

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

福田宏樹[†] 早瀬康裕[‡] 北川博之[‡]

[†]筑波大学情報学群情報科学類

[‡]筑波大学システム情報系情報工学科

1 はじめに

保守性の高いプログラムを書くためには、識別子にその役割を表す名前を付けておくことが必要である。保守作業者は、識別子名の意味からそのプログラム要素の役割を推測する [5] ためである。

しかしながら、そのような命名には、ソフトウェアが扱う問題領域の語彙といった専門知識が必要になる。そのため、経験豊富でない開発者が識別子名を自分で考えるのは難しいという問題があった。

柏原は、識別子のうち、メソッド名の改名時を対象として、メソッド名の動詞-目的語構造を利用し、メソッドボディや周囲に現れる識別子の情報を元にメソッド名を推薦する手法を提案した [6]。この手法は、周辺の識別子とメソッド名とを結びつける知識を、相関ルール [3] を用いて表現している。

本稿では、柏原手法のアイデアをクラス名の推薦に応用し、内容のあるクラスへの命名時を対象として、クラス名を推薦する手法を提案する。具体的には、我々が数種に分けたクラス名の構造ごとに、その構造を利用し、クラスボディや周囲に現れる識別子の情報を元にクラス名を推薦する。内容のあるクラスへの命名時とは、クラス名の改名時と、クラスが新たに出現するリファクタリング [1, 2] でのクラスへの命名時を想定している。提案手法を利用することで、開発者は、リファクタリング時により適切なクラス名を付けやすくなることを期待される。

本稿では、第2節で提案手法を説明し、第3節でまとめをそれぞれ述べる。

2 提案手法

提案手法において、クラス名の文法構造とクラス周辺の識別子を用いてクラス名を推薦する流れを、図1に沿って説明する。提案手法は、大きく分けて、予め、

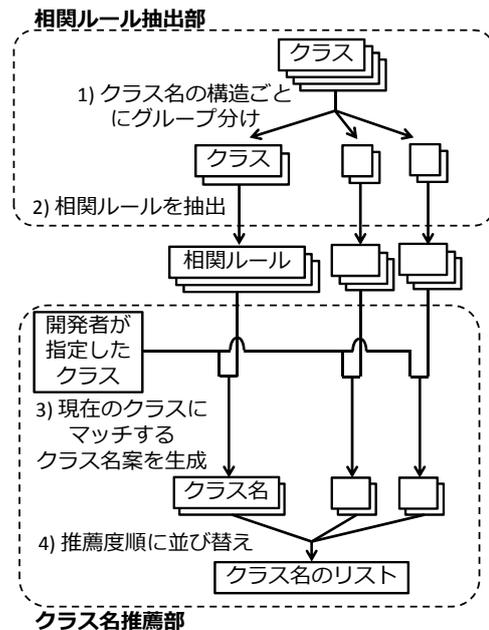


図 1: 提案手法の概要図

既存のクラスに対する大量の命名事例から、クラス名が同一の構造を持つクラスのグループごとに、周辺の識別子とクラス名とを結びつける相関ルールの抽出を行っておく部分（相関ルール抽出部）と、抽出したルールを用いて、クラス名の構造ごとにクラス名案を生成し、それらを推薦度順に並び替えて開発者に提示する部分（クラス名推薦部）との2段階から成る。相関ルール抽出部については2.1節で、クラス名推薦部については2.2節で、それぞれ説明する。

2.1 相関ルール抽出部

相関ルール抽出部では、既存のソースコード群中のクラスから、クラス名が同一の構造を持つクラスのグループごとに、周辺の識別子とクラス名とを結びつける相関ルールの抽出を行う。相関ルール抽出部は、より細かくは、以下に述べる1)と2)の2ステップから成る。

1) 既存のソースコード群中からクラスを集め、それらを、クラス名が同一の構造を持つと見なせるもの同士でグループ分けする。このとき、WordNet[4]とヒュー

Suggesting class names based on grammatical structure of class names and surrounding identifiers

Hiroki FUKUDA[†](hfukuda@kde.cs.tsukuba.ac.jp),

Yasuhiro HAYASE[‡](hayase@cs.tsukuba.ac.jp) and

Hiroyuki KITAGAWA[‡](kitagawa@cs.tsukuba.ac.jp)

