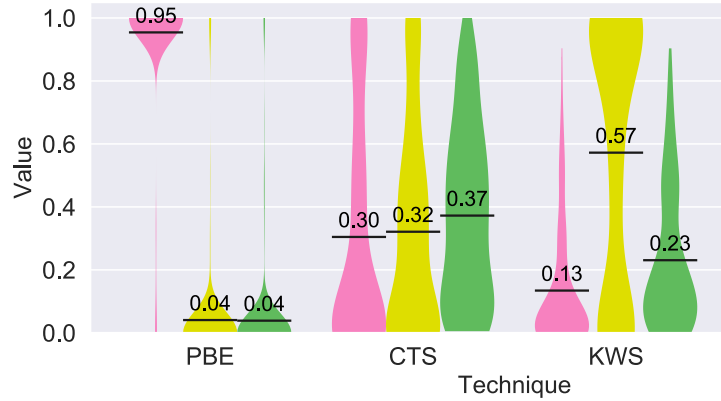
Measurement

- Precision
- Recall
- F1-Score

Multiple structural categories in training & test set:

Failure:
XXX

====stderr====
XXX
====stdout====



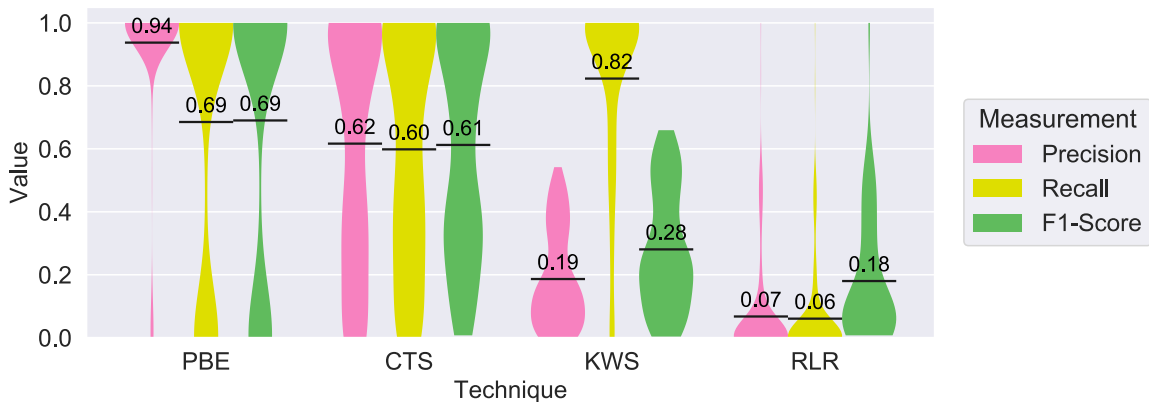
Measurement

- Precision
- Recall
- F1-Score

One structural category in training & test set:

