

学 位 論 文 要 旨

氏 名 長谷川 卓也

題 目 プログラミング教育支援システムの開発とコーディングログの分析による学
習支援に関する研究

学位論文要旨（和文2,000字又は英文1,000語程度）

本研究の目的は、高等学校におけるテキスト型プログラミング言語を用いたプログラミング教育において、生徒のプログラミング過程を明らかにすることで、効果的な学習支援を行うための知見を得ることである。

本論文は、6章で構成されている。大きく2つの主題を含んでおり、一方はプログラミング教育支援システムの開発、もう一方は同システムを通じて得られたコーディングログに基づく学習分析である。

第1章では、本研究の背景について述べている。高等学校学習指導要領（平成30年告示）で、「情報Ⅰ」が必修科目となり、プログラミング教育の機会が大幅に拡大した。本研究では、まず、プログラミング学習と関連学習の習熟度の関係性に着目した分析を実施した。その結果、プログラミング教育を円滑に導入するための事前学習について示唆が得られたが、一方で、生徒のプログラミング過程そのものを対象としない分析方法の限界が課題として示された。

第2章と第3章は、プログラミング過程を記録する機能をもつプログラミング教育支援システムの開発を主題とする。

第2章では、プログラミング教育支援システムの開発における予備調査からシステムテストの工程について述べている。プログラミング教育支援システムは、主に生徒に操作が容易なプログラミング環境を提供する機能、及び生徒のプログラミング過程を記録し教師に提示する機能を備えるように設計され、Webサーバ上で稼働するように開発された。本システムのテストにおいて両機能が正常に動作することを確認した。

第3章では、本システムの授業における利用実績、及び運用を通じた評価について述べている。利用生徒に対するユーザビリティ調査によって操作性を検証し、概ね高い評価が得られていることを確認した。また、授業で得られたコーディングログの内容を目視によって検証した。ログに欠損や誤り等は見つからず、正常にプログラミング過程が記録されていることを確認した。

第4章と第5章は、本システムによって得られたコーディングログに基づく学習分析を主題とする。

第4章では、プログラムの類似度と習熟度の関係性に関する時系列分析について述べている。情

報科の単元「探索のプログラム」において、プログラム作成を伴う授業を実施した。プログラム作成にはプログラミング教育支援システムを用い、生徒のプログラミング過程をコーディングログとして記録した。次いで、コーディングログを用いて生徒が作成したプログラムと模範解答プログラムとの類似度を時系列に抽出した。類似度には、ユニグラムを用いたコサイン類似度、バイグラムを用いたコサイン類似度、及びレーベンシュタイン距離に基づく類似度を用いた。これらの類似度と習熟度の関係について分析した。比較のため、プログラミング過程における操作数、プログラムの文字数と習熟度の関係についても分析した。その結果、いずれの類似度も、操作数や文字数よりも強い相関を示し、特に、バイグラムを用いたコサイン類似度と習熟度の相関が強いことが明らかとなった。また、プログラム作成における初期段階から相関が強くなっていることが分かり、類似度に着目することで早期に習熟度を予測できる可能性が示唆された。

第5章では、機械学習による成績の下位群の検出と、プログラムにおける記述が不完全な箇所の可視化について述べている。プログラムの最小構成単位となる字句をトークンと言う。生徒が作成したプログラムに含まれるトークンの出現数に基づき、下位群を検出するための分類モデルを複数の機械学習の手法を用いて構築した。最も高い識別性能を示したのは勾配ブースティングを用いたモデルであり、再現率は0.77、適合率は0.78であった。プログラミング過程の操作数やプログラムの文字数に基づき構築した分類モデルよりも高い識別性能を示した。さらに、非下位群と比べて下位群で充足率が低かったトークンを特定し、そのトークンが模範解答プログラムの中で記述されている位置を、下位群のプログラムにおける記述の不完全な箇所として可視化した。これにより、重点的に指導すべき内容を明確にするための根拠を得ることができた。

第6章では、本研究の成果と課題について述べている。コーディングログの分析で得られた主要な知見として、①類似度に着目することにより、プログラミング過程の初期段階で習熟度を予測できる可能性があること、②プログラムを構成するトークンに着目することにより、習熟度の低迷が予想される生徒を検出するとともに、指導すべき学習内容を明確化するための根拠が得られることを示した。今後の課題として、教師がコーディングログを活用しやすい仕組みの構築、コーディングログによる分析手法の汎用性の検証、及び①の類似度と②の機械学習と可視化の手法を統合した、より有用性の高い分析手法の開発が挙げられる。