

科学研究費助成事業 研究成果報告書

平成 28 年 6 月 15 日現在

機関番号：12102

研究種目：若手研究(B)

研究期間：2013～2015

課題番号：25730036

研究課題名(和文)過去の開発事例に基づくメソッド名とメソッドボディの補完システムの開発

研究課題名(英文)Recommendation system for a method name and method body content using code corpus

研究代表者

早瀬 康裕 (Hayase, Yasuhiro)

筑波大学・システム情報系・助教

研究者番号：40423090

交付決定額(研究期間全体)：(直接経費) 3,300,000円

研究成果の概要(和文)：オブジェクト指向プログラムのメソッド名とクラス名の命名を支援する手法として、メソッドやクラスの内容から適切と考えられる名前を推薦する手法を提案し、評価した。さらに、メソッド名や周辺の情報から、メソッドボディが空の状態で使用できるコード推薦手法を提案し、評価した。どちらの手法においても、既存の大量のソースコードをコーパスとして用いることで知識を学習し、そこから推薦される情報を生成する。

研究成果の概要(英文)：We developed and evaluated an approach for supporting naming a method and a class. The approach generates candidate of method names from content of the method and the class respectively. Furthermore, we developed a novel code completion approach, which can be used when a method body is empty, leveraging a name of the method and its context. Both approaches leverage the knowledge learned from huge size of code corpus for generating candidates.

研究分野：ソフトウェア工学

キーワード：コード補完 情報推薦システム 識別子名 メソッドボディ オブジェクト指向プログラミング

1. 研究開始当初の背景

一般に、ソフトウェアの長い寿命の中では多くの保守作業が実施されるが、その中でプログラムの理解に大きなコストが割かれている。ソフトウェア開発者は、プログラム理解に様々な手がかりを利用するが、その中でも識別子の名前は非常に重要な手がかりであるとされている。このため、適切な識別子名を与えることは、ソフトウェア開発において重要な意味を持つ作業となっている。この課題に対して、代表者は過去の研究において、開発者が新規にメソッドを作成する場合に、そのメソッドの周囲の情報から適切なメソッド名を推測し、候補を提示する手法を提案した。[早瀬科研 09, 鬼塚 12]

一方、ソフトウェア開発を効率的に行うための技術の 1 つに、コード補完 (code completion) がある。コード補完とは、開発者が編集しているソースコードに対し、次に入力されるであろうコード片を推測し、推薦する機能である。多様なライブラリやフレームワークを駆使する現代のソフトウェア開発では、コード補完システムは欠くことができない機能となっており、その機能を充実させる研究が数多く実施されている [Bruch09, 山本 12] が、メソッドを新規に作成した場合のような、メソッドボディが空の状態での効率的な推薦が行えるシステムは存在しなかった。

2. 研究の目的

本研究では、メソッド名とその内容との相互の関連に着目し、どちらか片方の情報と、周囲の情報とを合わせて用いることで、もう一方を推定・推薦するシステムを提案する。具体的には、以下の 2 つの手法を提案する。

- (1) 既存のメソッドに対して、その内容と周囲の情報を基準として、より良いと考えられるメソッド名を推薦する手法
- (2) 新規に作成するメソッドに対して、メソッド名と周辺の情報を基準として、そのメソッド内で呼び出すメソッドを推薦する手法

前者の手法により、開発者が少ない手間で、

良いメソッド名をつけられるようになると期待される。後者の手法により、高精度なコード補完が使用できる場面が増えると期待される。

3. 研究の方法

前準備として、コード補完に必要な知識の獲得のために、実績のあるソースコードの収集と、収集したソースコードのモデル化、そしてデータマイニングによる規則の獲得という順序で、作業を行なう。具体的には、オープンソースソフトウェアのリポジトリなどから、品質が高く、良いメソッド名が命名されていると期待されるソースコードを収集した。さらに、メソッド名の文法構造や、メソッドボディ内の処理内容、周囲の情報などを集合としてモデル化し、それに相関ルール分析を行うことで、知識を獲得した。

