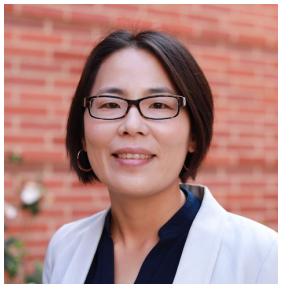


Analyzing and Supporting Adaptations of Online Code Examples

Tianyi Zhang,[1] Di Yang,[2] Crista Lopes,[2] Miryung Kim[1]

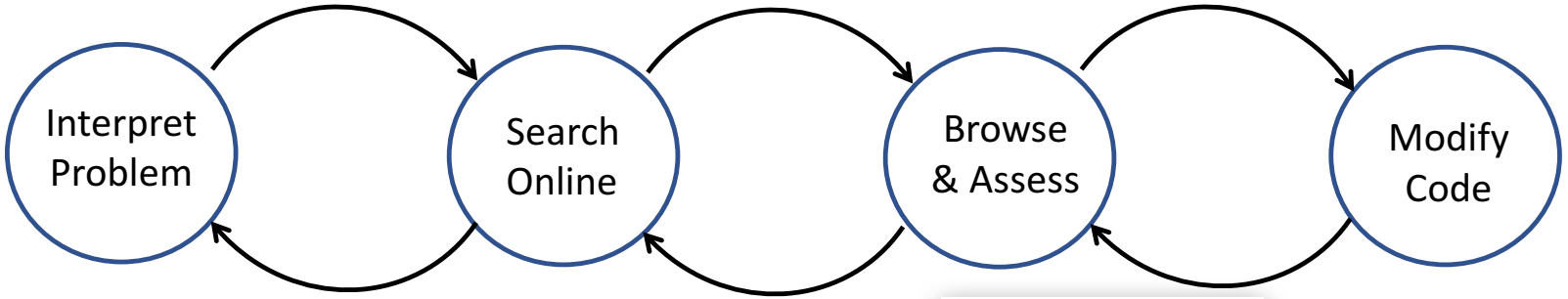
[1]University of California, Los Angeles

[2]University of California, Irvine

Dataset and Tool: https://github.com/tianyi-zhang/ExampleStack-ICSE-Artifact

* Both the first author and the second author contributed significantly.

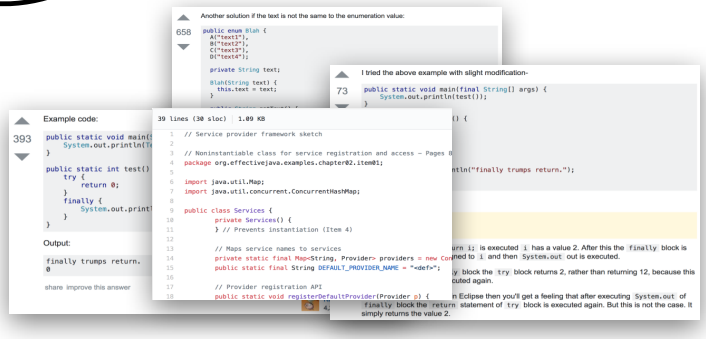
Modern Programming Workflow



"how to connect to MySQL in Java?"

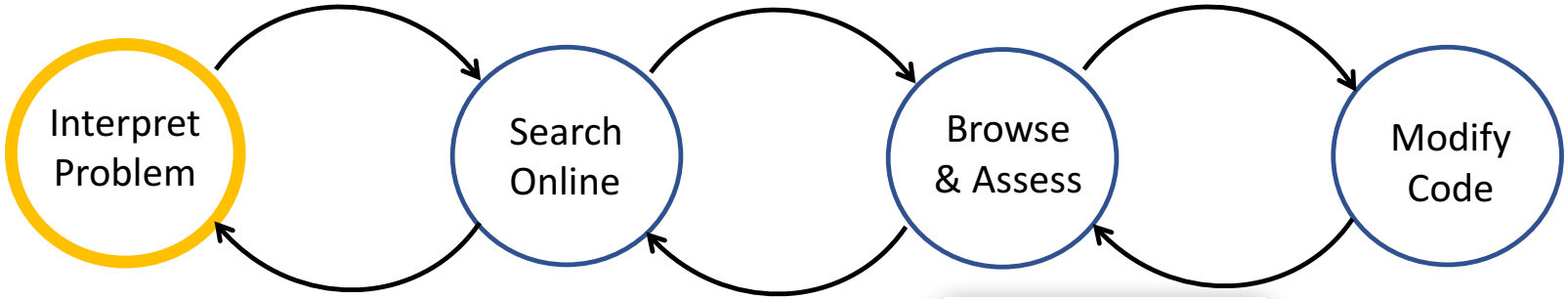
stackoverflow

github



Brandt *et al.* Two studies of opportunistic programming: interleaving web foraging, learning, and writing code. 2009.

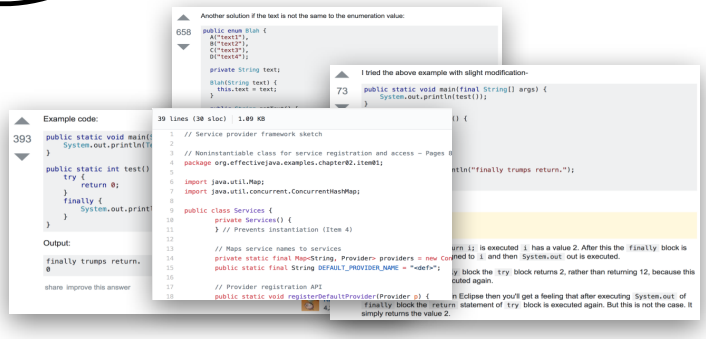
Modern Programming Workflow



"how to connect to MySQL in Java?"

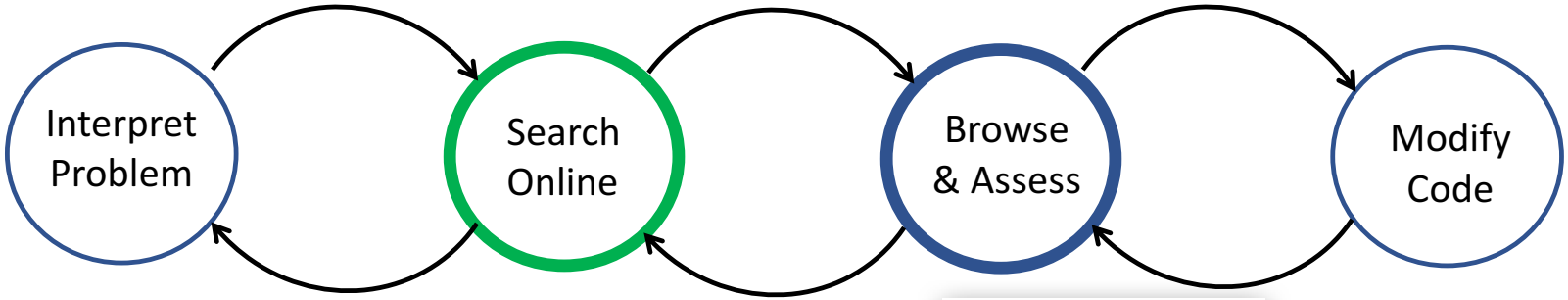
stackoverflow

github



Brandt *et al.* Two studies of opportunistic programming: interleaving web foraging, learning, and writing code. 2009.

