HOKKAIDO UNIVERSITY

Title	キーワードのあいまい一致を導入したキーワードプログラミングシステムに関する研究［論文内容及び審査の要旨］
Author(s)	坂本, 悠輔
Degree Grantor	北海道大学
Degree Name	博士(情報科学)
Dissertation Number	甲第11746号
Issue Date	2015-03-25
Doc URL	https://hdl.handle.net/2115/58714
Rights(URL)	https://creativecommons.org/licenses/by-nc-sa/2.1/jp/
Type	doctoral thesis
File Information	Yusuke_Sakamoto_abstract.pdf, 論文内容の要旨

Hokkaido University Collection of Scholarly and Academic Papers : HUSCAP

学 位 論 文 内 容 の 要 旨

博士の専攻分野の名称　博士（情報科学）　氏名　坂本　悠輔

学 位 論 文 題 名

キーワードのあいまい一致を導入したキーワードプログラミングシステムに関する研究
（Study on Keyword Programming Systems with Ambiguous Keyword Matching）

現在のソフトウェア開発では, ライブラリやフレームワーク, 参考書, ウェブサイトなどに存在する既存のプログラムコードを再利用して目的とするコードを記述することが一般的である。しかし, それらのライブラリや API(応用するときのインタフェースの仕様) の集合は巨大かつ複雑であるため, ユーザ (プログラマ) が記憶したり検索したりして活用することが非常に困難であり, この問題に対応するために様々なソフトウェアツールの研究・開発が行われている。その 1 つの例としては, ユーザが必要とするコードスニペット (よく使われる短いコードの断片) を, 大量に存在するコードの集合から抽出する Ohloh Code のようなコードサーチエンジンやそれに関する研究がある。また別の例として, 広く普及している Eclipse などの統合開発環境には, ユーザが現在編集中のソースコードの情報を利用して, 現在位置で欲しいコードを補完するコードコンプリーションという機能が備わっている。

これらの研究の延長上にある野心的な試みとして, Little と Miller が提唱したキーワードプログラミングがある。これはまるで Web 検索エンジンのように, ユーザが与える少数のキーワードから, システムが適切な完成型のコードを自動生成することにより, ユーザがプログラミング言語の文法規則や API 群の詳細を覚えている必要を軽減する技術である。ユーザが入力する複数のキーワードは, 現在の文脈においてユーザが入力したい表現を検索するキーワードであるかのように捉えることができる。キーワードプログラミング以前の研究においては, 望ましいコード片をコードのデータベースから取得するため, そのコード片の出力や入力の型 (基本データ型やクラスの名前) をキーワードの一部に入力しなければならなかったが, キーワードプログラミングでは, 入力するキーワードはプログラミング言語の型名に拘束されず自由になった。このために, ユーザは型名などのプログラミング言語の文法の詳細な知識に乏しくても, コーディングの支援を受けることが可能となった。しかし, この研究は, さらに発展させて, ユーザにとっての使いやすさを一層改善させられる余地がある。

本論文では主に, (1) キーワードプログラミングをさらに使いやすくするために, WWW の検索エンジンが多数のウェブページを順位付けて出力するのと同様に, キーワードプログラミングにおいても多数のコード片を順位を付けて出力する必要があることを論じ, その効率の良い実装を行ったことと, (2) ユーザのキーワードの入力に着目し, その言語に不

慣れなユーザが, あいまいな記憶によってキーワードの一部分だけを覚えていて短縮して入力したときや, 一部を間違えて入力してしまったときにも対応できるようなツールの改良について行ったこと, さらに, (3) ツールを統合開発環境 Eclipse のプラグインとして実装してその性能評価を行ったことについて論じる。

本論文では上記の内容を次の構成で述べている。

第 1 章では, 現在普及している複数のプログラミングツールに関する説明と, キーワードプログラミングの紹介を行う。

第 2 章では, コードコンプリーション分野の関連研究として, 本研究を含む代表的な 8 つのシステムについて, 入力の種類, 出力の種類, 及び主な知識源の 3 つの観点から各々の特徴を論じている。

第 3 章では, 本研究の先行研究である Little と Miller が提唱したキーワードプログラミングについて, そのモデル (出力コード片が関数木で表現されること) や, そのノードの評価と特徴ベクトル, 説明ベクトルを用いた関数木の評価, そして関数木生成アルゴリズムについて述べる。

第 4 章では, キーワードプログラミングの理解および分析を容易にするために, 本論文独自の表現を導入して, 先行研究では明確に分析されていなかったキーワードプログラミングの問題点を明確に述べ, それを解決するために, 複数候補を提示するようにキーワードプログラミングシステムを拡張することが本質的に必要であることを議論している。

第 5 章では, Little と Miller の研究において, 出力コード片を構成する関数や変数の名前に出現する単語と完全に一致するキーワードを入力する必要性から生じる, ユーザの認知的なあるいは物理的な負担を軽減するために, キーワードがあいまい (不正確) であっても出来るだけ良い出力を得ることを目的として, 単語間の類似度に基づくキーワードのあいまい一致手法を導入して, オープンソースプロジェクトを対象とした実験の結果に基づき, このような拡張が有効であることを議論している。

第 6 章では, 本研究において開発したツールの実装について解説している。具体的には, 候補の評価値を計算する際に必要な「重み」をある程度自動調整する機能や実行時間を短縮する実装上の工夫などについて述べている。