

## 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

## 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**

## 研究室で得られる(得て欲しい)こと Lab's Benefits

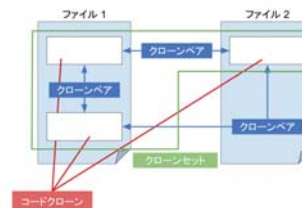
- ・ **ソフトウェアのプロになる**  
Professional of software
- ・ **問題解決能力が身に付く**  
Improving problem solving capability
- ・ **社会性が向上する**  
Improving social affair
- ・ **国際感覚が身につく**  
Global communication capability



## コードクローン分析

ソースコード中で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、...



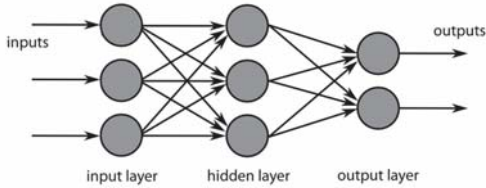
クローン検出ツール結果の比較

- 問題: ツールの出力結果がばらばら  
動作オプションで結果が変わる  
ツールによって得手不得手がありがち
- 解決: 複数の検出結果を可視化

# 深層学習を用いたコード分析

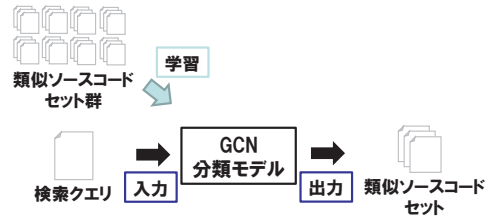
深層学習：神経回路をモデルにしたニューラルネットワークを用いる機械学習

- 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



# 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

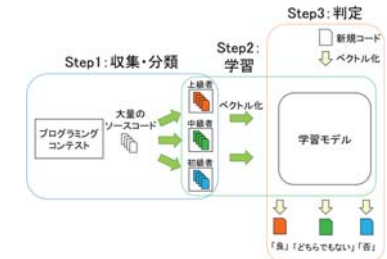
A screenshot of a Codeforces ranking table titled 'Rating: users participated in recent 6 months'. It shows a list of users with their ratings and the number of problems solved.

Who	#	=
1	162	3549
2	108	3520
3	196	3378
4	41	3351

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



# 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?

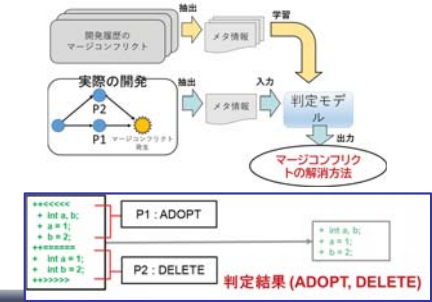
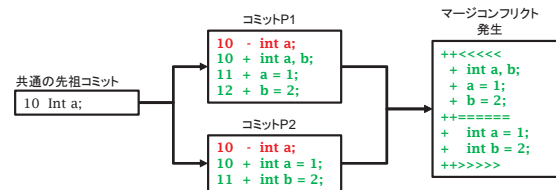
- 開発履歴を調査して、ライセンス違反や著作権表示の矛盾を発見

```
/*
 * spear machine family generic header file
 *
 * Copyright (C) 2009-2012 ST Microelectronics
 * Rajeev Kumar <rajeev-dih.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%
...		

# マージコンフリクトの解消モデル作成

- マージコンフリクト
  - 複数人で並行開発をした成果を統合するとき起きる
  - 自動で統合できないことがある
  - 手動での解消はコストがかかる
- 開発履歴から、過去のどのような修正がされたかを学習
  - 新しいコンフリクトが起きた時に、修正の指針を示す

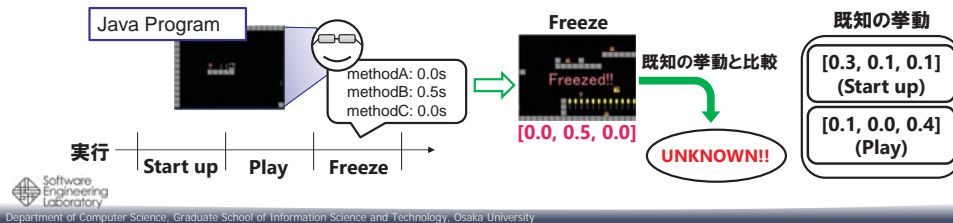


```
#####
+ int a;
+ b = 1;
+ b = 2;
#####
P1: ADOPT
P2: DELETE
#####
+ int a;
+ int b = 2;
#####
判定結果 (ADOPT, DELETE)
```

## 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



## 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
  - 新人輪講
  - 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**



## 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)



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！