

プログラミング演習における週単位の 課題進捗状況の可視化に関する基礎的検討

A Basic Study on Visualization of Weekly Assignment Progress in Programming Exercises

石川 健児^{*1}, 香山 瑞恵^{*2}, 新村 正明^{*3}

Kenji ISHIKAWA^{*1}, Mizue KAYAMA^{*2}, Masaaki Niimura^{*3}

^{*1} 信州大学

^{*1}Shinshu University

^{*2} 信州大学大学院

^{*2}Shinshu University Graduate School

^{*3} 信州大学 e-Learning センター

^{*3}Shinshu University e-Learning Center

Email: 18t2013d_ik@seclab.shinshu-u.ac.jp

あらまし：大学の情報系学部では、プログラミング演習授業を実施している組織が多い。本研究では、この学習を支援するために、ソースコードの質的評価に基づく学習者の学習進捗状況可視化システムについて探究している。本稿では、週単位での課題進捗状況の可視化を目的とし、基礎的検討の成果を述べる。

キーワード：プログラミング演習, 学習進捗状況, クラスタリング, 週単位, 課題

1. はじめに

近年, IT・ICT 市場の拡大により, 実践的なプログラミング能力を持った技術者の育成が求められている。特に情報系学部では, 指導者が学習者に課題を与え, 個別にプログラミングを行う形式(プログラミング演習)を採用している。既存の演習支援研究では, ソースコードの質的な評価指標は考慮されにくい。また, 演習中の評価は試みられているが, 週単位で進捗状況を分析する事例は少ない。

2. 研究目的

本研究の目的はソースコードの質的な評価指標⁽¹⁾を用いた。週単位で出題される課題進捗状況の可視化の検討である。ここで質的評価指標とは, 学習者と模範解答のソースコードが一致する部分割合を示す**類似度**と, 学習者コードの中で模範解答コードに含まれない部分割合を示す**相違度**である。

本研究ではこの研究目標を達成するために, 次の2つの研究課題を設ける。

1. 質的評価指標の情報量をできるだけ維持したまま週単位の学習状況を圧縮する方法は何か。
2. 質的評価指標を用いると週単位での学習進捗はどのように可視化されるのか。

なお, 本研究において学生コードは編集中には2秒間隔で個別に自動保存される。その1つ1つのソースコードをスナップショット(以下, SS)と称する。

3. 研究課題 1

2021 年度に実施された週単位の課題 20 種における各個人の SS 数は最少 2[ファイル], 最多 1068[ファイル]であり, SS 毎の時間間隔は最短 2[s], 最長 521337[s]であった。共に変化幅が広い。そこで,

(1)SS を同一ファイル数に標本化(数圧縮), (2)スナップショット間の時間間隔の閾値でまるめる(時間圧縮)ことで, SS 情報を圧縮する。

(1) については, 先行研究⁽²⁾を参考に, SS 数 20 個以上の学生を対象に, 全 SS を時間的に 10 等分するよう圧縮した。

(2) については, 70 分間で解答が求められた期末課題を対象に, 時間圧縮の閾値を検討した。

その結果, SS 間の時間間隔分布において 90%を閾値とした時間圧縮の妥当性を見出した。閾値を設けない場合と 90%を閾値とした場合で, (1)の標本化した結果の質的評価指標の相関を求めた。類似度と相違度それぞれの相関の, 学習者全体の中央値は 0.89 と 0.75 であり, 圧縮後も評価指標の情報量を維持できていると考えた。

4. 研究課題 2

先行研究⁽²⁾において, 期末課題における時間圧縮なしの SS 群での類似度と相違度は 5 クラスタに整理されている。それらに週単位課題において研究課題 1 で整理した圧縮後の SS の類似度と相違度をマッピングすることで, 学習進捗の可視化を検討した。

期末試験の結果より得られたクラスタをそれぞれ, Beginning(B), Indefinite(I), Favorable(F), Warning(W), Outlier(O)と呼称する。週単位の課題20種の内, 内容が類似し, かつ難易度が低くない 3 課題^{*1}について, 5 クラスタに分類される最終 SS 数を求めた。表 1 に結果を示す。3 課題において, 複数クラスタに分類され, 分類数の偏りが少ない課題 B を代表課題とする。

図 1 に代表課題の SS を期末課題で得られたクラスタにマッピングした結果を示す。図中, 薄青点は

*1 これらは C 言語における単方向リストの演習課題である。

各学習者の圧縮後の各 SS, 黒点は各学習者の圧縮後の最終 SS である。

代表課題においては, 期末試験と似た範囲に全 SS がプロットされた. 最終 SS は, **B, I, F, W** にプロットされ, **O** に含まれるものはなかった. 図 2 に, 最終 SS が **F** に含まれた学習者の SS 推移を示す. この学習者は **B-I-F** と推移していた. 一方, 最終 SS が **B** に含まれた学習者の SS 推移を図 3 に示す. この学習者は **B-I-W-I-F-I-B** と推移していた.

5. 考察

研究課題 1 では今回, 時間全体の 90%とした理由は, 短縮前と比較して時間間隔の疎密の情報が保持されており, その後の解析を行えると考えたためである. 時間閾値を 50%とした場合には, 類似度と相違度の相関係数が 0.85 と 0.68 となっていた. そのため, さらなる時間圧縮の可能性は否定できない.

