

数学オンラインテストにおける解答の推移の三次元可視化と分析

3D visualization and analysis of transition of answers to an online mathematics test

高田 知樹^{*1}, 中村 泰之^{*2}

Tomoki TAKADA^{*1}, Yasuyuki NAKAMURA^{*2}

^{*1}名古屋大学情報学部

^{*1}School of Informatics, Nagoya University

^{*2}名古屋大学大学院情報学研究科

^{*2}Graduate School of Informatics, Nagoya University

Email: takada.tomoki.j7@s.mail.nagoya-u.ac.jp

あらまし：本研究では、数式自動採点システム STACK を用いて行われた講義における解答データをもとに解答の推移の三次元可視化及びその自動化と、それに基づく分析を行った。この講義において学生は、正答に至るまで何度でも再解答できるようになっている。そこで、STACK のポテンシャル・レスポンス・ツリーという機能を用いて、学生の誤答をその原因ごとに分類することで、学生一人一人について、正答に至るまでにどのように解答が推移したのか、その時間変化の様子を三次元の折れ線グラフに可視化した。このグラフから、学生らがどのような誤答を経由して最終的な解答に辿り着いたのかや、「誤答 A をした学生は誤答 B に至りやすい」などそれぞれの誤答の関連性などの情報が得られる。また、解答が正答に近付いたり遠ざかったりする様子から「悩み度」という指標を設定し、最初の解答での正答率との相関関係を分析することで問題の難易度推定や特徴の抽出等に役立てられるのではないかと考えた。

キーワード：数式自動採点システム、STACK、三次元可視化

1. はじめに

ここ数年、新型コロナウイルスの感染拡大の影響もあり、オンライン学習やオンラインテストの需要は大きく上昇している。

従来のオンラインテストの解答形式に多く採用されていた多肢選択式や正誤判定式では、当てずっぽうによる正答を防ぐことができず、学生の能力に対する正しい評価が難しかった。数学のオンラインテストシステムに着目した際に、数式で解答させる解答形式を採用することで、学生の解答に適切な評価を与えることを可能としているのが、数式自動採点システムの STACK⁽¹⁾ である。また、STACK には、入力された解答を分類し解答の学生に適切なフィードバックを与えることができるポテンシャル・レスポンス・ツリー⁽²⁾ という機能が存在する。そこで今回の研究では、ポテンシャル・レスポンス・ツリーによる誤答の分類結果と学生の解答の時間推移に着目し、三次元の折れ線グラフへの可視化を行い、誤答同士の関係性に重きを置いた分析を行った。また、新たに指標を設定し、各問題に対して、新たな観点から特徴等の分析を行った。

2. STACK について

STACK は英国バーミンガム大学の Christopher Sangwin 氏らが開発した、数学のオンラインテストシステムであり、数式処理システムとしてオープンソースである Maxima を利用することでその正誤評価を行うことが可能となっている。また解答に対して、あらかじめ想定される解答を設定しておくこと

で、その誤答に対するフィードバックを表示する機能をポテンシャル・レスポンス・ツリーという。このポテンシャル・レスポンス・ツリーの各ノードにおいて、入力された解答があらかじめ設定した評価基準と一致しているかどうか判定を行うことで、それぞれの誤答に対して最適なフィードバックを行うことを可能にしている。

3. 解答過程の可視化

3.1 データの内容

分析の対象とするデータは、数学Ⅲ程度の微積分の問題に対する学生の解答データである。STACK では解答ごとに解答内容や提出時間などの情報が記録されている。

3.2 可視化方法

誰が・どのような順番で・どのような解答をしたのかを直感的に理解し分析をスムーズに行うため、三次元折れ線グラフに可視化した。また、選択したデータを瞬時に可視化するプログラムを作成、可視化過程を自動化した。

作成したグラフにおいて、Attempt 軸は「その問題における何回目の解答か」、Students 軸は「解答者の番号」、Type of Error 軸は「ポテンシャル・レスポンス・ツリーによる誤答の分類結果」を示している。Type of Error の値は、最小値が 0 (正答)、最大値が 6 (原因不明の誤答)となっており、この値が大きくなるほど正答から遠ざかっており、理解度が低いことによる誤答であるということを示している。

4. 分析方法

作成したグラフに基づいて、誤答の推移に着目した分析を行う。グラフにおいて、解答が推移していく中で、正答から遠ざかった、すなわち Type of Error の値が増加した一人当たりの平均回数を新たに「悩み度」という指標として設定し、問題ごとに算出する。また、問題の難易度を示す指標となる正答率を算出する。ただし、STACK では正答に至るまで複数回解答を提出できるため、今回は「最初の解答での正答率」を問題ごとに算出し、悩み度との相関関係を調べる。