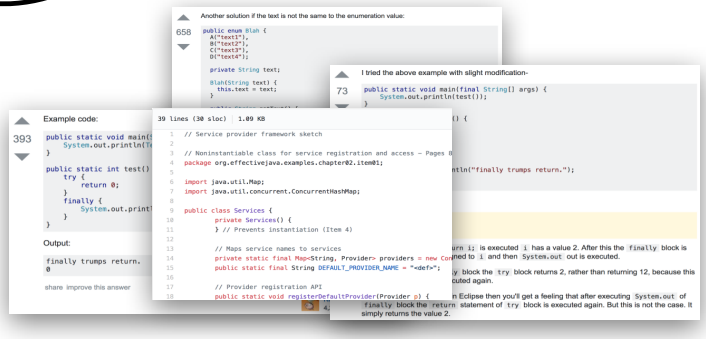
Modern Programming Workflow



"how to connect to MySQL in Java?"

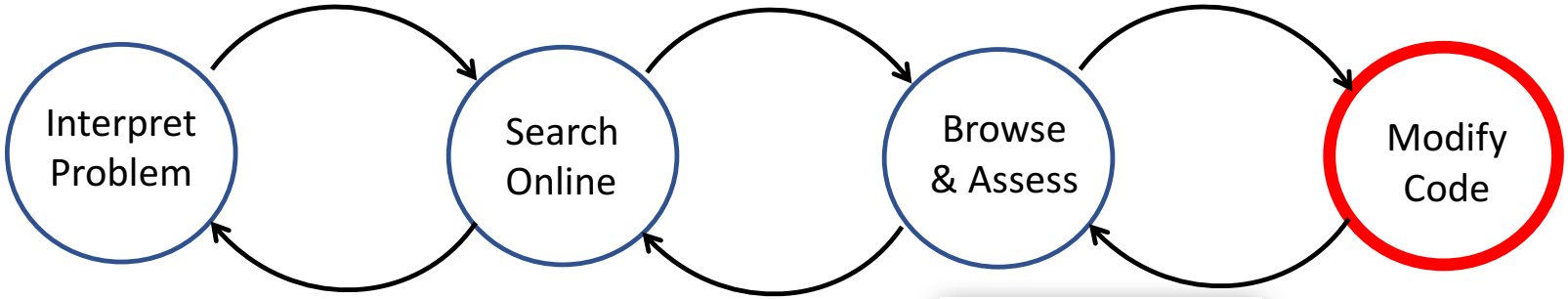
stackoverflow

github



Brandt *et al.* Two studies of opportunistic programming: interleaving web foraging, learning, and writing code. 2009.

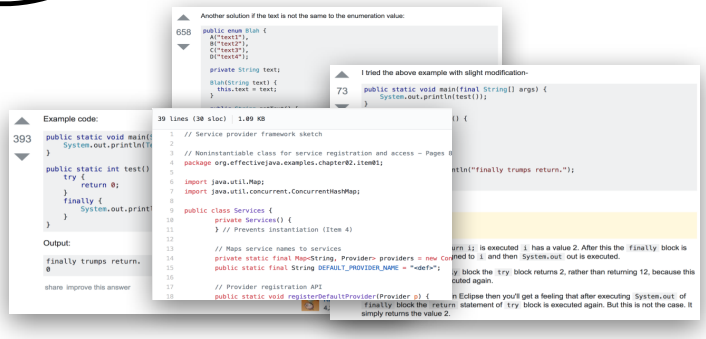
Modern Programming Workflow



"how to connect to MySQL in Java?"

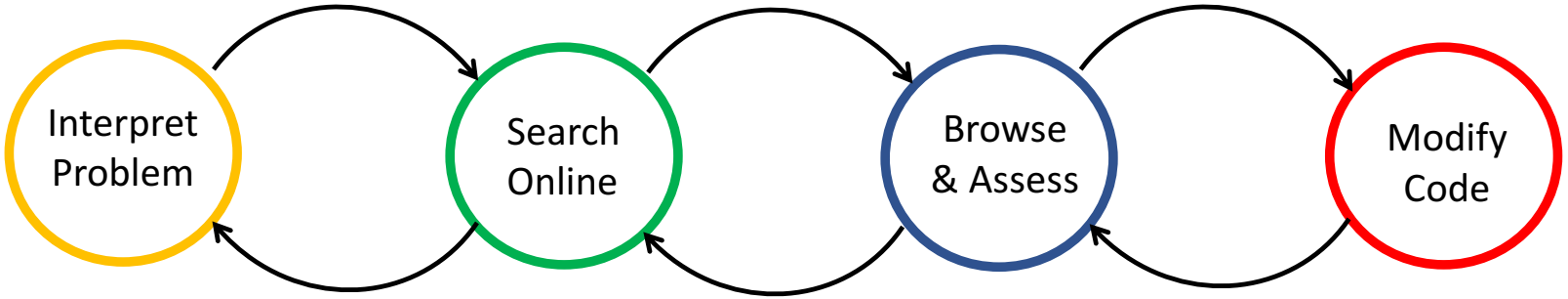
stackoverflow

github



Brandt *et al.* Two studies of opportunistic programming: interleaving web foraging, learning, and writing code. 2009.

Modern Programming Workflow



Test cases [CodeGenie]
I/O types [ParseWeb, Hunter]
I/O examples [Stolee et al., 2014, FlashFill]
Multimodal [Reiss, 2009]
…

StrathCona [Homles and Murphy, 2005]
Sourcerer [Bajracharya *et al.*, 2006]
Exemplar [McMillan *et al.*, 2012]
FaCoy [Kim *et al.*, 2018]
…

Prompter [Ponzanelli *et al.*, 2014]
AnswerBot [Xu *et al.*, 2017]
Deprecation Watchter [Zhou *et al.*, 2017]
ExampleCheck [Zhang *et al.*, 2018]
Examplore [Glassman *et al.*, 2018]
…

?

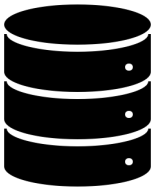
What we have known so far …

- Online Code Reuse Behavior
  - Copy and paste with adaptations [Wu *et al.*, 2018]
  - Seldom attribute to the sources of online code [Baltes and Diehl, 2018]
- Code Adaptation & Integration Support
  - Rename variables and port relevant program statements [SnipMatch, Jigsaw]

What we don’t know yet…

- RQ1. What kinds of adaptations do developers make in practice?
- RQ2. Are these adaptations done repetitively?
- RQ3. How can we provide effective tool support?

