

## 論文内容の要旨

博士論文題目

コードクローンの検出・分析技術の活用手法に関する研究

氏名

上村恭平

(論文内容の要旨)

ソースコード中に存在する、互いに一致、あるいは類似するソースコードの断片をコードクローンと呼ぶ。コードクローンの存在はソフトウェア開発の効率を低下させる要因の1つであり、適切に管理することが重要であるといわれている。そのため、ソースコード中から自動でコードクローンを検出する技術や、コードクローンの除去を支援する技術、一貫した修正を支援する技術などが提案されている。しかし、コードクローン関連技術を実際開発の中で利用するには課題がある。まず、既存のコードクローン検出ツールが特定の言語のみを対象としていることが挙げられる。次に、コードクローンの影響を分析するために、コードクローンの変更履歴を調査することがあるが、既存手法では大規模なソフトウェアへの適用が困難であること。最後に、コードクローン検出ツールの検出結果はソースコードの管理に利用するには余分な検出結果が多いこと。本論文ではこれら3つの問題を解決する手法についての検討結果がまとめられている。

初めに、コードクローン検出技術のハードウェア記述言語を用いた回路開発への応用に関する研究成果が報告されている。この研究では既存のコードクローン検出ツールが対応していない言語で記述されたソースコードに対して、簡易な変換を適用することでコードクローンを検出する手法を提案している。この手法は、本論で対象とした言語以外にも適用可能である。次に、メソッド単位の変更情報を記録した構文情報リポジトリである`hstorage`を利用した、コードクローンの履歴情報を抽出する手法を提案している。最後に、ソースコード削減可能量見積もり作業の支援を目的としたツールの開発を行っている。開発したツールはコードクローン検出結果から重複を除去し、ファイル単位で類似するソースコードをまとめた類似プログラム集合を利用者に提供する。

本論文の貢献は、以下の3つといえる：1) コードクローンの検出対象となる言語を容易に拡張する手段の提案。2) 長期間かつ細かなバージョン単位でのコードクローンの履歴分析の容易化。3) コードクローン検出結果の要約技術による利用性の向上。これらの研究成果は、コードクローンに関する研究や、実際開発の中でのコードクローン検出技術の活用の促進に貢献するものと言える。

氏名	上村 恭平
----	-------

(論文審査結果の要旨)

本論文は、ソフトウェアのプログラムコード中に発生する、互いに一致、あるいは類似するソースコードの断片（コードクローン）の検出および分析と、それを保守作業等に活用するための手法を提案するものである。

コードクローンの存在はソフトウェア開発の効率を低下させる要因の1つであり、適切に管理することが重要であるといわれている。そのため、ソースコード中から自動でコードクローンを検出する技術や、コードクローンの除去を支援する技術、一貫した修正を支援する技術などが提案されている。

コードクローンの研究分野には10年以上の歴史があり、プログラムテキストに対する高速な検出への取組から始まり、現在では意味の類似にまで踏み込んだ検出や、クローンの履歴の追跡、検出結果を用いた品質評価や保守への応用などに裾野が広がっている。

本論文では、クローン分野に於ける具体性・実用性の高い複数の研究テーマへの取組をまとめたものであり、以下の3つの貢献を有する。1) コードクローンの検出対象となる言語を容易に拡張する手段の提案。2) 長期間かつ細かなバージョン単位でのコードクローンの履歴分析の容易化。3) コードクローン検出結果の要約技術による利用性の向上。いずれの内容も新規手法の提案と評価のための実装、オープンソースや企業の実プロジェクトへの適用による評価を含んでおり、当該分野に於ける高い有用性を示したものである。

以上より、本論文の内容は、コードクローンに関する研究や、実際の開発中でのコードクローン検出技術の活用の促進に貢献するものであり、実社会における工学的な問題の定式化とそれに対する一定の解決策を実用的レベルにおいて提示しており、博士(工学)の学位を授与するに値すると判断できる。