5F-05

編集履歴可視化システムを用いたLearning Analytics ～Cプログラミング 科目における編集履歴と評価得点データを統合した分析モデル

後藤 正幸[†]
早稲田大学[†]
経営システム工学科

三川 健太[†]
早稲田大学[†]
経営システム工学科

雲居 玄道[‡]
早稲田大学[‡]
理工学術院総合研究所

小林 学[††]
湘南工科大学[††]
情報工学科

荒本 道隆[‡‡]
アドソル日進(株)[‡‡]
先端IT技術部

平澤 茂一[‡]
早稲田大学[‡]
理工学術院総合研究所

1. はじめに

近年，様々なレベルや分野の教育現場において情報技術を活用とする取り組みが広まっている．このような仕組みは，e ラーニング，反転学習など，様々な新しい教育・学習形態を生み出しており，現在もさらに発展傾向にある．情報技術による援用の意義は，単にそれ自体の機能のみならず，学習者の行動履歴のようなログ情報を蓄積し，教育に活用できるといった利点を生み出している．このように学習者のログを活用し，教育に活用しようとする試みは，ラーニングアナリティクス[1]という研究領域を形成し，様々な研究や取り組みが開始されている．近年では益々重要度が高まっているプログラミング技術やアルゴリズム構築に関する情報教育分野においても，学習者のログを集積し，教育に生かすことができれば，多大なメリットが享受できる可能性がある．

以上の背景のもと，著者らの研究グループでは，機械学習などの統計情報分野の専門家と情報技術の実務家が協働し，プログラミング科目で利用可能な編集履歴可視化システムを開発した[2]．本報告では，大学の正規演習科目においてこの編集履歴可視化システムを利用し，学習者のプログラミングプロセスの詳細ログを活用すると共に，各回の小テストとレポート点の分析を機械学習分野で良く用いられる潜在クラスモデルを使って分析し，プログラミング履歴データとの紐付けを行った事例に付いて報告する．この実証研究を通じ，詳細ログの活用と機械学習手法の可能性について考察する．

Learning Analytics via Visualization System of Edit Record ～ Analytics Model Based on Edit Record and Evaluation Score Data for C-Programming Courses

[†] Masayuki Goto, Kenta Mikawa, Gendo Kumoi, Shigeichi Hirasawa: Waseda University
[††] Manabu Kobayashi: Shonan Institute of Technology
[‡‡] Michitaka Aramoto: Ad-Sol Nissin Corp.

2. 授業形態とデータ取得の仕組み

本研究で対象とした演習授業は，理工系の大学 1 年生を対象とした情報処理の基礎的な演習科目であり，計 10 回の演習授業によって，一通りの C プログラミングの技術を学ぶものである．内容的には基本的な文法などのプログラムの知識だけでなく，再帰構造やポインタを駆使した高度なアルゴリズムまでを含む内容となっている．各回の演習授業では，毎回冒頭に筆記の小テストを実施し，前回までの演習内容の理解度確認のチェックを行うと共に，各回の演習内容理解を促すため，3 題程度のプログラムを作成し考察させるレポート課題を課している．これらの小テストとレポートは，それぞれ計 9 回ずつ実施しており，教員によって各回 30 点満点で得点化されている．加えて，本演習授業の後半回では，編集履歴可視化システムを活用し，このシステムを用いたプログラミング小テストを実施したり，授業中の演習課題の達成確認のために，このシステムにプログラムを入力させて実行させるという演習を行った．履修人数は 139 名であり，2 教室同時進行で演習を実施した．

3. 小テストとレポート点に対する潜在クラスモデル分析

本稿ではまず，複数の小テストとレポートの得点から，学生をクラスタリングするための手法として，潜在クラスモデルを新たに提案し，活用する．いま，M回の小テスト得点を集めたベクトル $\boldsymbol{x} = (x_1, x_2, \cdots, x_M)$ とM回のレポート点を集めたベクトル $\boldsymbol{y} = (y_1, y_2, \cdots, y_M)$ に対し，

$$P(\boldsymbol{x}, \boldsymbol{y}) = \sum_{z} P(\boldsymbol{x}|z) P(\boldsymbol{y}|z)\, P(z)$$

という確率モデルを仮定する．ただし，z は離散値の潜在クラスを表し，$P(z)$ は多項分布，$P(\boldsymbol{x}|z)$ と $P(\boldsymbol{y}|z)$ は無相関M次元正規分布を仮