Outline



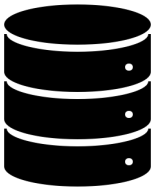
1. A Comprehensive Dataset

2. Qualitative Analysis

3. Quantitative Analysis

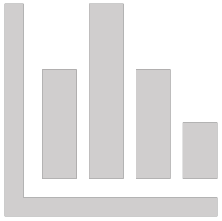
4. Tool Design & User Study

Outline



1. A Comprehensive Dataset

2. Qualitative Analysis

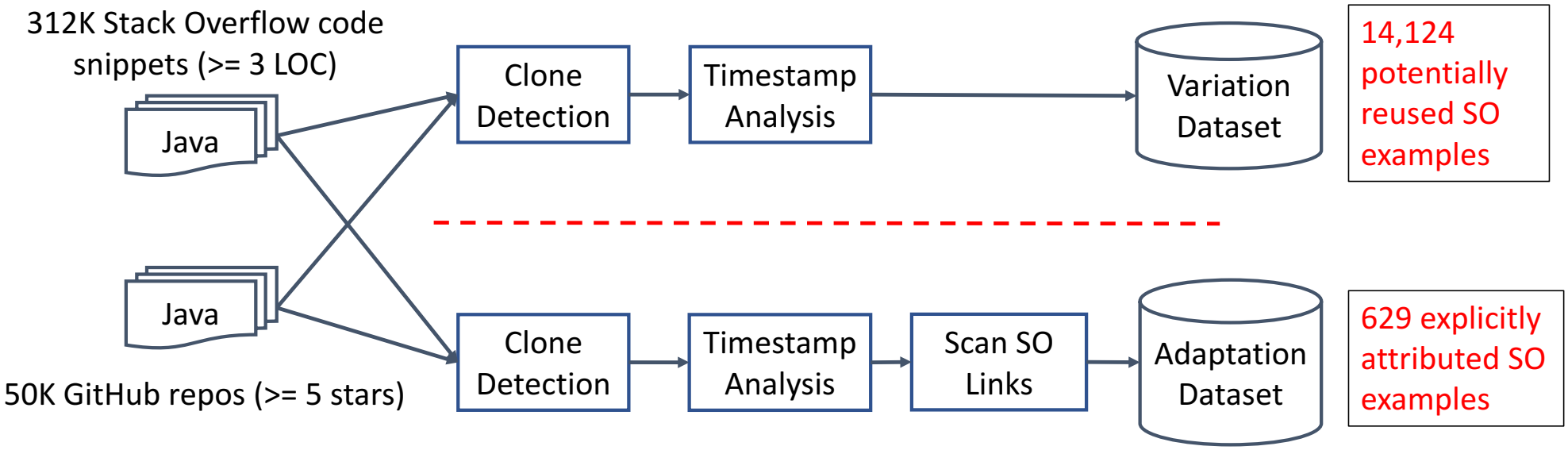


3. Quantitative Analysis

4. Tool Design & User Study

Identify Reused Stack Overflow Examples

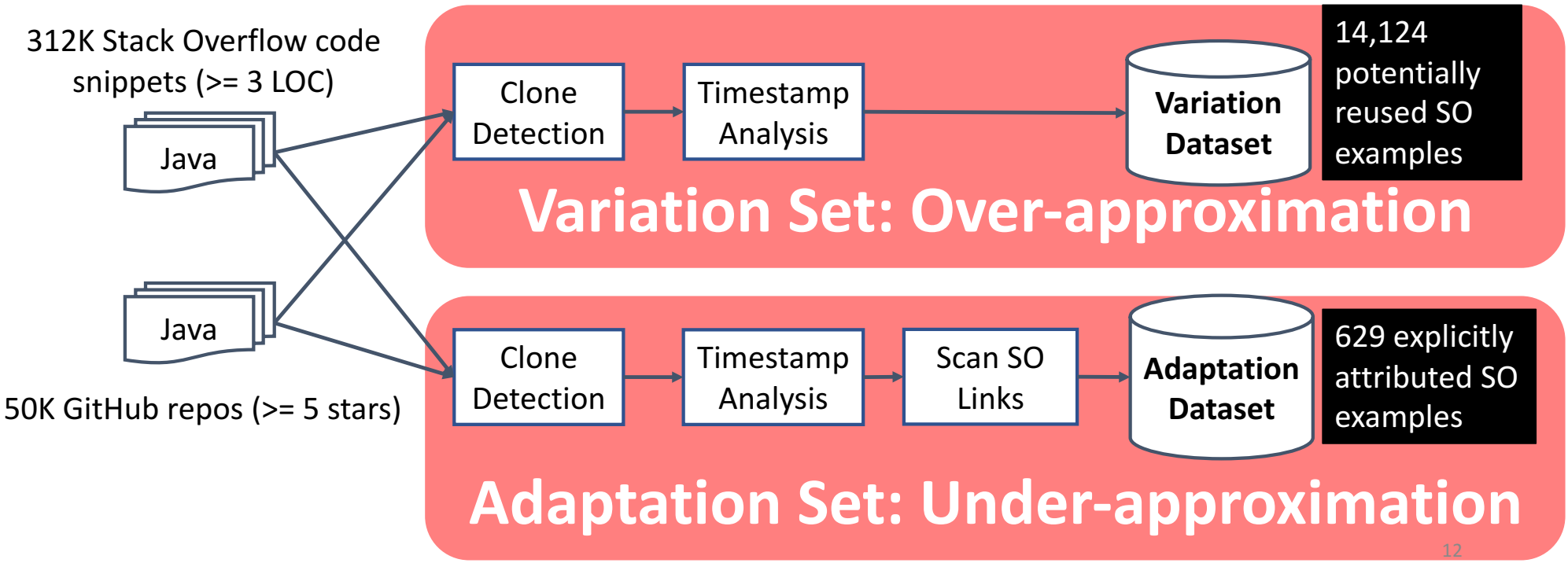
- **Challenge**: the lack of attribution [Baltes and Diehl, 2018]



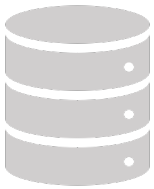
Sajnani et al. SourcererCC: Scaling Code Clone Detection to Big Code. 2016

Identify Reused Stack Overflow Examples

- **Challenge**: the lack of attribution [Baltes and Diehl, 2018]



Outline



1. A Comprehensive Dataset

2. Qualitative Analysis

3. Quantitative Analysis

4. Tool Design & User Study

Qualitative Analysis

- Randomly sample 200 pairs of clones from each dataset
- Manually inspect their differences using GumTree [Falleri *et al.*, 2014]
- Label program changes with short descriptions and group similar ones.