5. 分析結果

5.1 グラフの分析

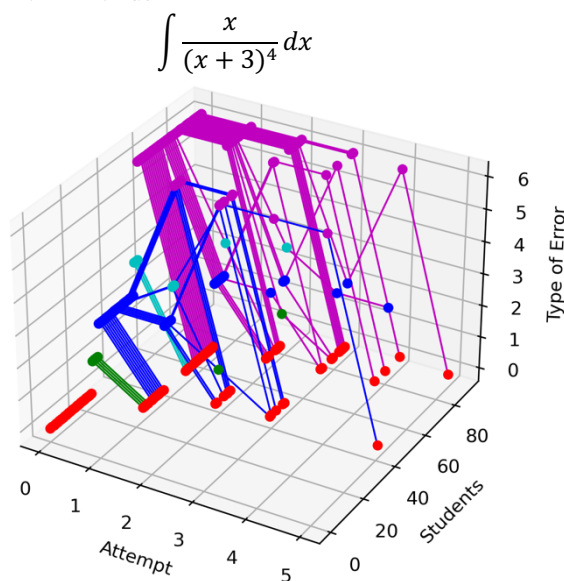


図1：解答過程を表した三次元グラフ

- 0 (赤)：正答
- 1 (緑)：積分定数忘れ
- 2 (青)：部分積分の原始関数のミス
- 3 (水)：合成関数の微分の逆数のかけ忘れ
- 6 (紫)：原因不明の誤答

不定積分の問題であるため、数人の学生に積分定数の入力忘れによる誤答が見られた。これは単純なケアレスミスであるため、この誤答をした全員がその次の解答で正答している。また、部分積分のミスによる誤答が多く、また、その誤答の後にすぐに正答にたどり着いているのではなく、解答数が多くなり、他の誤答を経由している学生が多くみられるため、ただのケアレスミスではなく部分積分の数学的理解が足りていない・すっかり解き方を忘れてしまっている学生が多いと考えられる。

5.2 悩み度と最初の解答での正答率

図2は、悩み度と正答率に関する、問題の分布を表したものである。青い点線は近似曲線となっている。この近似曲線から逸脱している2つの問題5-9と7-1について分析を行う。

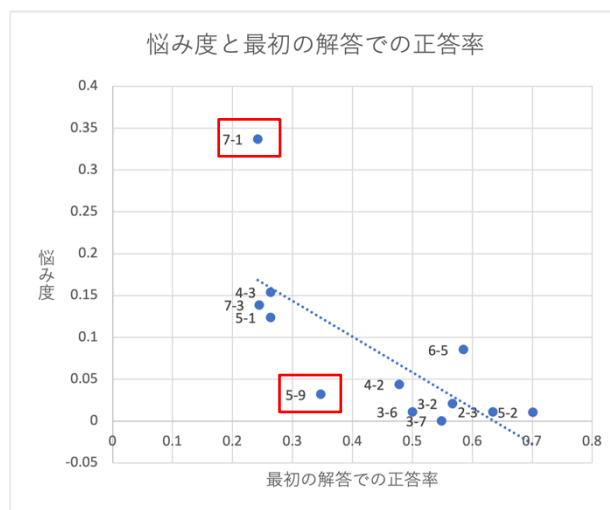


図2：問題の悩み度と正答率の関係

問題5-9は、悩み度、正答率は共に低くなっている。これは、比較的難易度は低く、迷った学生は少なかったが、積分の問題群の1問目だったために、積分定数を忘れるというケアレスミスによる誤答が他の問題よりも多くなり、正答率が低くなったことが原因であると考えられる。

問題7-1は、指定された方法以外で入力してしまったことによる、オンラインテスト特有の誤答である数式入力ミス⁽³⁾の割合が非常に高いことが原因で、悩み度が圧倒的に高くなってしまっていると考えられる。この数式入力ミスは、ポテンシャル・レスポンス・ツリーの作成者も予測できないものが多く、適切なフィードバックを送ることが難しくなっている。

6. まとめと考察

誤答分析において、数式オンラインテスト特有の数式入力ミスが多く見られた。数式入力ミスは、学生の数学の能力を正確に測る上でノイズとなるため、この改善が、今後オンラインテストがさらに普及していくための課題であると考えた。また、各問題で算出した、最初の解答での正答率と「悩み度」には強い負の相関関係があり、それを分布図に起こすことでケアレスミスや、数式入力ミスが発生しやすい問題を発見することができた。新たに設定した指標「悩み度」は、平均解答回数や学生の能力値等、他の指標との相関関係を調べることで、様々な活用方法があるのではないかと考える。

参考文献

- (1) STACK | The University of Edinburgh
"https://www.ed.ac.uk/math/stack", 参照 2023-2-6
- (2) 中村泰之: “数学eラーニング 数式解答評価システム STACK と Moodle による理工系教育”, 東京電機大学出版局, (2010)
- (3) 樋口三郎: “従来型と数学オンラインテストによる事前事後テスト結果の数式入力ミスを考慮した分析 (数学ソフトウェアとその効果的教育利用に関する研究)”, 京都大学数理解析研究所, 数理解析研究所講義録, 第2067巻, pp.152-155 (2018)