

プログラミング演習における TA のための学生行動マイニング

Student's actions mining for teacher assistants in programming exercises

加藤 利康

Toshiyasu KATO

日本工業大学情報工学科

Nippon Institute of Technology

Email: katoto@nit.ac.jp

あらまし: プログラミング演習の授業では課題解決に時間がかかる学生が多く存在しているため TA がサポートしていることが多い。しかしながらプログラミング行動は学生ごとに異なっているため、机間巡視や課題の提出確認だけでは行動の特徴を把握することが困難である。そこで本研究は、TA に行動の特徴を提示し学習指導の手がかりを得られるようにすることを目的として、コンパイルと実行の回数やコンパイルと実行の平均間隔、および課題解決までのエラー数や同じ種類のエラー数などの行動履歴にデータマイニングを行った。この結果とプログラミングの仕方についての回答に相関があったことから、課題解決の要因として、プログラミングの仕方を指導することで課題解決が早くなる仮説を得た。

キーワード: プログラミング演習, TA 指導支援, 行動分析

1. はじめに

高等教育機関におけるプログラミング演習では、学生が個別に課題に取り組むため、順調に課題を進めている学生と比較して遅延や停滞している学生が多く存在している。そのため、授業には TA (Teaching Assistant) がサポートしていることが多く、指導者には、学生からの質問に対応するだけでなく、課題解決に時間がかかっている学生を把握して、その状況に応じた学習の支援を行うことが望まれている⁽³⁾⁽⁴⁾。しかしながら、プログラミング行動（プログラミングスタイルを含む）は学生ごとに異なっているため、エラーの対処以外に TA は何を指導すればいいのかわからないことがある。

そこで本研究は、TA に学生の行動の特徴を提示し指導の手がかりを得られるようにすることを目的として、学生行動マイニングを行う。学生行動マイニングは、課題解決が遅い学生のプログラミング行動の特徴を提供する。課題解決の要因には、プログラミングの仕方、プログラミング言語の文法、コンパイラの動作など、多くのプログラミング知識とさまざまなプログラミングスキルが関連している。本研究では、プログラミング演習で利用されている演習支援システム⁽¹⁾でリアルタイムに学生の行動データを収集できることから、プログラミングの仕方について着目する。

2. 学生行動マイニング

学生行動マイニングの目的は、TA がどのような特性の学生に対して、どのようなプログラミングの仕方を指導すればよいかを把握できるようになることである。

本研究では学生の特性を推定するため、演習支援システムで集計している行動内容を分析し、学生主観のプログラミングの仕方に対する回答との関連性

を調べる。ここでの行動内容は、課題解決時間(秒)、コンパイルの平均間隔、実行の平均間隔、コンパイル回数、実行回数、課題解決までのエラー数、最も多い種類のエラー数である。プログラミングの仕方については、授業担当教員が授業で教えているプログラミングの基本事項 20 項目である。基本事項の内容は、本稿の制限により割愛する。

2.1 行動特徴分類のためのクラスタ分析

プログラミング演習の授業 1 回において出題された課題のうち、全員（個体数 80）が提出した 2 つの課題に対する行動内容（特徴変数 7）を対象にクラスタ分析を行なった。その結果を図 1 と図 2 に示す。

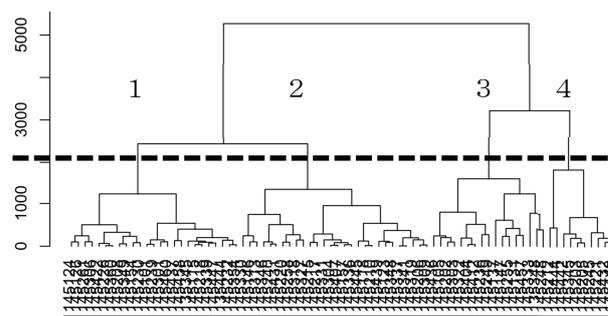


図 1 ward 法によるクラスタ分析

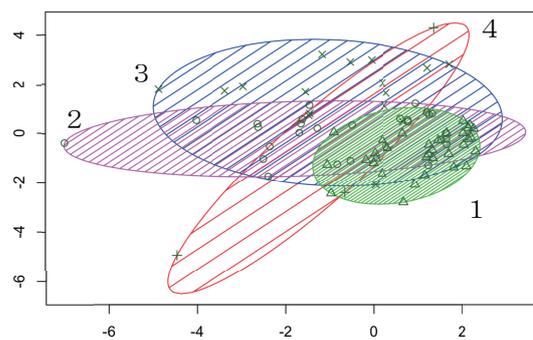


図 2 k-means 法によるクラスタ分析

分析にはクラスタ分析で主に使われている⁽²⁾階層的クラスタリングの ward 法と非階層的クラスタリングの k-means 法を用いた。k-means 法では、分類結果が初期値に依存されるため、初期値はクラスタリング結果の評価尺度基準 Pseudo-F の最も良い結果を用いた。Pseudo-F は、クラスタ同士は疎、クラスタ内は密になっていることを評価する指標である⁽⁵⁾。

2.2 プログラミングの仕方との関連性

学生の行動内容とプログラミングの仕方(以下 PG 仕方)について関連性を調べるため、クラスタ分析の結果に PG 仕方の回答集計を合わせた結果を表 1 と表 2 に示す。