研究課題 2 では, 今回は 20 課題の内の 3 課題のみ取り上げた. 代表課題以外の 2 課題では, 最終 SS のクラスターが **F** と **W** のみであり, **F** に偏って割り当たっていた. **F** の SS は, 相違度が低く類似度が高いため, より模範解答に近いコードであると考えられる. **W** の SS は, 相違度が高く類似度が高いため, 模範解答とは異なる正答コードである可能性がある. これより, 課題 A と課題 C に比べて代表課題の内容が学習者にとって難しかった可能性がある.

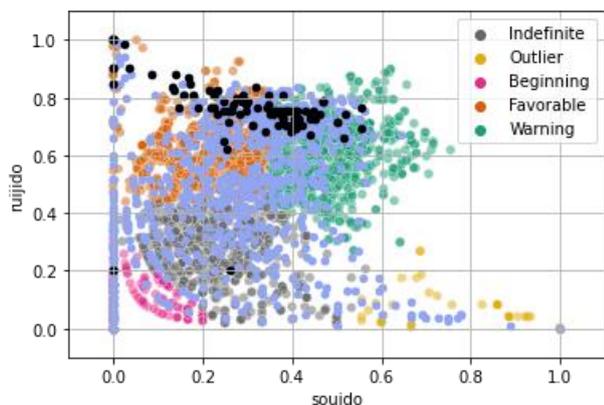


図 1 期末課題クラスターに, 週単位の代表課題の SS をマッピングした結果

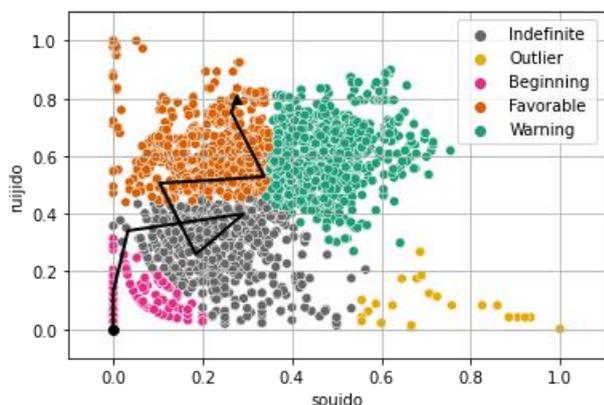


図 2 最終 SS が **F** に含まれた学習者の学習状況

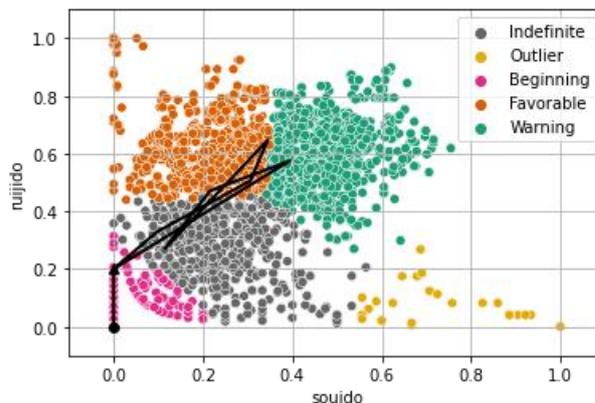


図 3 最終 SS が **B** に含まれた学習者の学習状況

表 1 類似 3 課題の各クラスターの SS 数

| | 課題 A | 課題 B | 課題 C |
|----|------|------|------|
| B | 0 | 1 | 0 |
| I | 0 | 1 | 0 |
| F | 59 | 43 | 65 |
| W | 16 | 63 | 16 |
| O | 0 | 0 | 0 |
| 合計 | 75 | 108 | 81 |

6. まとめ

本研究では, 週単位での課題進捗状況の可視化を目的とした. 研究課題 1 では, SS の圧縮法を検討した. SS 間の時間間隔の上限閾値を時間全体の 90% と設定し, 時間圧縮を行った. また, 課題に取り組んでいた時間を 10 等分し, その等分時間に近い SS を選択することで数圧縮を行った. 研究課題 2 では, 期末試験の SS から得られた 5 クラスターに対して, 週単位課題の圧縮後の SS をマッピングすることで, 課題進捗状態の可視化を試みた. 代表課題を対象に, 学習者の学習推移を具体化した. 一方で, 時間圧縮の閾値をさらに短縮できる可能性も考えられる. また, 代表課題以外の週課題における学習者の推移は, 今回取り上げた学習者とは異なる可能性がある. これらの検討が今後の課題となる.

参考文献

- (1) 田中空来, 香山瑞恵, 新村正明, 館伸幸: プログラミング演習における学習者の行き詰まり同定を目的とした評価指標の妥当性に関する考察, 教育システム情報学会第 47 回全国大会講演論文集, p105-106 (2022).
- (2) 田中空来, 香山瑞恵, 新村正明, 館伸幸: プログラミング演習における学習者の進捗状況の把握を目的とした評価指標に関する研究, 教育システム情報学会 2022 年度第 6 回研究会, pp.1-8 (2023, 発表予定).