次に、獲得した知識を利用してメソッド名やメソッドボディの推薦を行うシステムを、統合開発環境 Eclipse プラグインとして実装し、評価を行なった。

さらに、当初目標になかった研究成果として、メソッド名の推薦の技術を、応用・発展させ、クラス名の推薦を行う手法を、提案・開発した。

4. 研究成果

まず、メソッド名の推薦手法については、大量のソースコードに対してメソッド名とボディなどの関連情報との分析を行い、過去の開発履歴から得られたメソッド名の一部を組み合わせて生成するメソッド名候補を、一覧の形で提示する手法を開発した。システムを評価した結果、特にメソッド名のうち動詞部については、高い精度で推薦できることが分かった。これらの成果について、国内学会 FOSE2013 [学会発表 7] と、国際会議 CSMR-WCRE [学会発表 6] においてそれぞれ成果を発表した。特に FOSE2013 における発表について、第一著者に対して IEEE Computer Society Japan Chapter FOSE Young Researcher Award 2013 が授与された。

次に、メソッドボディの推薦手法については、初年度から実現方法の検討を進め、国際ワークシ

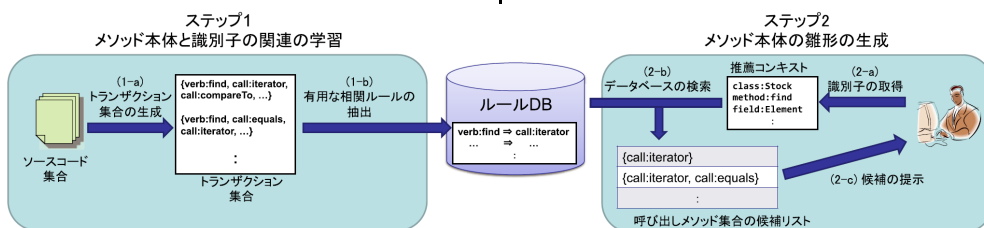


図 1. メソッドボディの推薦手法

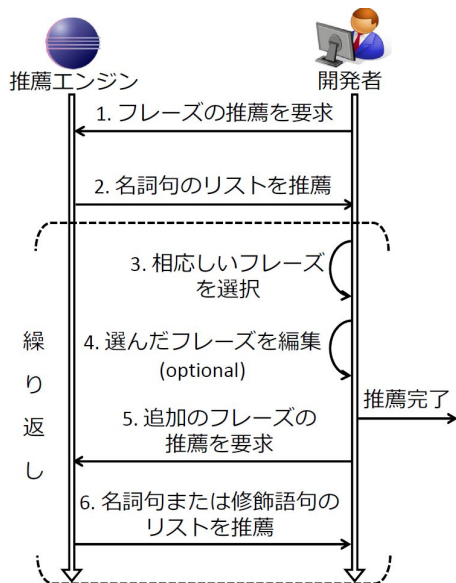


図 2. クラス名の推薦手法

ヨップにおいて手法を提案 [学会発表 8] して、国内外の研究者と議論を深めた。その後提案手法を実装 (図 1) して評価を行い、その成果を研究会と論文誌で報告した。 [雑誌論文 1, 学会発表 4]

クラス名の推薦手法については、まず基礎的な調査として、クラス名の構造と、クラス名で用いられている単語やその数を確認し、その結果を報告した [学会発表 3]。続いて、その調査結果に基づいて、メソッド名の推薦手法と同様にクラス名候補の一覧を提示する手法を提案し、これを評価した結果とあわせて報告 [学会発表 5] した。この評価の結果、メソッド名と同様の手法はクラス名にはうまく適合しないことが判明した。そこで、クラス名の特徴である文法構造や、単語の多様性を利用し、段階的かつ対話的なクラス名推薦手法 (図 2) を提案 [学会発表 1, 学会発表 2] し、問題を克服した。

<引用文献>

[早瀬科研 09] 早瀬 康裕, “シソーラス情報を用いたソフトウェア開発の支援”, 科研若手研究(B), 研究課題番号 21700031, 2009年度~2011年度

[鬼塚 12] 鬼塚 勇弥, 早瀬 康裕, 石尾 隆, 井上 克郎: “ソースコード中に出現する動詞-目的語関係を利用したメソッド名の命名支援手法”, 電子情報通信学会技術研究報告, vol. 111, no. 481, SS2011-57, pp. 1-6, 2012/3/13

[Bruch09] Marcel Bruch, Martin Monperrus, and Mira Mezini. “Learning from examples to improve code completion systems”, In

Proceedings of the the 7th joint meeting of the European software engineering conference and the ACM SIGSOFT symposium on The foundations of software engineering (ESEC/FSE '09), pp. 213-222, 2009.

