

プログラミング演習支援システムの定量的学習データを用いた成績不振兆候者の早期発見に関する研究

畝川 みなみ[†] 宇野 健[†]

県立広島大学 総合学術研究科 経営情報学専攻[†]

1. はじめに

初学者に対するプログラミング教育では、試験成績が落ち込む可能性のある成績不振兆候者を早期の段階で発見し対処することで、より効果的な指導が可能となる[1]。しかし、成績不振兆候者の抽出に関しては、学生へのアンケート等による主観的な情報では正確性に欠ける場合があるため、客観的なデータによる判定が望ましい。そのため、学習者の定量的学習データを収集・蓄積する学習環境の構築が必要である。

そこで本研究室では、Web ブラウザのみで利用できる初学者向けのプログラミング演習支援システムの開発・運用をおこなってきた。学習環境の構築を不要とするのみでなく、演習操作記録を取得・集計し、学習の過程を定量的に評価して学習者にフィードバックすることで、自発的な取り組みを向上させることができた[2]。

本論文では、システム運用によって蓄積した学習データを活用することにより、定量的データのみで成績不振兆候者の早期発見手法を提案し、実際の学習データに適用してその有効性を検証した。

2. プログラミング演習支援システム

本システムの開発は、初学者の自学習促進、および定量的学習データの収集を目的としておこなった。エディタとコンパイラを一つの画面に統合した C 言語演習システム(以下、演習システム)と、学習記録を集計し、学習状況を評価して表示するフィードバックシステム(以下、FBシステム)から構成される Web アプリケーションである。演習システム上でのコンパイルは、Ajax による非同期通信を用いて、サーバに設置した gcc でおこなう。コンパイルしたプログラムの実行は、PHP で開発したエミュレータを使用し、Web ブラウザ上で scanf 等の入力待ちを含むプログラムの実行を可能とした。

A study on at-risk Students Prediction Utilizing Quantitative Learning Data from The Support System for Programming Exercises.

[†]Minami UNEKAWA, Takeshi UNO

[†]Graduate School of Comprehensive Scientific Research of Prefectural University of Hiroshima

演習システムと共に開発した FB システムは、授業中および授業時間外の学習状況を A+~D の 5 段階で評価し、学習者および教員にフィードバックする。評価には課題達成率と操作得点(ファイルの保存、コンパイル、実行の回数)を用い、学習者や教員がシステムにアクセスした時点での学習記録をリアルタイムに集計することが可能である。

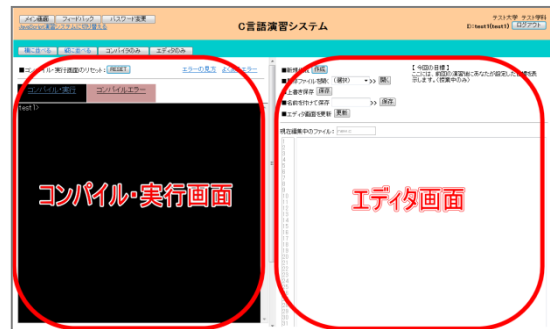
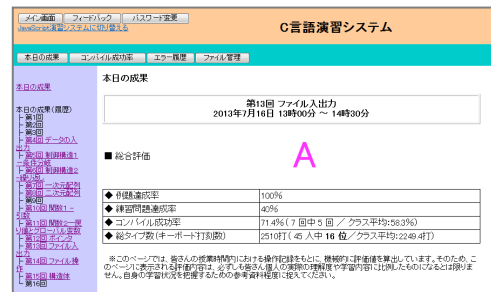


図1 演習システムのインターフェース



題のうち 1 問がメールによる提出課題で、それ以外は自学習用の任意課題であった。

3-2. 実験結果

3 年間にわたりおこなった本システムの運用実験の結果を以下に示す。ソースプログラムの編集履歴等詳細なデータも合わせると、累計 50 万件以上の学習データを蓄積した。2013, 2014 年度は FB システムを導入したため、2012 年度と比較すると全項目において値が向上した。

表 1 運用期間中の実績(一人あたり平均)

運用年度 (受講者数)	2012 (46)	2013 (44)	2014 (43)
保存(回)	270.4	414.3	490.3
コンパイル(回)	243.9	354.0	432.7
実行(回)	185.5	246.6	287.8
例題達成率(%)	84.9	94.6	91.8
練習問題達成率(%)	33.4	72.2	75.1

