



Software Engineering Laboratory

井上研究室(ソフトウェア工学講座) Software Engineering Lab



Software Engineering Laboratory, Department of Computer Science, Graduate School of Information Science and Technology, Osaka University

1

研究室のミッション Lab's Mission

- ・ **素晴らしいソフトウェア**を作り、**よりよい社会の発展**を目指す！
Make great software and rich society
- ・ **ソフトウェアに関する技術**を探求し、**ツールとして公開**する
Develop software technologies and distribute their tools
 - 何が問題か、よく考えよう Find and think issues
 - 作って動かし、確かめよう Implement and see what happens
 - 結果を発表し、評価されよう Present results and get evaluation

ソフトウェア工学 Software Engineering

- ・ 品質の高いソフトウェアを、限られたコストと時間内で開発・保守するための技術を扱う学問分野
- ・ 井上研究室はソフトウェアの分析や再利用などが中心



Software Engineering Laboratory
Department of Computer Science, Graduate School of Information Science and Technology, Osaka University

2

研究室で得られる(得て欲しい)こと Lab's Benefits

- ・ **ソフトウェアのプロ**になる
Professional of software
- ・ **問題解決能力**が身に付く
Improving problem solving capability
- ・ **社会性**が向上する
Improving social affair
- ・ **国際感覚**が身につく
Global communication capability



Introduction

ring Bimodal program analysis via a
bliness lens

fix: Error Detection
Innaturalness
fix: Error Correction
ensibility
sh Report Clustering
Information Retrieval + Naturalness



Software Engineering Laboratory
Department of Computer Science, Graduate School of Information Science and Technology, Osaka University

3

多様な検出ツールに対応する コードクローン分析環境

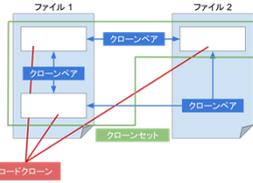
コードクローン:
ソースコード中の類似または一致した部分

- コピーペーストなどで生じる
- 一括でいろいろできると便利
バグ修正、集約、部品化、...

コードクローン検出ツールの問題

- ツールインストールが大変
- 入出力仕様ばらばら、結果比較できず
- パラメータばらばら、能力差が大きい

解決: ツール実行環境をコンテナ化・ツール差を吸収するインターフェイスをWebで提供



 ユーザ

→ 利用 →

Web ブラウザ

アクセス

 CCX





Software Engineering Laboratory
Department of Computer Science, Graduate School of Information Science and Technology, Osaka University

4

深層学習を用いたコード分析

深層学習：神経回路をモデルにしたニューラルネットワークを用いる機械学習

- 計算機の発展によって実用的に
- 多様な特徴を認識・高い精度

グラフ畳み込みネットワーク(GCN)を用いた類似コード検索

- GCN: グラフを対象とした深層学習
- 抽象構文木を学習させ、「意味が似ている」コードを高精度で探せるように

5

機械学習を用いたコードの良さ判定

プログラミングコンテスト

- 与えられた問題を解けるコードを書く
- 正解数や回答時間で成績(レーティング)が付けられる

問題：コードの良し悪しを判定したい

- レーティングが高い人はきっと良い
- 良いコードかどうかを自動判定したい

解決：機械学習を用いた判定手法

6

OSSの開発・発展をサポートする

- **オープンソースライセンス**
 - ソフトウェアやそのソースコードを再利用・再配布する際の約束事
 - 著作権表示とともに、コメントやテキストファイルで記述される
- **ライセンス・著作権の問題**
 - コピーしたコードのライセンスを守れているか？
 - コードを書いた人とコメントの著作権表示に矛盾がないか？

- **開発履歴を調査して、ライセンス違反や著作権表示の矛盾を発見**

```

/* spear machine family generic header file
 * Copyright (C) 2009-2012 ST Microelectronics
 * Rajeev Kumar <rajeev-dlh.kumar@st.com>
 * Viresh Kumar <vireshk@kernel.org>
 * This file is licensed under the terms of the GNU
    
```

Person	Tokens	Prop
Arnd Bergmann	170	82.52%
Viresh Kumar	29	14.08%
...		

7

ログ取得量の動的変更

- プログラムを実際に動かして解析すると、ログの量が膨大になる
- **プログラムの動作を監視して、未知の挙動が起きたときだけ詳細なログをとる**
 - 怪しい箇所のログを効率的に確認できる

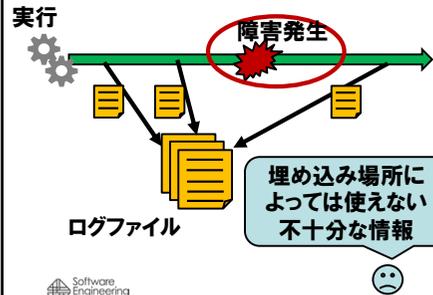
8

Javaを対象としたNear-全知デバッグ



デバッグしたくてlog出力を埋め込んで実行することは多いけど...

全命令を記録すると解決するが、素直にやると大量のログ(10MB/sec)
解決: Near-全知デバッグ
(Near Omniscient Debugging)



```
void methodA(int var) {
  var = methodB(var);
  while (var > 0)
    var = methodC(var);
  System.out.println(var);
}
```

時系列 提案手法

1	1
2	2
3	3
4	4
5	5
6	6
7	7
8	8
9	9
10	10
11	11
12	12
13	13
14	14
15	15
16	16
17	17
18	18
19	19

9

新4年生の年間計画 Annual Schedule

- 新人研修(春から初夏にかけて) Basic studies
 - 新人論議
 - Android を用いた携帯ソフト開発演習: Javaプログラミングのスキル向上、グループ開発経験、各種ツールの習得
- 院試勉強(夏) Preparation for exam
- 研究(秋から初冬にかけて) Research
 - 関連文献の読み込み
 - 手法の考案、他のツールの利用、新たな実装など
 - 評価実験の計画、実施
- 卒論の仕上げ(冬) Bachelor thesis
 - 論文の執筆、タイトル決め
 - スライド作り、発表練習



10

各種活動 Other activities

- 花見 Hanami (April)
- 新歓コンパ Welcome party (April)
- スポーツ Sports (Anytime)
- 随時コンパ Party (Summer)
- 研究室旅行 Lab's trip (Aug., Sep.)
- 忘年会 End-year party (Dec.)
- 追いコン Farewell party (Feb.)
- 卒業バーベキュー BBQ party (March)



昨今の事情でどこまでできるかわかりませんが...

11

ぜひソフトウェアの研究に興味を持って
井上研に参加してください!

12