[†]College of Information Science, University of Tsukuba

[‡]Faculty of Engineering, Information and Systems, University of Tsukuba

リストとを用いて、クラス名を構成する単語に対して品詞判定を行った上で、単語数と判定した品詞情報とを基に、クラス名の構造を判定する。具体例を挙げると、1単語目が形容詞で2単語目が名詞単数形のクラス名と、1単語目が動詞過去分詞形で2単語目が名詞単数形のクラス名とは、両方とも、2単語目の名詞が意味の中心となり、1単語目が2単語目を限定するという構造を持つことから、同一の構造を持つと判定する。単語数が異なるクラス名は構造も異なると見なす。

2) 1) でグループ分けした個々のクラスを、それぞれトランザクションとして表現し、グループごとに相関ルールマイニングを行う。トランザクションは表1の右列に示すものをアイテムとして含む。本稿では、表1の上側にある6種のアイテムを総称してクラスのコンテキストと呼び、下側にあるアイテムを総称してクラス名の構成単語と呼ぶ。相関ルールマイニングでは、条件部にクラスのコンテキストのみが現れ、帰結部にクラス名の構成単語のみが現れる形の相関ルールのみを、クラス名推薦部で利用するために保存しておく。

2.2 クラス名推薦部

クラス名推薦部では、相関ルール抽出部で得たルールを利用して、クラス名の構造ごとにクラス名案を生成し、それらをまとめて推薦度順に並び替え、開発者に提示する。クラス名推薦部は、より細かくは、以下に述べる3)と4)の2ステップから成る。

3) 開発者がクラス名の推薦を希望して指定したクラスのコンテキストにマッチする候補クラス名集合を、クラス名の構造ごとに生成する。まずは、開発者がクラス名の推薦を希望して指定したクラスのコンテキストに対して、この現在のコンテキストにマッチする相関ルールを、それぞれのグループごとに検索する。具体的には、グループごとに、条件部が現在のコンテキストの部分集合となる相関ルールを、全て取得する。その後、

それぞれのグループごとに、得た相関ルールの帰結部に現われている単語を組み合わせて、候補クラス名の集合を作る。具体的には、帰結部に現れているクラス名の k 単語目の集合を W_k 、このグループが表すクラス名の構造の単語数を n として、 $W_1 \times W_2 \times \dots \times W_n$ のように組み合わせを行い、候補クラス名の集合を作る。これにより、現在のコンテキストにマッチしたクラス名の集合を、クラス名の構造ごとに、相関ルールを利用して得ることができる。

4) 3) で得たクラス名を全てのグループから集め、そのクラス名の単語の由来元である相関ルールの確信度の積を推薦度として用い、推薦度が高い順に並び替えたクラス名のリストを、開発者に提示する。相関ルールの確信度は、現在のクラスのコンテキストから単語が得られる確率と見なせ、相関ルールの確信度の積は、現在のクラスのコンテキストからそれぞれの単語を組み合わせたクラス名が得られる確率と見なせることから、確信度の積を推薦度として用いる。

3 まとめ

本稿では、クラス名の文法構造とクラス周辺の識別子を用いて、内容のあるクラスへの命名時にクラス名を推薦する手法を提案した。

今後は、提案手法で元のクラス名を再現できるかどうかを評価する実験と、この手法を用いるとクラス名を付けやすくなるか、この手法で付けられたクラス名は分かりやすいかという観点からの被験者実験を行う予定である。

参考文献

- [1] M. Fowler. extractClass. <http://www.refactoring.com/catalog/extractClass.html>
- [2] M. Fowler. introduceParameterObject. <http://www.refactoring.com/catalog/introduceParameterObject.html>
- [3] R. Agrawal et al. Mining Association Rules between Sets of Items in Large Databases. SIGMOD 1993.
- [4] G. A. Miller. WordNet: A lexical database for English. Comm. ACM 38(11), 1995.
- [5] N. Pennington. Comprehension Strategies in Programming. Empirical Studies of Programmers: 2nd Workshop, 1987.
- [6] 柏原由紀. メソッド名とその周辺の識別子の相関ルールに基づくメソッド名変更支援手法. 大阪大学基礎工学部情報科学科特別研究報告, 2013.

表 1: トランザクションに含むアイテムの一覧

コンテキストのト	親クラス名
	親インターフェース名
	定義しているフィールドの型名
	定義しているフィールド名
	定義しているメソッドの返り値の型名
	定義しているメソッド名
クラス名の構成単語	クラス名の1単語目
	クラス名の2単語目
	⋮