定する．これらの分布のパラメータは，EM アルゴリズムによって推定することができる．

以下に，小テスト 9 回，レポート 9 回のそれぞれに対し，30 点満点で評価した結果を，先の潜在クラスモデルの潜在クラス数 4 で学習させた結果を示す．なお，学生全体での平均得点は，小テストが 15.00 点，レポートが 18.73 点であった．

表 1．潜在クラスモデルによるクラスタリング

	クラス1	クラス2	クラス3	クラス4
人数	46	29	5	59
小テスト平均点	13.12	19.25	12.61	14.61
レポート平均点	17.36	17.11	20.30	20.38

クラス 1 は「小テストが平均よりやや劣るグループ」，クラス 4 は「レポートの点が良いグループ」であり，これらが大きな集団を構成している．クラス 3 は，クラス 4 に傾向が似ているが，それよりはやや下位に来るグループ，クラス 2 は「小テストの出来は良いにも関わらず，レポート点がさほどでもないグループ」である．

4. 詳細ログを利用したプログラミングプロセスの分析

本演習講義では，開発した編集履歴可視化システムを利用して，学生が「どのようにプログラムを構築していくか」の過程であるプログラミングプロセスのログデータを取得した．過去のプログラムのコピー操作などが行われないように，プログラミング試験の形式により監視下で操作を行わせた．後半の演習授業においてシステムを活用したが，ここでは下記の 3 問についての詳細ログを分析してみることにする．

・Q1：過去の小テストで実施した do… while ループと if, else if 分岐を駆使するプログラム
・Q2：直前の小テストで出題し，模範解答例について解説もした，2 つの整数の最大公約数を再帰的関数定義で解く問題
・Q3：配列とその長さを受け取り，中身を降順に並び替えるソートの関数を作る問題

これらの問題に対して編集履歴可視化システム上で回答させた結果は，個々の学生が“保存し実行”コマンドを押し，プログラムをコンパイルして実行する度毎に記録される．本研究では，それらの行動を以下のように分類し，個々の学生のプログラミングプロセスを追うこととした．

1) 入力： コードを初めから入力し，実行する
2) 修正： プログラム構造を変更し，実行する
3) 微修正： 軽微な変更をし，実行する
4) pnt 挿入： 途中経過把握のための printf 文を挿入し，実行する
5) 削除： プログラムの一部を大幅削除して，実行する
6) 実行のみ： プログラムを全く書き換えずに，再度実行をする

これらの定義に基づき，各学生のプログラミングプロセスについて，統計量を計算し，考察を行う．その結果の一部を表 2 に示す．

表 2．プログラミングプロセスの統計量

	クラス1	クラス2	クラス3	クラス4
全体回数	14.61	18.88	20.80	16.62
微修正回数	2.52	3.76	5.80	4.24
実行のみ回数	3.98	5.45	9.80	6.63

この結果，レポートの得点は高いが，小テストの得点が低い学生のグループであるクラス 3 は，「実行のみ」「微修正」を他のグループよりも多く行っていることがわかる．何らかの操作を行っている“全体回数”の点から見ても，クラス 3 が最も多い．このタイプの学生は，自宅に帰ってゆっくりと PC を使って取り組むことができるレポート課題では力を出せるが，筆記の小テストでは PC が使えないため，力が出せないようである．頭で考えるよりも先に PC でデバックして誤りを発見している可能性があり，学生のプログラミングの仕方に適切な指導が必要であることも示唆される．

5. まとめと今後の課題

本稿では，C 言語のプログラミング演習科目を事例とし，小テストとレポート点の潜在クラスモデル分析手法を提案すると共に，編集履歴可視化システムを用いた学生のプログラミングプロセスの分析の方法について検討を行った．学生のプログラミングプロセスについては，保存・実行の際のソースコードや実行結果の全ログが残っているが，非常に自由度の高いテキストデータであるため，その履歴情報から，学生間のプログラミングプロセスの特徴を表す空間に射影するための前処理が大変重要である．その特徴量の生成法については，さらに様々な観点や切り口があるため，検討の余地がある．

参考文献

[1] Johann A. L., Brandon W. (Eds.), *Learning Analytics From Research to Practice*, Springer (2014).
[2] 荒本，小林，中澤，中野，後藤，平澤：“編集履歴可視化システムを用いた Learning Analytics ～システム構成と実装”，情報処理学会第 78 回全国大会講演論文集，(2016)