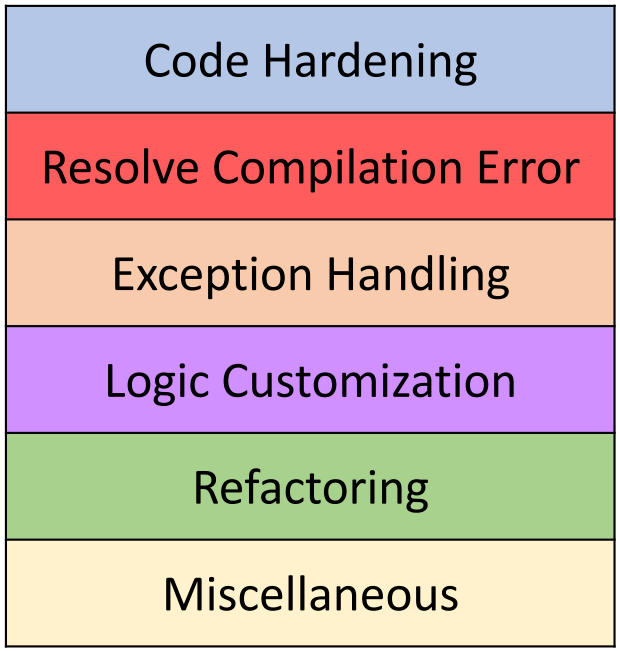
so-37273871-0-1.java

```
public class foo {
public String getJSONFromAssets() {
        String json = null;
        try {
            InputStream inputData = getAssets().open("locations.json");
            int size = inputData.available();
            byte[] buffer = new byte[size];
            inputData.read(buffer);
            inputData.close();
            json = new String(buffer, "UTF-8");
        } catch (IOException ex) {
            ex.printStackTrace();
            return null;
        }
        return json;
```

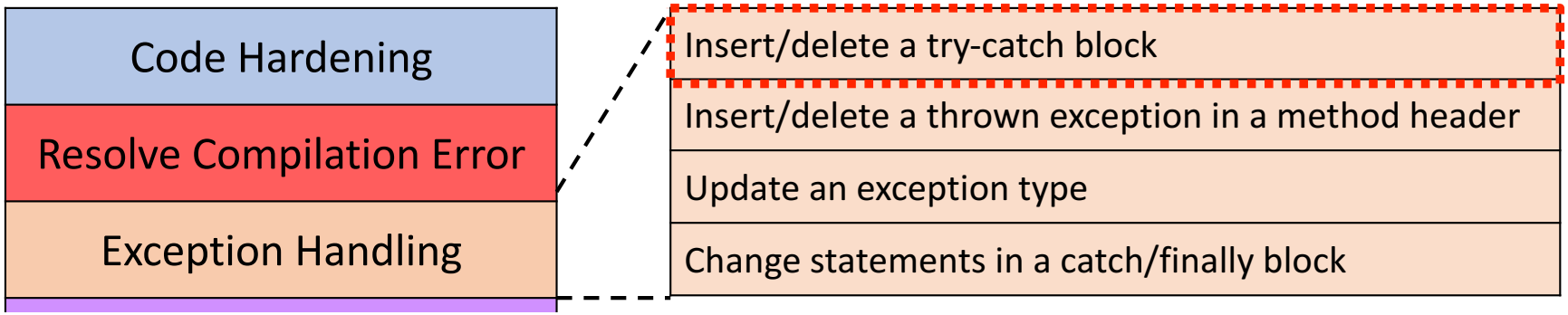
carved-gh-0-1-517-531.java

```
public class foo{
    public String loadJSONFromAsset(String jsonFileName) {
        String json = null;
        try {
            InputStream is = getAssets().open(jsonFileName);
            int size = is.available();
            byte[] buffer = new byte[size];
            is.read(buffer);
            is.close();
            json = new String(buffer, "UTF-8");
        } catch (IOException ex) {
            ex.printStackTrace();
            return null;
        }
        return json;
```

24 Frequent Adaptation Types in 6 Categories



24 Frequent Adaptation Types in 6 Categories



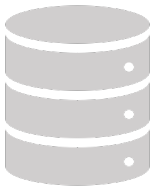
so-15409446-1-2.java

```
public class foo {
public static void addLibraryPath(String pathToAdd) throws Exception{
    final Field usrPathsField = ClassLoader.class.getDeclaredField("usr_paths");
    usrPathsField.setAccessible(true);

    //add the new path
    final String[] newPaths = Arrays.copyOf(paths, paths.length + 1);
    newPaths[newPaths.length-1] = pathToAdd;
    usrPathsField.set(null, newPaths);
}
}
```

carved-gh-3-1-131-153.java

```
public class foo{
        private static void addLibraryPath(String pathToAdd) {
                try {
                        final Field usrPathsField = ClassLoader.class.getDeclaredField(
                        usrPathsField.setAccessible(true);

                        // add the new path
                        final String[] newPaths = Arrays.copyOf(paths, paths.length + 1
                        newPaths[newPaths.length - 1] = pathToAdd;
                        usrPathsField.set(null, newPaths);
                } catch (NoSuchFieldException | SecurityException | IllegalArgumentExce
                        throw new RuntimeException(e);
                }
```

Outline



1. A Comprehensive Dataset

2. Qualitative Analysis

3. Quantitative Analysis

4. Tool Design & User Study

Automated Rule-based Classification

- Codify each adaptation type as a logic rule
  - e.g., **Insert**(t_1, t_2, i) $\wedge$ **NodeType**(t_1, `TryStatement`) $\Rightarrow$**Insert_Try_Catch_Block**
- 98% precision and 96% recall on another 100 clone pairs

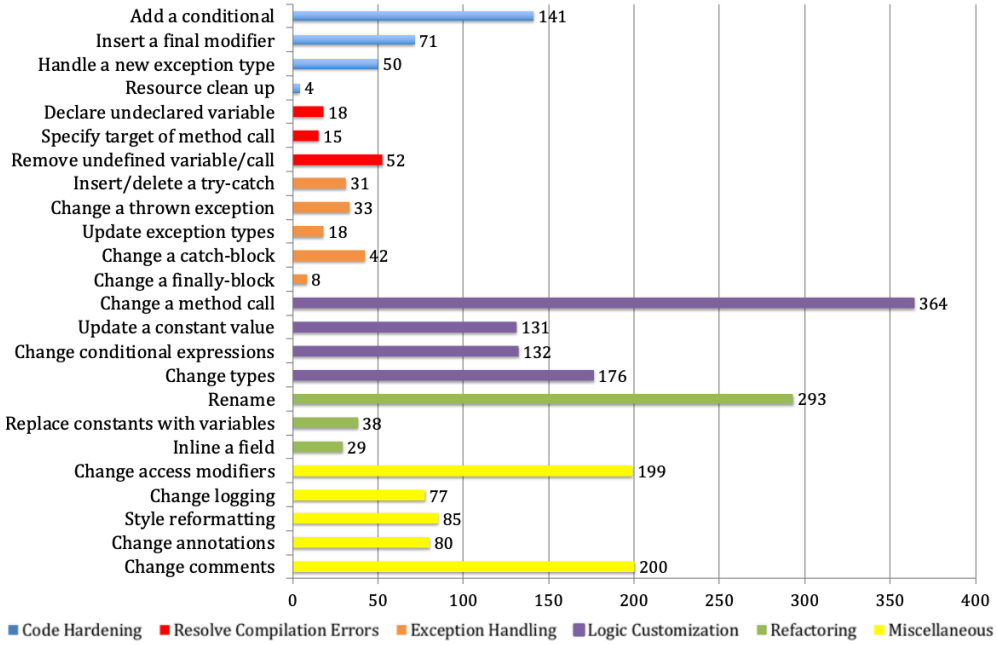
so-15409446-1-2.java

```
public class foo {
public static void addLibraryPath(String pathToAdd) throws Exception{
    final Field usrPathsField = ClassLoader.class.getDeclaredField("usr_paths");
    usrPathsField.setAccessible(true);

    //add the new path
    final String[] newPaths = Arrays.copyOf(paths, paths.length + 1);
    newPaths[newPaths.length-1] = pathToAdd;
    usrPathsField.set(null, newPaths);
}
}
```