表 1 と表 2 の数値はクラスタごとの個体の平均値であり、人数は個体数である。PG 仕方の数値は、「かっこの対応」や「インデント」などの 20 項目について、している (+1)・していない (-1) の 4 段階評価の回答を点数化したものである。また、課題解決時間と PG の仕方について相関分析を行なった結果、20 項目中 18 項目が正の相関 (0.24) を示した。

表 1 図 1 の点線上階層の内訳

| | 解決時間 | Com. 間隔 | 実行 間隔 | Com. 回数 | 実行 回数 | e 数 | 同じ e 数 | PG 仕方 | 人数 |
|---|------|---------|-------|---------|-------|-----|--------|-------|----|
| 1 | 528 | 209 | 267 | 7 | 5 | 1 | 1 | 10 | 25 |
| 2 | 812 | 379 | 606 | 6 | 3 | 3 | 3 | 6 | 28 |
| 3 | 1405 | 590 | 1160 | 7 | 3 | 4 | 3 | 5 | 17 |
| 4 | 1506 | 228 | 337 | 18 | 13 | 5 | 4 | 4 | 10 |

表 2 図 2 のクラスタ内訳

| | 解決時間 | Com. 間隔 | 実行 間隔 | Com. 回数 | 実行 回数 | e 数 | 同じ e 数 | PG 仕方 | 人数 |
|---|------|---------|-------|---------|-------|-----|--------|-------|----|
| 1 | 537 | 275 | 333 | 6 | 4 | 1 | 1 | 10 | 33 |
| 2 | 1007 | 321 | 568 | 10 | 5 | 5 | 4 | 5 | 28 |
| 3 | 1410 | 521 | 1119 | 8 | 4 | 4 | 4 | 6 | 16 |
| 4 | 2137 | 588 | 646 | 2 | 1 | 1 | 1 | 2 | 3 |

2.3 行動特徴分類の考察

図 1 のクラスタ分析について、デンドログラムの中間で線を引くと 4 つに分割される。図 2 のクラスタは 4 つあり、表 1 と表 2 のクラスタ内訳が近似していることから、合わせて行動の特徴を考察する。

1. 課題解決時間が短く、PG 仕方の数値が高い。コンパイルと実行の平均間隔が短く、コンパイル回数も少ないことから、エラー内容を理解しプログラミングしていると考えられる。
2. 課題解決時間は 3 より短い、PG 仕方の数値が低い。課題解決までのエラー数と同じ種類のエラー数が多く、3 と比べてコンパイルと実行の平均間隔が短いことから、3 より入力が速いだけでエラー内容を理解せずにプログラミングしていると考えられる。
3. 課題解決時間が長く、PG 仕方の数値が低い。同じコンパイルエラーを繰り返し発生させていることから、エラー内容を理解せずにプログラムを直してはコンパイルを行なっていると

考えられる。

4. 課題解決時間が長く、PG 仕方の数値が低い。表 1 では、コンパイル回数と実行回数が多く、その間隔が短いことから、トライアンドエラーをひたすら行なっていると考えられる。表 2 では、コンパイル回数と実行回数が少なく、エラーを発生させずに課題を提出していることから、演習支援システムを使わずにプログラミングをしているか、友人からソースを得ていることが考えられる。

これらの考察から、プログラミングの仕方ができるようになると課題解決が早くなることから、行動特徴の原因は、プログラミングの仕方の違いにある。

3. TA の指導による評価実験

第 2 章の考察から、仮説として行動特徴の原因がプログラミングの仕方に対するプログラミング知識の理解度の違いにあるとし、TA が指導するプログラミング知識の理解度不足が行動の特徴と対応していることを検証する。

検証方法は、TA の指導前後において、実験群と統制群の課題解決時間を定量評価する。群分けについては、過去 3 回のプログラミング演習の受講生 80 名のうち、課題提出の遅い半数について、ランダムに実験群と統制群を各 20 名とする。

4. おわりに

本論文では、プログラミング演習における TA 指導支援として、学生のプログラミング行動をマイニングし、プログラミングの仕方との関係について分析を行なった。その結果、行動の特徴とプログラミングの仕方には関係があることから、演習の課題解決の要因として、プログラミングの仕方を指導することで課題解決が早くなる仮説を得た。

今後は、仮説を実証するため TA の指導による評価実験を行う。

参考文献

- (1) 加藤 利康, 石川 孝: “プログラミング演習のための授業支援システムにおける学習状況把握機能の実現”, 情報処理学会論文誌, Vol.55, No.8, pp.1918-1930, (2014)
- (2) 神島 敏弘: “データマイニング分野のクラスタリング手法(1) - クラスタリングを使ってみよう! -”, 人工知能学会誌, Vol.18, No.1, pp.59-65 (2003)
- (3) 匂坂 智子, 渡辺 成良: “プログラミング初学者の学習方略と段階的理解度に関する調査および支援ルールの作成について”, 教育システム情報学会誌, Vol.26, No.1, pp.5-15 (2009)
- (4) 安田 光, 井上 亮文, 市村 哲: “学生とティーチングアシスタント間でトラブル解決過程を共有できるプログラミング演習支援システム”, 情報処理学会論文誌, Vol.53, No.1, pp.81-89 (2012)
- (5) Calinski, T., and J. Harabasz: “A dendrite method for cluster analysis”, Communications in Statistics, Vol.3, pp.1-27 (1974)