4. 成績不振兆候者の早期発見手法の提案

4-1. 提案手法の概要

本研究では、各回の授業終了時までの定量的学習データを用いて学習者の試験成績(中間試験と期末試験の平均得点の偏差値)を推定し、成績不振兆候者の早期発見を図る。推定時にしきい値を任意に設定し、それを下回る学習者を下位判定者として抽出する。そのうち、実際の試験成績もしきい値未満である学習者(成績下位者)の割合を「的中率」、成績下位者の中で下位判定者の占める割合を「抽出率」として推定精度を評価する。

また、中間試験後の推定については、推定精度向上のため中間試験結果をもとに補正をおこなう。具体的には、2014 年度のデータをもとに作成した補正式を適用し、得られた結果を新たな推定値としてその精度を検証する。

4-2. 推定式

試験成績の推定にあたり、2014 年度の運用で蓄積された全学習データに対して、試験成績を目的変数とした重回帰分析をおこなった。その結果、各説明変数の係数の有意確率 p が 5%未満で、且つ決定係数が 50%を超える重回帰式を 2 式作成できた。これらの式を各回の授業終了時までの学習データに適用し、試験成績の推定をおこなう。2 式のうち少なくとも一方の推定値がしきい値を下回れば「単純下位者」(単純判定)、単純下位者と判定された回数が平均値を上回れば「累積下位者」(累積判定)とすることとした。

4-3. 有効性の検証

提案手法を 2013 年度および 2014 年度の学習データに適用し、有効性を検証した。2013, 2014 年度ともに、FB システムの本格的な利用を開始した第 4 回の授業からのデータを用いた。

中間試験は 2013, 2014 年度ともに第 9 回の授業で実施した。

図 3, 4 に 2013, 2014 年度の的中率および抽出率を示す。縦軸は、各回での任意のしきい値(50~35)における結果の平均値である。単純判定よりも累積判定の方が比較的安定しており、その抽出率は 2 回目の推定である第 5 回の授業以降、2013 年度は 60%以上、2014 年度は 70%以上と高い結果となった。中間試験後の抽出率は、2013 年度が 70%以上、2014 年度が 80%以上となり、補正により抽出率の向上が確認できた。以上より、的中率のみでは単純判定が、抽出率のみでは累積判定が優位であり、成績不振兆候者の発見には両判定方法の併用が望ましい。

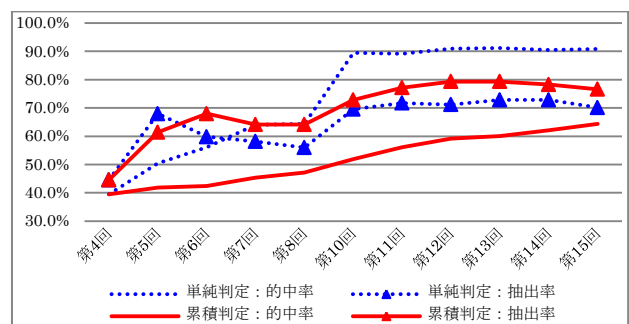


図 3 2013 年度の的中率および抽出率

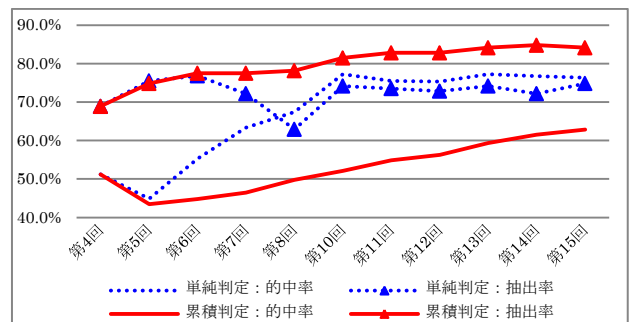


図 4 2014 年度の的中率および抽出率

5. まとめ

本研究では、開発したシステムによって蓄積された定量的な学習データを用い、試験成績の推定と成績不振兆候者の抽出をおこなった。検証の結果、第 5 回の授業という早期の段階でも、成績下位者を 60~70%以上と高い割合で抽出できた。

今後の課題として、提案手法を実際の授業期間中に適用し、その有効性を検証することや、学習データの収集を継続し、より精度の高い推定をおこなうことが挙げられる。

参考文献

[1] 高岡詠子,大沢佑至,吉田淳一:「e-Learning 学習履歴を用いたドロップアウト兆候者早期抽出手法の提案,検証および今後の可能性」情報処理学会論文誌,Vol.52,No.12,pp.3080-3095 (2011)
 [2] 畷川みなみ,宇野健:「プログラミング演習における学習状況のフィードバックシステムの開発」情報処理学会第 76 回全国大会講演論文集,No.4,pp.549-550 (2014)