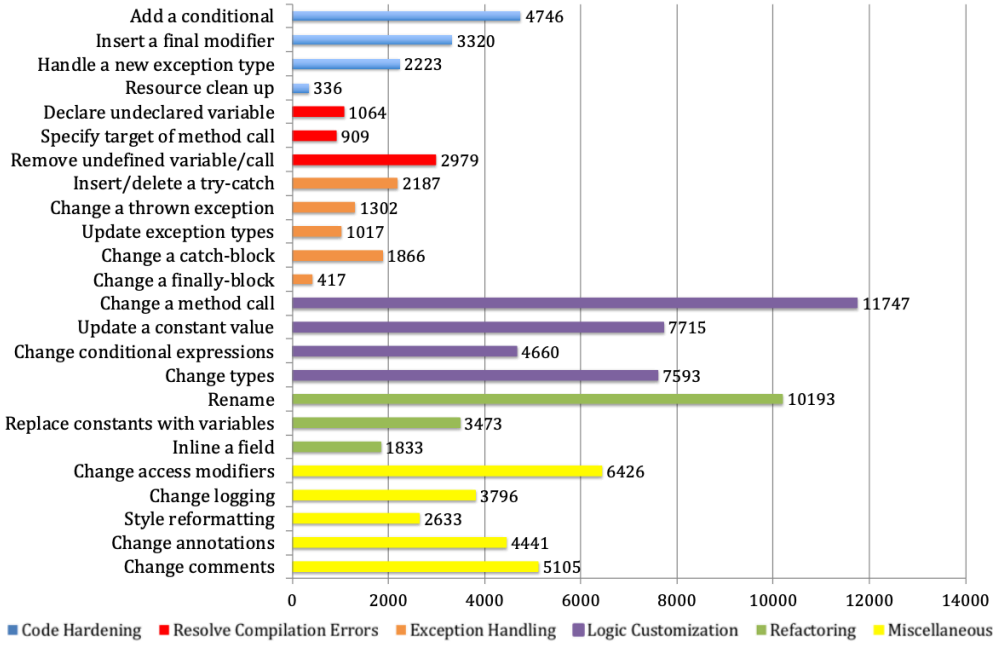
carved-gh-3-1-131-153.java

```
public class foo{
        private static void addLibraryPath(String pathToAdd) {
                try {
                        final Field usrPathsField = ClassLoader.class.getDeclaredField(
                        usrPathsField.setAccessible(true);

                        // add the new path
                        final String[] newPaths = Arrays.copyOf(paths, paths.length + 1
                        newPaths[newPaths.length - 1] = pathToAdd;
                        usrPathsField.set(null, newPaths);
                } catch (NoSuchFieldException | SecurityException | IllegalArgumentExce
                        throw new RuntimeException(e);
                }
        }
```