Failure:
XXX

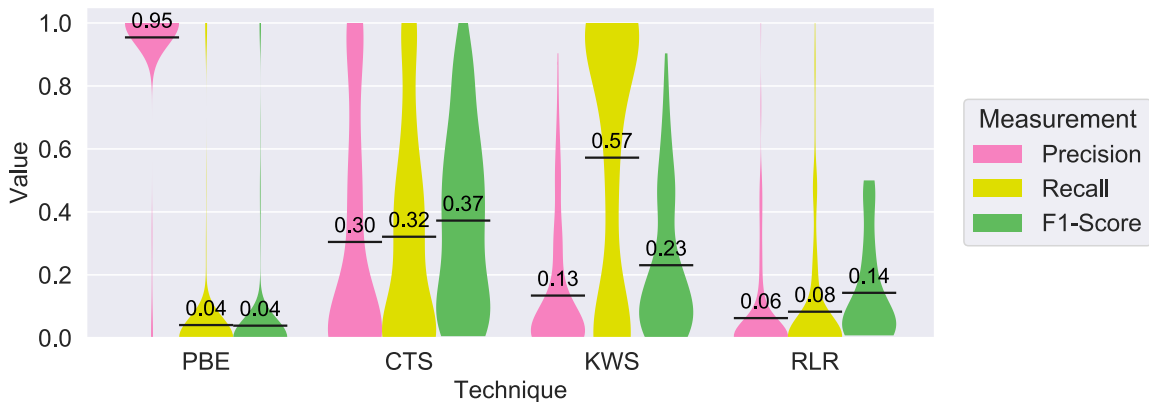
Failure:
XXX



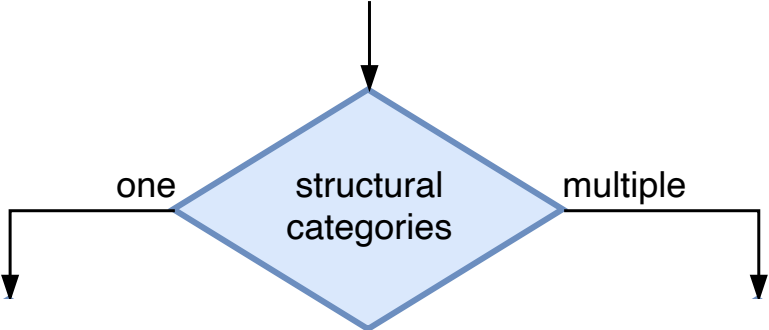
Multiple structural categories in training & test set:

Failure:
XXX

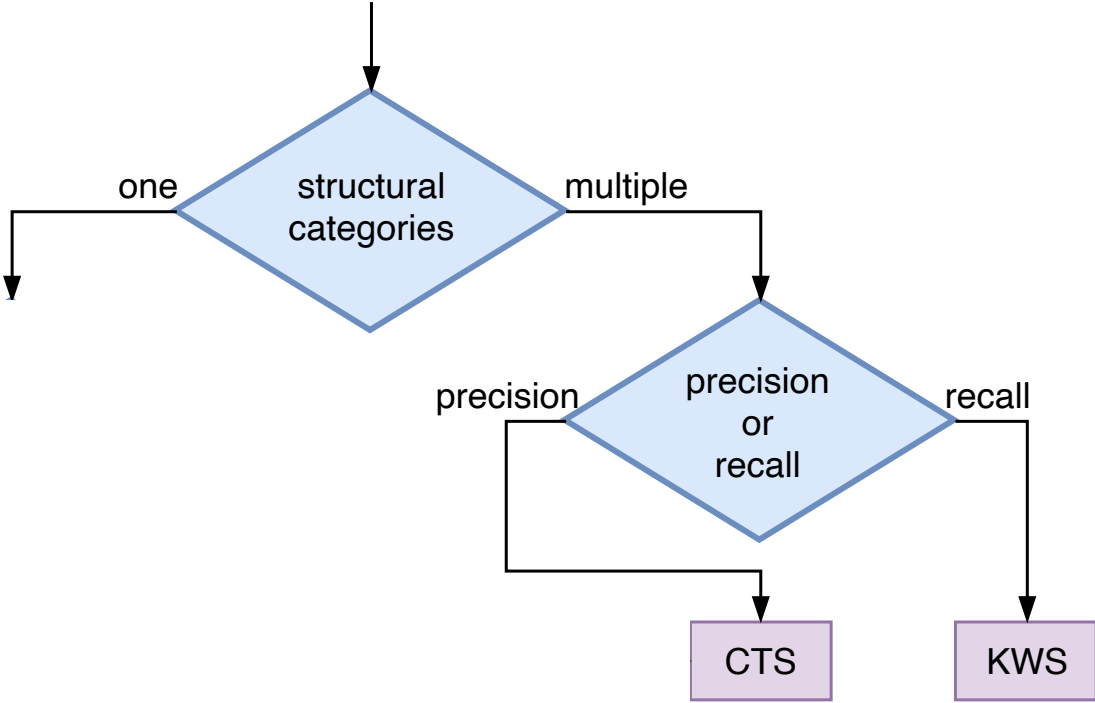
====stderr====
XXX
====stdout====



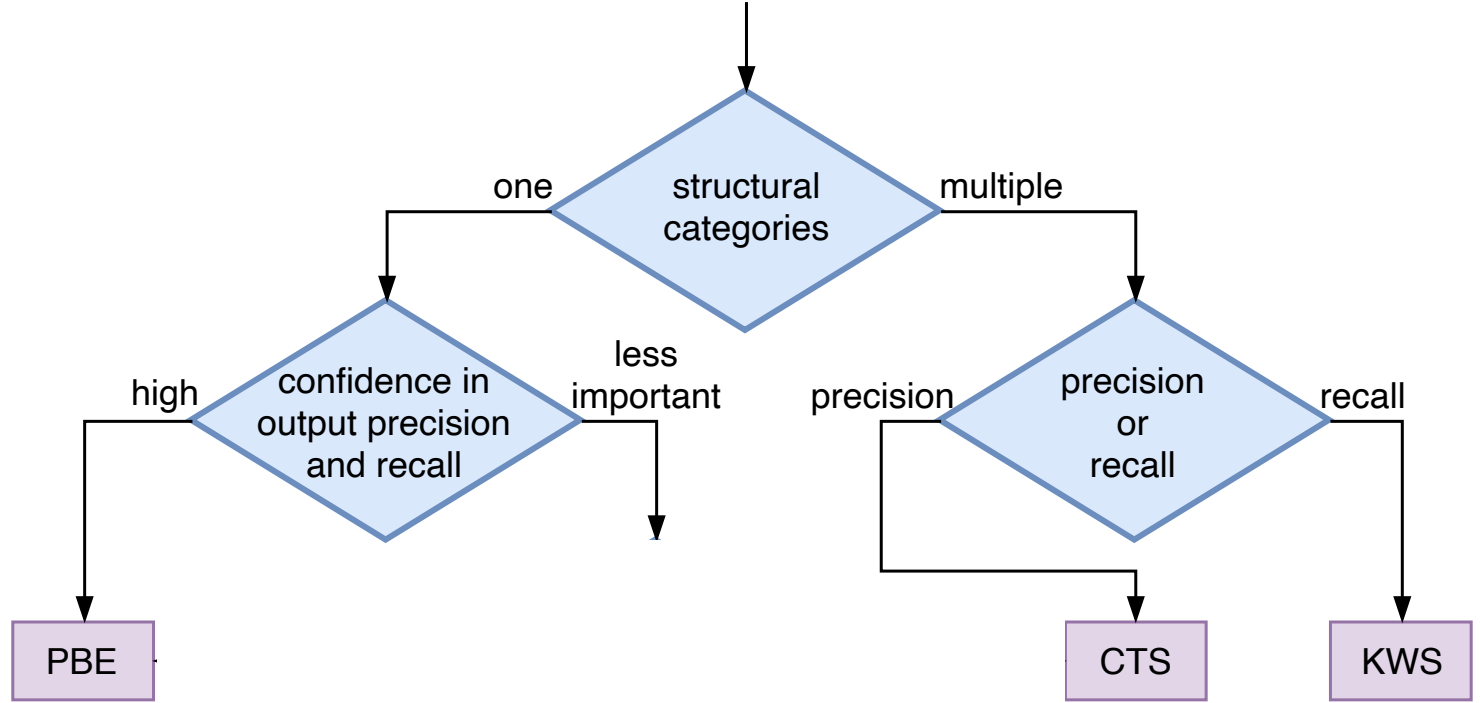
Choosing a Chunk Retrieval Technique



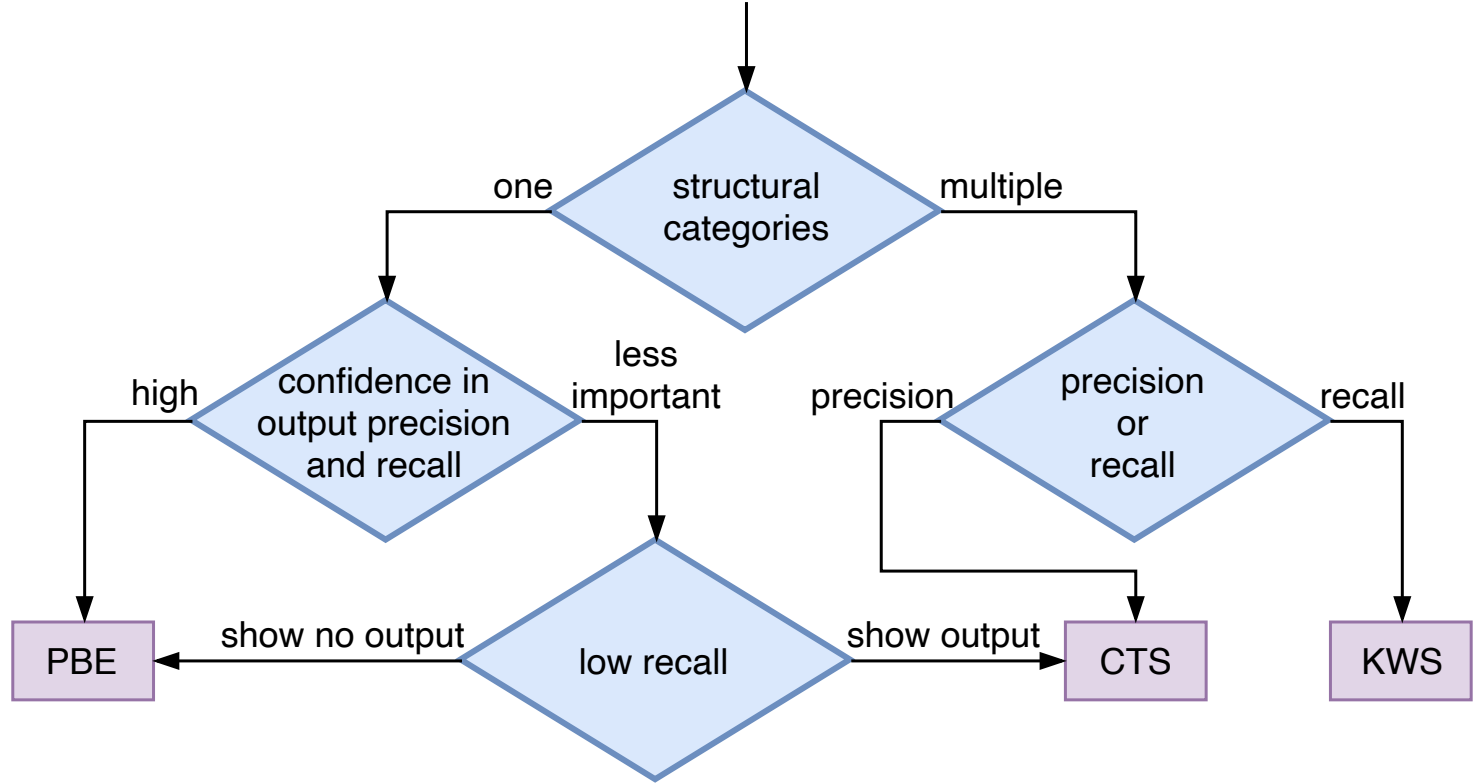
Choosing a Chunk Retrieval Technique



Choosing a Chunk Retrieval Technique




Choosing a Chunk Retrieval Technique



Summary

Build Logs




2

Chunk
= Substring of the log that describes a specific information

```
FAIL: TestSeparator (0.03s)
repo_test.got75: expected entry to be in form of '*
[link] = description', got * [ops](https://github.com/
nanovms/ops) = Unikernel Builder/Orchestrator'
```


Build Logs

2



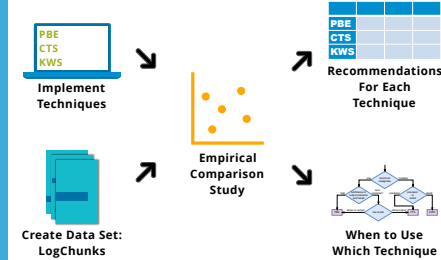
```

FAIL: TestSeparator (0.03s)
repo_test.go:75: expected entry to be in form of '*
[[link = description, got = [ops](https://github.com/nanovms/ops) - Unikernel Builder/Orchestrator'
```

