

生徒の回答から物理的概念理解を推定・可視化する 個別最適な学び支援ツールの開発

松田竜弥†
東京学芸大学†

加藤直樹‡
東京学芸大学‡

1 はじめに

1.1 物理教育の現状と課題

物理教育,特に力学領域において,生徒が持つ経験則に基づく素朴な概念と,物理法則との間に不一致がしばしば見られる.この不一致を解消しない限り,力学の理解は難しいとされている.[1]本稿では「概念」という言葉を多用するが,共通する性質を抽出し普遍化した知識としてのあくまで一般的な概念という意味で用いる.特に物理法則と整合しない素朴な概念は誤概念と呼ばれる.この誤概念の分析には FCI (Force Concept Inventory) という概念調査紙を用いる方法がある.これは多肢選択問題を通じて生徒の誤概念を明らかにし,教育効果を定量的に評価するツールである.

また,PI (ピアインストラクション) のようなアクティブラーニングを採用した授業手法では,

- 生徒が個別に概念理解を問う選択式問題に回答する.
- 3~4 人のグループで議論し,もう一度,回答する.
- クラス全体で回答の変化を共有する.

しかし,このアプローチでは,選択式のため,概念に対する深い理解が図れないことや個々の生徒の概念理解状況を教師が把握しきれないという課題がある.[2]

1.2 先行研究

西村の研究[3]によると,形成的評価を頻繁に使用する生徒群は FCI や小テストでの概念理解が向上する傾向がある.しかし,教師の業務的負担は大きい.そのため,FCI に基づく物理的概念理解状況をリアルタイムで自動化・可視化することが重要だと考える.そのことにより,形成的評価の容易な実施を図ることができる.

1.3 研究の目的

本稿では,先述の課題を解決するため,生徒の筆記回答から FCI に基づく物理的理解状況を推定・可視化し,教師が適切な支援を行うことができる,個別最適な学び支援ツールを提案する.

2 個別最適な学び支援ツールの設計

2.1 基本設計

本ツールの目的は,物理のコンセプトテストにおける各生徒の筆記回答から,その物理的概念の正誤を判定し,分類結果から誤概念を特定することである.これにより生徒の理解状況を深く洞察する.また,その個別な状況を,教師が授業中に認識できるように可視化する.このことにより,個別的な教育的介入のための基盤を提供し,指導の個別化の実現を目指す.

2.2 生徒の回答分類アルゴリズムの設計

2.2.1 基本原理とアプローチ

問題文に基づいて事前に作成した選択肢を使用し,生徒の筆記回答を最も適切な選択肢に分類する方法を提案する.正誤のラベル付けをするとともに,物理的誤概念を元に各選択肢を生成する.このことにより,回答した生徒の物理的概念理解状況を把握することが可能となる.

選択肢作成と分類には生成 AI を使用する.生成 AI とは大規模なデータセットで訓練された自然言語処理モデルであり,複雑な言語問題に対応する能力を有している.先の筆記回答がどの選択肢に近いかは,生徒の回答の文脈を深く理解し,複雑な言語的問題に対応する能力を持つ生成 AI が適している.生成 AI は,単純な自然言語処理モデルに比べて文脈全体の意味を把握し,高度な言語理解を行う能力を有しているため,生徒の記述から物理的概念理解の正確な推定が可能となる.

2.2.2 分類プロセス

生成 AI を用い,プロンプトエンジニアリングにより分類する.

ただし,コオアルゴリズムの適応には限界があり,全ての回答を完璧に分類することは不可能である.特に生徒の回答が,非常に独特な回答や,曖昧な表現を含む場合には,分類が難しい.