Distribution of Common Adaptation Types

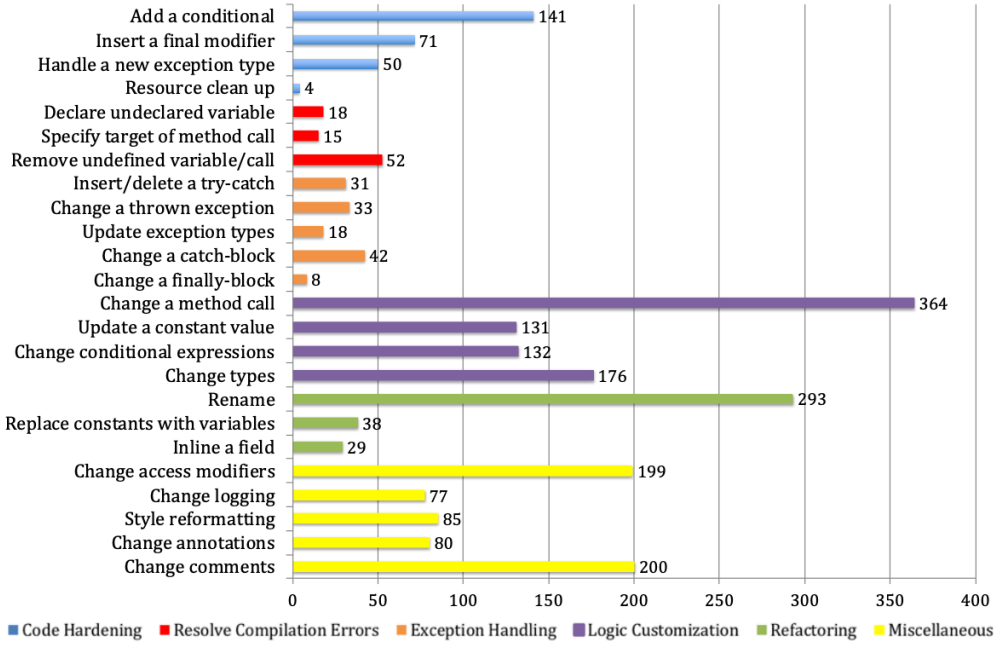


(a) Adaptations: 629 explicitly attributed SO examples

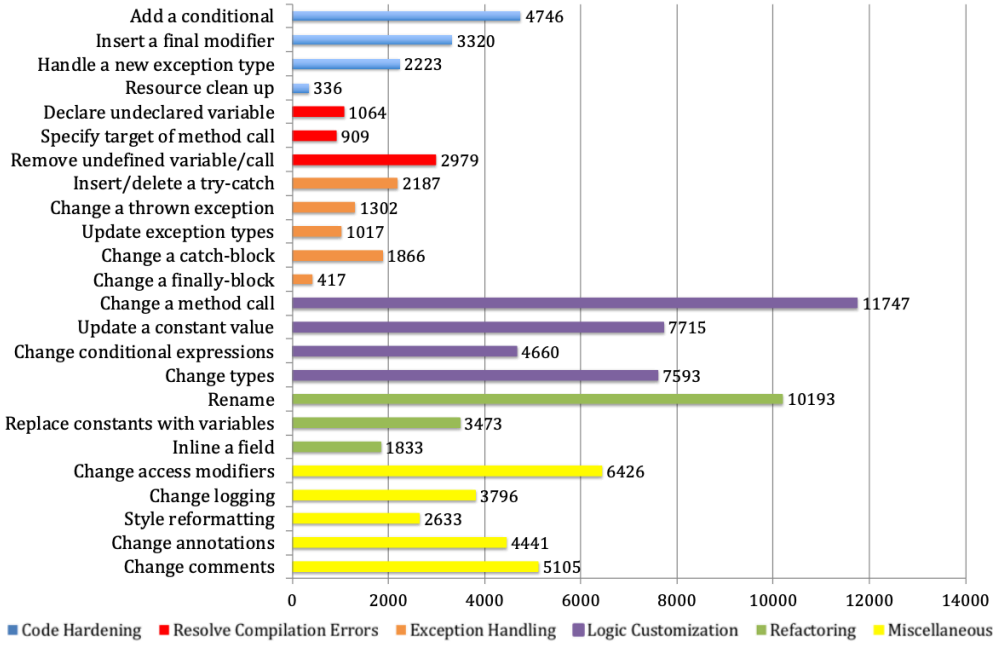


(b) Variations: 14,124 potentially reused SO examples

Finding 1. Variation patterns resemble adaptation patterns

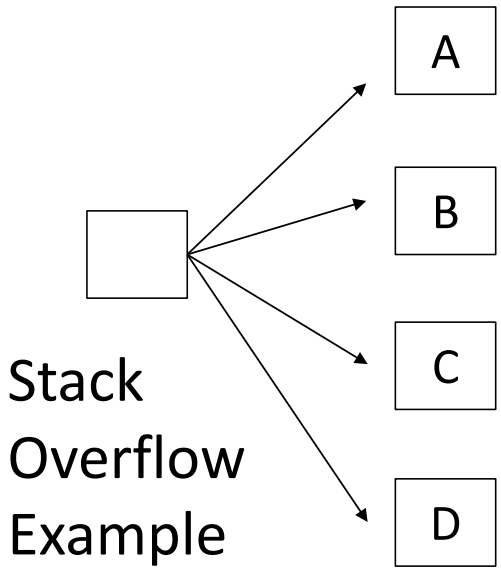


(a) Adaptations: 629 explicitly attributed SO examples



(b) Variations: 14,124 potentially reused SO examples

Finding 2. Different GitHub clones of the same example share common adaptation types.



Stack Overflow Example

GitHub Counterparts

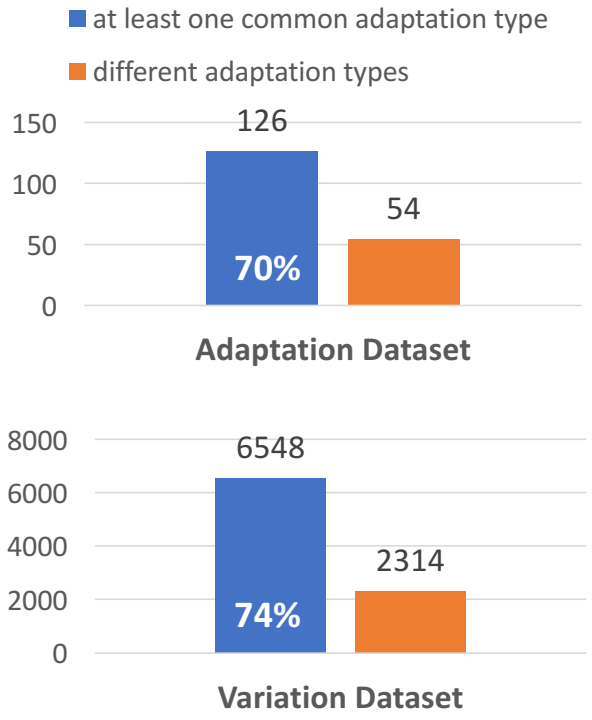
Add an if check, renaming

Add an if check, change a method call

Add an if check, renaming

Change a method call, renaming

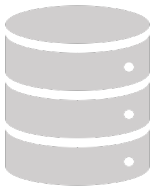
Adaptations



Implications and Hypothesis Development

- **Implications**
  - Variations in similar code resemble real adaptations made by developers
  - Different GitHub developers make similar adaptations independently
- **Hypothesis:** Displaying variations in similar GitHub code can inspire more careful reasoning when adapting code