Chunk
= Substring of the log that describes a specific information


Approach

4



Summary

Build Logs



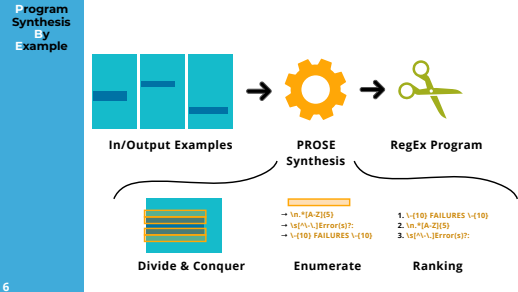
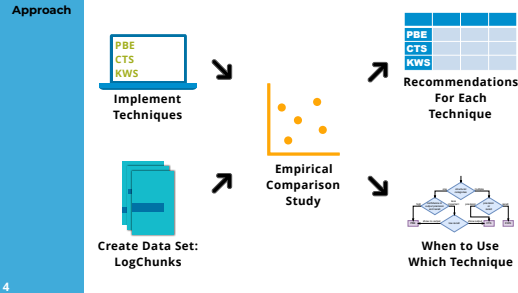
```

    ...
    ...
    ...
  
```

Chunk
= Substring of the log that describes a specific information


```

    FAIL: TestSeparator (0.03s)
    repo_test.go:75: expected entry to be in form of '[link] = description', got '[ops] (https://github.com/nanovms/ops) - Unikernel Builder/Orchestrator'
  
```



Summary

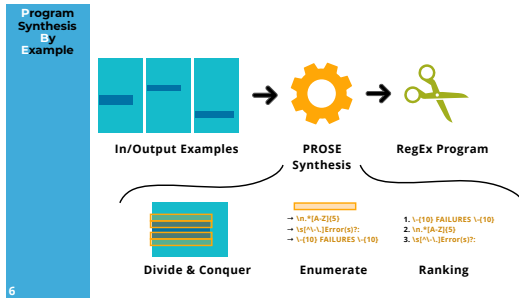
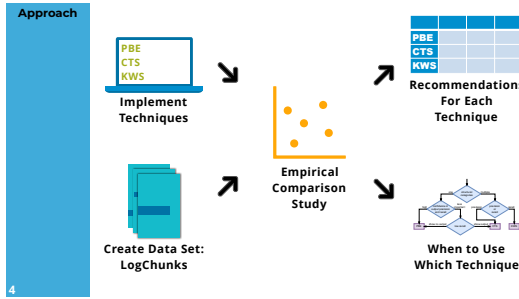
Build Logs



Chunk
= Substring of the log that describes a specific information

```

FAIL: TestSeparator (0.03s)
repo_test.go:75: expected entry to be in form of '*
[link] = description', got '* (ops) (https://github.com/
nanovms/ops) = Unikernel Builder/Orchestrator'
    
```



PBE Example

```

FAILURE: Build failed with an exception.
> What went wrong:
Execution failed for task ':app:connectedAndroidTest'.
> com.android.build.test.runner.DeviceException: java.lang.RuntimeException: ...
> Try:

FAILURE: Build failed with an exception.
> What went wrong:
Could not determine the dependencies of task ':app:jacocoTestDebugReport'.
> Task with path 'testDebug' not found in project ':app'.
> Try:

StartExtraction(RegexPair("Colon-Line Separator", "ALL CAPS")) in
EndExtraction(RegexPair("C", "Line Separator-Line Separator"))

=== RUN TestSeparator
--- FAIL TestSeparator (0.03s)
repo_test.go:75: expected entry to be in form of '* [link] = description', got '* ...
=== RUN TestGeneratorURL

no program found
    
```

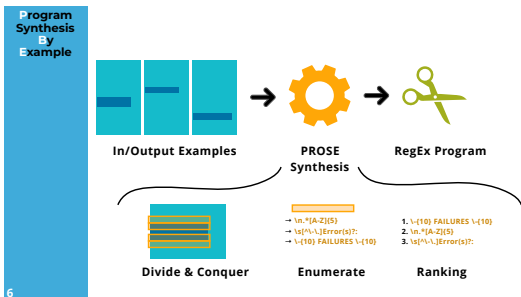
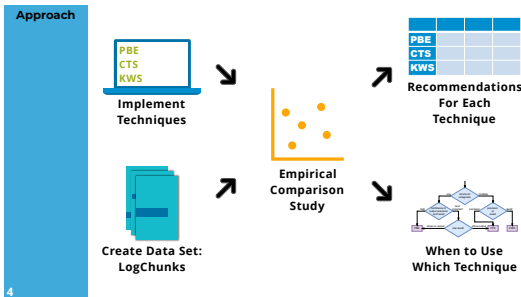
Summary

