

# 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<sup>†</sup>加藤直樹<sup>‡</sup>東京学芸大学大学院<sup>†</sup>東京学芸大学<sup>‡</sup>

## 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．また，特に物理法則と整合しない素朴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する．単元の前後で実施することで，授業の教育効果を定量的に評価することができる．

さらに，PI (ピアインストラクション) というアクティブラーニングの手法は，誤概念の訂正と理解の深化を促進するための有効な手段である．生徒が個別に概念理解を問う選択肢式のコンセプトテストに回答し，その後，3～4人のグループで議論し，もう一度，回答する．最後にクラス全体で回答の変化を共有するといったアプローチである．

しかし，これらのアプローチでは解答が選択肢式であるため，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授業中には把握しきれず，個別に適切な形成的評価を行うことができないという課題がある[2]．

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．

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

<sup>†</sup>Tatsuya Matsuda, Graduate School of Education, Tokyo Gakugei University

<sup>‡</sup>Kato Naoki, Tokyo Gakugei University

## 1.2 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．

本ツールは，誤概念の特定，FCI および PI の限界を克服し，形成的評価の実施を目指す．

## 2. 個別最適な学び支援ツールの設計

### 2.1 基本設計

本ツールでは，物理の概念理解を問うコンセプトテストにおける各生徒の筆記回答から，その物理的概念理解の正誤を判定し，誤概念を特定する．このプロセスを通じ，生徒の理解状況をより深く洞察することを支援する．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とで，個別的な教育的介入による，指導の個別化の実現を支援する．

### 2.2 物理的概念理解状況推定アルゴリズム設計

#### 2.2.1 基本原理とアプローチ

ツールの核心部分は，生徒の筆記回答を最も適切に反映する選択肢に分類するアルゴリズムである．これらの選択肢には，それぞれに対応する物理的概念や誤概念が定義されている．このアプローチにより，生徒の物理的概念理解状況を推定することができる．

#### 2.2.2 選択肢の推定

生徒の解答がどの選択肢に当てはまるかの分類には生成 AI を使用する．生成 AI は大規模なデータセットで訓練されており，生徒の筆記解答の文脈を理解し，それを適切な選択肢に分類する能力を持っていると考え，採用する．

#### 2.2.3 選択肢の作成

選択肢の作成にも生成 AI を用いる．問題文から5つの選択肢を生成する．このとき正答には身につけて欲しい物理的概念を，誤答には誤概念に基づいた選択肢となるようにする．また，教師の意図を反映できるように，選択肢・誤概念の修正は自由に行えるようにする．

## 2.3 機能設計

### 2.3.1 座席表による可視化機能

本ツールでは，生徒の筆記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り教師はPIでの議論中の生徒同士の議論傾向を予測し、特定のグループへの集中的な援助やファシリテートなど、教育的な意思決定を行うことが可能となる(図1)。

### 2.3.2 次の学習の提案機能

次の学習の提案を表示する機能を提供する。この機能は、PIでの生徒同士の議論をより活発かつ対話的にするための論点や議論のきっかけを提供する。たとえば、生徒が誤答選択肢を選んだ場合、その誤概念に基づいた考えるべき点や、正答を選んだ生徒には他の条件下や応用的な質問を投げかける。提案は生成AIを用いて、正答選択肢と生徒の回答を比較して作成する。

### 2.3.3 生徒の記述支援機能

選択肢式問題と違い、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ツールと対話をしながら回答を作成できるようにする。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支援する。問題の答えは生徒に教えないようにする。対話文は生成AIを用いて、生徒の回答と問題文、選択肢を参考にして作成する。

## 3. 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。

## 4. 評価実験

### 4.1 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「本ツールの推定結果」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、その回答と本ツールの推定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2に示すFCIの一問目を評価問題として使用した。

### 4.2 結果と考察

アルゴリズムの全体的な精度(一致率)は、

約89.09%となった。また各選択肢に対する、その他の評価指標による評価結果を表1に示す。選択肢1の適合率が低く、選択肢4の再現率が低い原因は、生徒が選択肢を与えられない状態だと、半分の時間という詳細状況まで記述してくれないということが考えられる。選択肢2と5は、選んだ生徒が1人もおらず、アルゴリズムによっても選ばれなかったことがわかる。

### 5. おわりに

本稿では、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、教師が教育的介入を行い、個別最適な学びを支援するツールを提案した。評価実験からは、アルゴリズムの精度の限界が見える。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。

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。今後、アルゴリズムのみならず可視化によるツール全体の有用性を評価していきたい。

### 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] 西村壘太：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ス

“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。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) 分からない

図2 評価実験で用いた問題

表1：選択肢ごとの評価結果

選択肢番号	適合率(Precision)	再現率(Recall)	F1-score
①	0.50	0.86	0.63
②	0.00	0.00	0.00
③	0.99	0.99	0.99
④	0.60	0.19	0.29
⑤	0.00	0.00	0.00