2.3 インターフェイスの設計

本ツールでは,生徒の回答とそれに基づく分類結果,次の学習への提案を座席表に紐付け,視覚的に表現する.また,分類結果ごとに生徒の座席を色分けして表示し,回答,分類結果,次の学習への提案は座席をクリックするとドロップダウンメニューとして出現するようにした.このインターフェイスによって

Development of an individually optimal learning support tool that estimates and visualizes understanding of physical concepts from students' answers

†Matsuda Tatsuya, Tokyo Gakugei University

‡Kato Naoki, Tokyo Gakugei University

教師は PI での議論中の生徒同士の議論傾向を予測し、特定のグループへ集中的な援助の提供やファシリテートなど、教育的な意思決定を行うことができる。(図1)

2.4 生徒の記述支援機能

先述した通り、分類アルゴリズムの精度には限界がある。そこで生徒の筆記内容の文章的完成度を向上させるために、生徒の回答作成を補助し、理由不足や論理矛盾などを添削する機能を提供する。

選択式テキストと違い、筆記式では思考の手がかりが少ない。この問題を解決するため、チャット形式で生徒が AI アシスタントと対話をしながら回答を作成することができる機能を開発した。これにより、理由不足、論理矛盾、不明瞭な表現を指摘し、生徒がよりクリアな回答を作成できるようになる。この機能は生徒の思考を促進し、深い理解を支援する。また、問題の答えは絶対に生徒に教えないよう、分岐とプロンプトエンジニアリングを用いて設計した。(図2)

4 ツールの試作

実装言語はバックエンドでは Python を使用し、ChatGPT-4 の API を活用している。また、データベースは SQLite、フロントエンドの動的な座席表では JavaScript を使用し、Web アプリケーション全体では Flask フレームワークを用いて構築した。教師が授業中に容易に扱うことができるよう、GoogleClassroom の UI 設計様式を採用した。

5 評価実験

5.1 評価方法

アルゴリズムの精度は、FCI に基づいた選択肢式テストの「生徒の選んだ選択肢」とアルゴリズムによる分類結果」との比較により評価する。具体的には筆記回答後に選択肢を与え、生徒がどの選択肢を選ぶかを尋ね、その回答とアルゴリズムによる分類結果の一致率をもとに精度を測定する。東京都私立高等学校(5クラス、166名)で評価実験を行なった。

図3に示す FCI の一問目を評価問題として使用した。

“”同じ大きさの2つの金属球がある。一方の重さは他方の2倍である2階の窓から、2つの金属球を、同時に落とすとき、地面に着くまでの時間はどうか？”

(記述回答後に選ばせる選択肢)

- 1) 重い方が、軽い方の半分くらいの時間で着く。
- 2) 軽い方が、重い方の半分くらいの時間で着く。
- 3) ほぼ同時。
- 4) 重い方が先に着くが、軽い方の半分の時間とは限らない。
- 5) 軽い方が先に着くが、重い方の半分の時間とは限らない。

図3 評価実験の問題

5.1 結果

分類結果の一致率は、約 84.24%となった、また、分類結果は正答選択肢に集中するため、各選択肢のサンプル数を重み付けした加重平均を用いると、その他の評価指標は、適合率: 82.16%, 再現率: 84.24%, F1 スコア: 82.96%となった。

5 おわりに

本稿では、生徒の物理的概念理解を推定・可視化するツールを開発し、教育の質を向上させる新たな方法を提供する。生徒の記述回答から物理的概念を把握し、形成的評価を支援することで、教師による指導の個別化を実現することが可能となる。また、記述支援機能により生徒の文章が整えば、分類精度は向上するだろう。

本ツールの最大の利点は、記述式回答に基づく概念理解の評価と可視化能力にある。しかし、精度には限界があり、特に独特な回答や曖昧な表現の分類が難しい。また、現段階では記述支援機能は未評価でありその効果は今後の研究で検証する必要がある。

参考文献

- [1]新田英雄：素朴概念の分類, 物理教育, vol.60, No. 1, pp.17-22 (2012)
- [2]後藤敬祐：ピア・インストラクションにおける生徒間相互作用の分析, 物理教育, vol.67, No. 4, pp.227-234 (2019)
- [3]西村 聖太：形成的評価の活用は概念理解と学習姿勢を向上させるか, 物理教育, vol.66, No. 4, pp.243-248 (2018)



図1 教師画面：座席インターフェイ



図2 生徒画面：記述支援機能