Build Logs

Chunk
= Substring of the log that describes a specific information

```

FAIL: TestSeparator (0.03s)
repo_test.go:75: expected entry to be in form of '*
[link] = description', got '* (ops) (https://github.com/
nanovms/ops) - Unikernel Builder/Orchestrator'
    
```



PBE Example

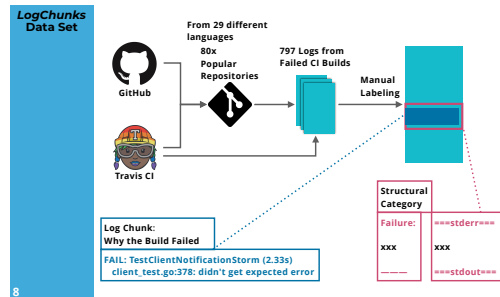
```

FAILURE: Build failed with an exception.
> What went wrong:
Execution failed for task ':app:connectedAndroidTest'.
> com.android.build.api.instrumentation.ApiLang.RuntimeException: ...
> Try:

FAILURE: Build failed with an exception.
> What went wrong:
Could not determine the dependencies of task ':app:jacocoTestDebugReport'.
> Task with path 'testDebug' not found in project ':app'.
> Try:

StartExtraction(RegexPair("Colon-Line Separator", "ALL CAPS")) in
EndExtraction(RegexPair("C", "Line Separator-Line Separator"))

=== RUN TestSeparator
--- FAIL TestSeparator (0.03s)
repo_test.go:75: expected entry to be in form of '* [link] = description', got '* ...
=== RUN TestGeneratorURL
no program found
    
```



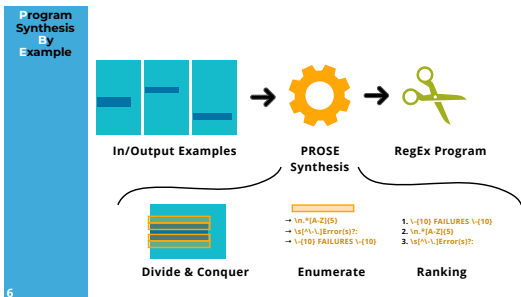
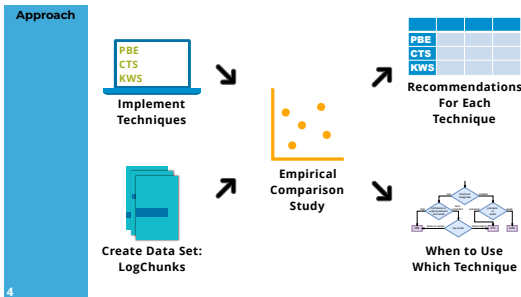
Summary

Build Logs

Chunk
= Substring of the log that describes a specific information

```

FAIL: TestSeparator (0.03s)
repo_test.go:75: expected entry to be in form of '*
[link] = description', got '* (ops) (https://github.com/
nanovms/ops) - Unikernel Builder/Orchestrator'
    
```



PBE Example

```

FAILURE: Build failed with an exception.
> What went wrong:
Execution failed for Task ':app:connectedAndroidTest'.
> java.lang.RuntimeException: java.lang.RuntimeException: ...
> Try:
    
```

```

FAILURE: Build failed with an exception.
> What went wrong:
Could not determine the dependencies of task ':app:jacocoTestDebugReport'.
> Task with path 'testDebug' not found in project ':app'.
> Try:
    
```

```

StartExtraction(RegexPair("Colon-Line Separator", "ALL CAPS")) in
EndExtraction(RegexPair("C", "Line Separator-Line Separator"))

--- RUN TestSeparator
--- FAIL: TestSeparator (0.03s)
repo_test.go:75: expected entry to be in form of '* [link] = description', got '* ...
--- RUN TestGeneratorPHL

no program found
    
```