Outline



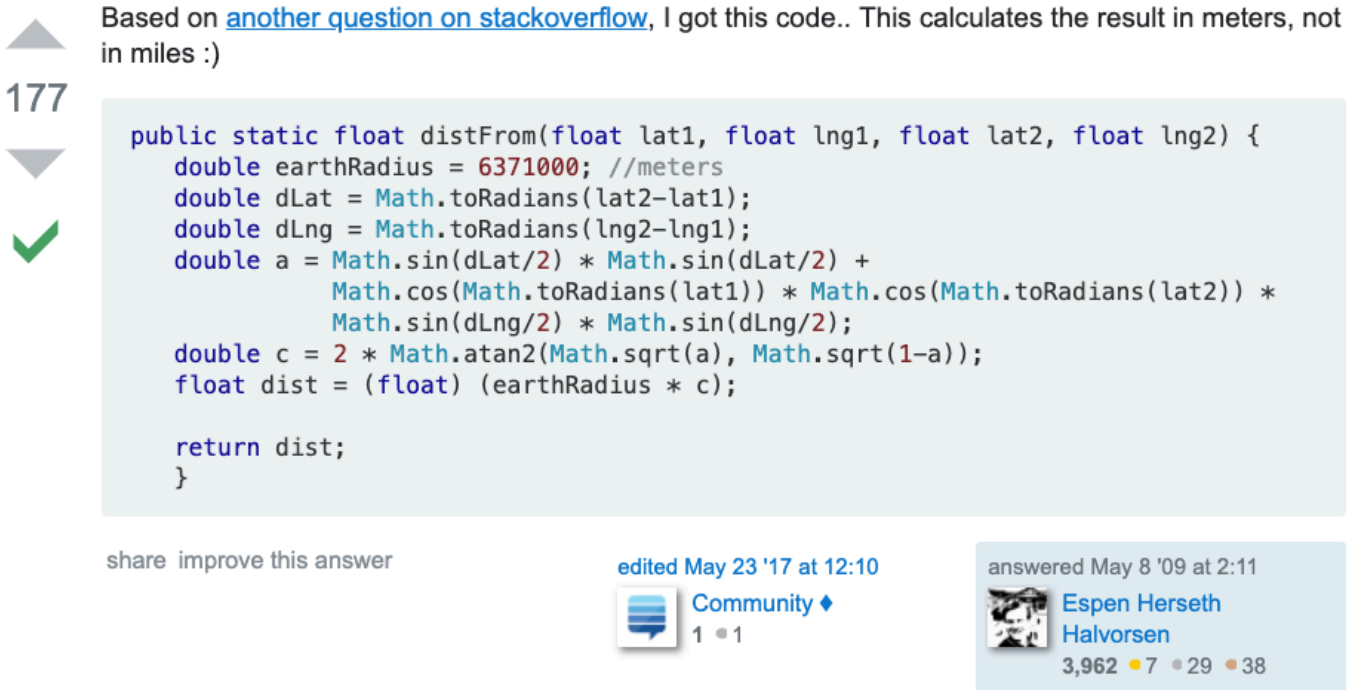
1. A Comprehensive Dataset

2. Qualitative Analysis

3. Quantitative Analysis

4. Tool Design & User Study

“How to calculate the distance between two coordinates?”



177

Based on another question on stackoverflow, I got this code.. This calculates the result in meters, not in miles :)

```
public static float distFrom(float lat1, float lng1, float lat2, float lng2) {
   double earthRadius = 6371000; //meters
   double dLat = Math.toRadians(lat2-lat1);
   double dLng = Math.toRadians(lng2-lng1);
   double a = Math.sin(dLat/2) * Math.sin(dLat/2) +
              Math.cos(Math.toRadians(lat1)) * Math.cos(Math.toRadians(lat2)) *
              Math.sin(dLng/2) * Math.sin(dLng/2);
   double c = 2 * Math.atan2(Math.sqrt(a), Math.sqrt(1-a));
   float dist = (float) (earthRadius * c);

   return dist;
   }
```

share improve this answer

edited May 23 '17 at 12:10
Community ♦
1 1

answered May 8 '09 at 2:11
Espen Herseth Halvorsen
3,962 7 29 38

“How to calculate the distance between two coordinates?”

Based on another question on stackoverflow, I got this code.. This calcul in miles :)

```
public static float distFrom(float lat1, float lng1, float
   double earthRadius = 6371000; //meters
   double dLat = Math.toRadians(lat2-lat1);
   double dLng = Math.toRadians(lng2-lng1);
   double a = Math.sin(dLat/2) * Math.sin(dLat/2) +
              Math.cos(Math.toRadians(lat1)) * Math.cos(Mat
              Math.sin(dLng/2) * Math.sin(dLng/2);
   double c = 2 * Math.atan2(Math.sqrt(a), Math.sqrt(1-a));
   float dist = (float) (earthRadius * c);

   return dist;
   }
```

share improve this answer

edited May 23 '17 at 12:10

Community ♦

1 • 1

```
public static float distFrom(double lat1, double lng1, double lat2, double lng2) {
        double earthRadius = 3958.75;
        double dLat = Math.toRadians(lat2 - lat1);
        double dLng = Math.toRadians(lng2 - lng1);
        double a = Math.sin(dLat / 2) * Math.sin(dLat / 2) + Math.cos(Math.toRadians(
        double c = 2 * Math.atan2(Math.sqrt(a), Math.sqrt(1 - a));
        double dist = earthRadius * c;

        int meterConversion = 1609;

        return new Double(dist * meterConversion).floatValue();
}
```

```
public static int distFrom(float lat1, float lng1, float lat2, float lng2) {
    double earthRadius = 6371; //kilometers
    double dLat = Math.toRadians(lat2-lat1);
    double dLng = Math.toRadians(lng2-lng1);
    double a = Math.sin(dLat/2) * Math.sin(dLat/2) +
        Math.cos(Math.toRadians(lat1)) * Math.cos(Math.toRadians(lat2)) *
            Math.sin(dLng/2) * Math.sin(dLng/2);
    double c = 2 * Math.atan2(Math.sqrt(a), Math.sqrt(1-a));

    return Math.abs((int) (earthRadius * c));
}
```

chrome

wiki/Haversine_formula – dotjoe May 8 '09 at 1:55

ormula -- movable-type.co.uk/scripts/latlong-vincenty.html -- if you care about the Earth not quite being a sphere – mob Nov 10 '09 at 23:22

1 refer to this blog ..... xebee.xebia.in/2010/10/28/working-with-geolocations – Robin Nov 1 '10 at 10:28

show 5 more comments

Answers

active oldest votes

177

Based on another question on stackoverflow, I got this code.. This calculates the result in meters, not in miles :)

```
public static float distFrom(float lat1, float lng1, float lat2, float lng2) {
    double earthRadius = 6371000; //meters
    double dLat = Math.toRadians(lat2-lat1);
    double dLng = Math.toRadians(lng2-lng1);
    double a = Math.sin(dLat/2) * Math.sin(dLat/2) +
               Math.cos(Math.toRadians(lat1)) * Math.cos(Math.toRadians(lat2)) *
               Math.sin(dLng/2) * Math.sin(dLng/2);
    double c = 2 * Math.atan2(Math.sqrt(a), Math.sqrt(1-a));
    float dist = (float) (earthRadius * c);

    return dist;
    }
```

share improve this answer

edited May 23 '17 at 12:10 Community ♦ 1 ● 1

answered May 8 '09 at 2:11 Espen Herseth Halvorsen 3,962 ● 7 ● 29 ● 38

17 Why convert to Float and then back to float? – Steve Kuo May 8 '09 at 4:05

3 `return (float) (dist * meterConversion)` – mob Nov 10 '09 at 23:18

6 The earth radius is 3958.75 in *miles* not kilometers. So your code above returns miles. – swinefeaster Nov 12 '11 at 8:55

5 @swinefeaster: There's meterConversion constant in the second to last line which turns miles into meters. – Keith Irwin Nov 21 '11 at 2:30

Welcome to ExampleStack!