[山本 12] 山本 哲男 他, “ソースコードコーパスを利用したシームレスなソースコード再利用手法”, 情報処理学会論文誌, Vol.53, No.2, pp.644-652, 2012

5. 主な発表論文等

[雑誌論文](計 1 件)

[1] 早瀬 康裕, 鬼塚 勇弥, 山本 哲男, 石尾 隆, 井上 克郎: “API 呼び出しとメソッド周辺の識別子の実績に基づいた API 集合推薦手法”, 情報処理学会論文誌, Vol.56, No.2, pp.692-700, 2015/02/16 (査読あり)

[学会発表](計 8 件)

[1] Hiroki Fukuda, Yasuhiro Hayase, Hiroyuki Kitagawa: “Towards Interactive Phrase Suggestion for Naming Classes” (poster), 23rd IEEE International Conference on Software Analysis, Evolution, and Reengineering (SANER 2016), Osaka University, Suita, Osaka, Mar 15, 2016

[2] 福田 宏樹, 早瀬 康裕, 北川 博之: “クラス名命名を支援する段階的なフレーズ推薦手法”, 情報処理学会 第 191 回 SE 研究発表会, 大阪大学吹田キャンパス, 大阪府吹田市, 2016/3/14

[3] 福田 宏樹, 早瀬 康裕, 北川 博之: “クラス名の単語列に対する品詞列ごとのクラス名数の定量的調査”, 情報科学技術フォーラム講演論文集 /13(1), pp.131-132, 筑波大学, 茨城県つくば市, 2014/9/4

[4] 鬼塚 勇弥, 早瀬 康裕, 山本 哲男, 石尾 隆, 井上 克郎: “メソッド周辺の識別子名とメソッド本体の API 利用実績に基づいた API 集合推薦手法の提案と評価”, 情報処理学会研究報告 2014-SE-183(16), 化学会館, 東京都千代田区, 2014/3/20

[5] 福田 宏樹, 早瀬 康裕, 北川 博之: “クラス名の文法構造と周辺の識別子を用いたクラスの命名支援”, 情報処理学会第 76 回全国大会, 東京電機大学東京千住キャンパス, 東京都足立区,

2014/03/11

- [6] Yuki Kashiwabara, Yuya Onizuka, Takashi Ishio, Yasuhiro Hayase, Tetsuo Yamamoto and Katsuro Inoue: “ Recommending Verbs for Rename Method Using Association Rule Mining ”, IEEE CSMR-WCRE 2014 ERA, pp. 323-327, Antwerp, Belgium, February 3-6, 2014
- [7] 柏原 由紀, 鬼塚 勇弥, 石尾 隆, 早瀬 康裕, 山本 哲男, 井上 克郎: “ 相関ルールマイニングを用いたメソッドの命名方法の分析 ”, ソフトウェア工学の基礎XX (第 20 回 ソフトウェア工学の基礎ワークショップ), ゆのくに天祥, 石川県加賀市, pp.25-35, 2013/11/28
- [8] Yuya Onizuka, Yasuhiro Hayase, Tetsuo Yamamoto, Yuki Kashiwabara, Takashi Ishio and Katsuro Inoue: “ Towards Generating Templates of Method Body Based on Method Name and Related Identifiers ”, 8th International Workshop on Advanced Modularization Techniques (AOAsia/Pacific 2013), pp. 13-14, Fukuoka International Congress Center, Fukuoka, Japan, March 25, 2013

6 . 研究組織

(1)研究代表者

早瀬 康裕 (HAYASE, Yasuhiro)
筑波大学・システム情報系・助教
研究者番号 : 4 0 4 2 3 0 9 0