Code Template

Undo Selection Copy

```
public static [ ] distFrom( [ ] lat1, [ ] lng1, [ ] lat2,
[ ] lng2) {
    double earthRadius = [ ]; //meters
    double dLat = Math.toRadians(lat2-lat1);
    double dLng = Math.toRadians(lng2-lng1);
    double a = Math.sin(dLat/2) * Math.sin(dLat/2) +
               Math.cos(Math.toRadians(lat1)) * Math.cos(Math.toRadians(lat
2)) *
               Math.sin(dLng/2) * Math.sin(dLng/2);
    double c = 2 * Math.atan2(Math.sqrt(a), Math.sqrt(1-a));
    [ ]
    [ ]
    [ ]

    [ ]
    }
```

Adaptation Categories: Code Hardening Resolve Compilation Error Exception Handling Logic Customization Refactoring Miscellaneous

2 Similar GitHub Examples

GitHub link watch: 0 star: 0 fork: 0

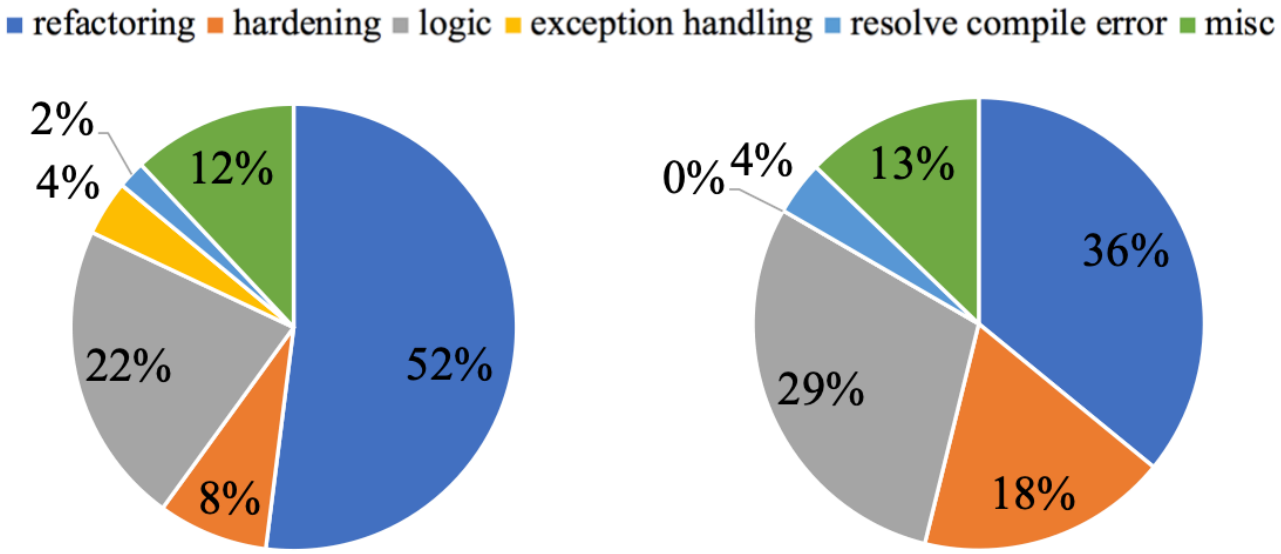
```
        public static float distFrom(double lat1, double lng1, double
lat2, double lng2) {
                double earthRadius = 3958.75;
                double dLat = Math.toRadians(lat2 - lat1);
                double dLng = Math.toRadians(lng2 - lng1);
                double a = Math.sin(dLat / 2) * Math.sin(dLat / 2) + Math.c
os(Math.toRadians(lat1)) * Math.cos(Math.toRadians(lat2)) * Math.sin(dLng /
 2) * Math.sin(dLng / 2);
                double c = 2 * Math.atan2(Math.sqrt(a), Math.sqrt(1 - a));
                double dist = earthRadius * c;

                int meterConversion = 1609;

                return new Double(dist * meterConversion).floatValue();
        }
}
```

Within-Subjects User Study

- Sixteen students from UCLA Computer Science
- Two code reuse tasks
  - Control: view a code example and search online
  - Experiment: view similar code in GitHub using ExampleStack

	Task Description	LOC	GitHub Clone#
Task I	compute the distance between two coordinates on earth	12	2
Task II	get the relative path of a given file and a root folder	74	2
Task III	encode an array of bytes to a hexadecimal string	12	17
Task IV	add animation to an Android view	29	4

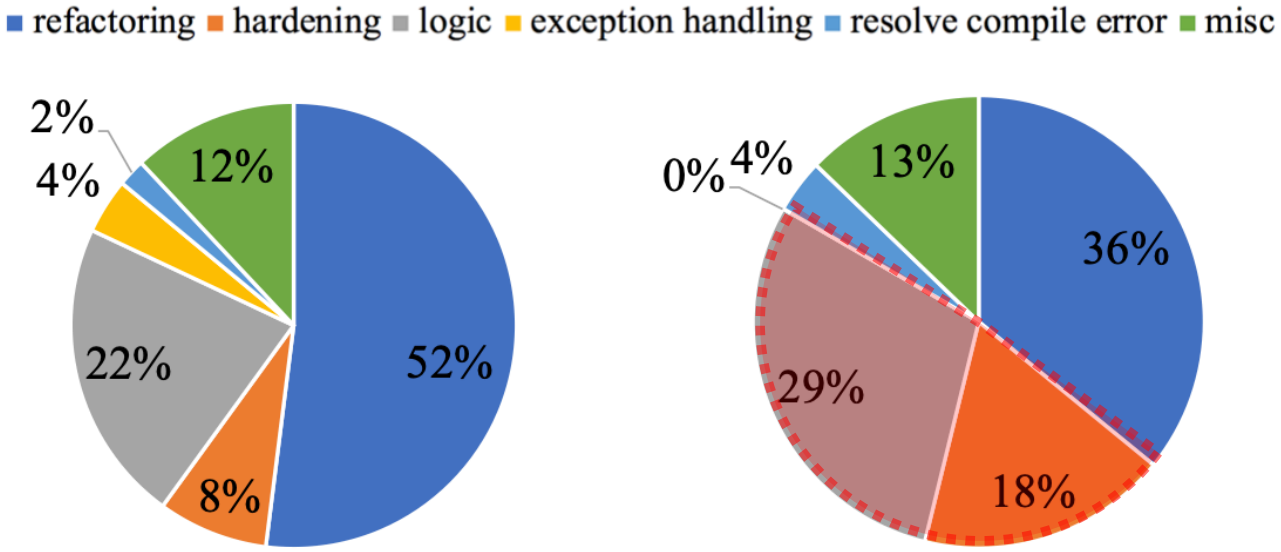
Finding 1. Viewing variations in similar GitHub code inspires new adaptations that are otherwise overlooked.



Without ExampleStack

With ExampleStack

Finding 1. Viewing variations in similar GitHub code inspires new adaptations that are otherwise overlooked.



Without ExampleStack

With ExampleStack

Finding 2. Seeing similar code is more useful than overwhelming.

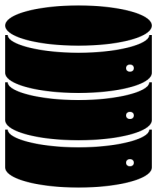
P5: “It **highlights the best practices followed by the community** and **prioritizes the changes** that I should make first”

P6: “Super nice, it seems like **the fast path to reach consensus** on a particular operation”

P9: “[It is] reassuring to know that the same code is used in production systems and to **know the common pitfalls**”

P14: “I would have completely **forgotten about the null check without seeing it in a couple of [GitHub] examples**”

Contributions



1. Make available a large-scale dataset of reused code between SO and GitHub.

2. Rigorously codify common adaptation patterns and create a taxonomy

3. Quantify the frequencies of common adaptations

4. Build a prototype and conduct a user study

Dataset and Tool: https://github.com/tianyi-zhang/ExampleStack-ICSE-Artifact