

大規模言語モデルの思考プロセスを用いた授業発話生成

児嶋祥成¹ 大西朔永² 椎名広光² 保森智彦³¹ 岡山理科大学大学院 理工学研究科 ² 岡山理科大学 情報理工学部³ 岡山理科大学 教育学部

r24smd8qv@ous.jp {s-onishi,shiina,yasumori}@ous.ac.jp

概要

本研究は、教員の授業改善支援を目的に、大規模言語モデルの思考プロセスを活用して若手教員と経験度が高い教員の授業発話の差異を比較し、若手教員の発話に対して経験度が高い教員の発話シミュレーションを生成する手法を提案する。具体的には、実際の小学校の教員と児童の授業対話データから教員ごとの教育的特徴の差を明確化し、教育的特徴の差を大規模言語モデルを用いて同じ授業を他の教員で発話するシミュレーションを行う。若手教員の発話と経験度が高い教員の教育的特徴をLLM-as-a-Judgeを用いた自動評価で高い評価を得る傾向を確認した。

1 はじめに

近年、学校現場において教員に求められる能力は高度化している。特に小学校学習指導要領 [1] において、「主体的・対話的で深い学び」の実現が求められている。しかし、日本の教員は多忙である場合が多く [2]、省察活動 [3] の機会を確保することは困難となっており、個々の教員が自身の授業改善にリソースを割きづらい現状が常態化している。そのため、教員の省察活動を効率的かつ効果的に支援する仕組みの構築 [4] が急務となっている。

一方、大規模言語モデル (以下、LLM) では、近年においては回答に至るまでの論理的な思考プロセスを出力可能な推論モデル (Reasoning Model [5]) が登場しており、表面的なテキスト生成だけでなく、文脈に潜む意図や背景を汲み取った高度な推論が可能となりつつある。

実際の小学校の授業は、教員と児童間の相互作用的な対話によって進行する。その中で教員は、児童の表情や発言内容といった行動から学習状況や理解度を推定し、次の発問や支援を決定している。このように高度な判断が求められる対話において、若手

教員が経験不足から適切な対応に苦慮する場面も少なくない。

そこで本研究では、若手教員の授業データに対し、模範となる経験度が高い教員のデータを用いてLLMで教育的特徴の差を出力し、若手教員の発話を経験度が高い教員の特徴を持った発話に置き換えるシミュレーション手法を提案する。授業中の対話データに基づき、若手教員が経験度が高い教員の思考に近づくためにはどのような発話を選択すべきかを生成・提示する。生成された発話が経験度が高い教員の教育的特徴を適切に反映しているかを検証するため、LLM自身を評価者として用いるLLM-as-a-Judge [6] のアプローチを採用し、自動評価を行う。若手教員と経験度が高い教員の思考の差を用い、単なる正解の提示にとどまらない、深い発話分析及びアドバイス生成を行うシステムの構築を目指している。

2 授業対話データの概要

本研究では、実際の小学校における算数の授業風景を映した動画データを用いて、各動画データから人手による文字起こしを行い、教員と児童の次の4種類の発話データセットを作成した。なお、経験度が高い教員のデータ (b)~(d) については、同一人物による授業記録を用いた。

(1) 若手教員の授業

(a) 4年 階段の周りの長さを求める単元

(2) 経験度が高い教員の授業

(b) 5年 余りのある小数の割り算の仕方を考える単元

(c) 5年 正方形の棒の数を求める式を考える単元

(d) 6年 三角形の数を調べる単元

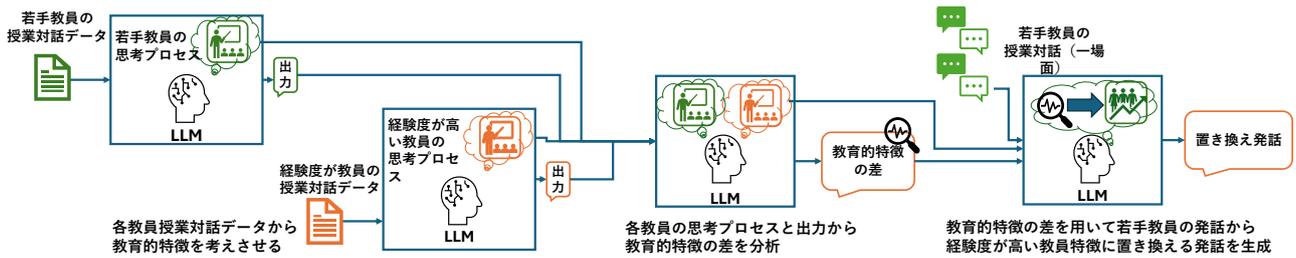


図1 システム概要図

3 システム概要

本システムは、若手教員と経験度が高い教員の授業対話データを入力とし、経験度が高い教員の教育的特徴に基づいた「置き換え発話」を出力する。本システムの全体構成を図1に示す。システムの機能は大きく分けて次の3つのモジュールで構成される。

- (1)教育的特徴の抽出
- (2)教育的特徴の教員間比較
- (3)比較を用いた発話置き換え

教育的特徴の抽出は、各教員の授業対話データを用い、教育的特徴の教員間比較は、教育的特徴の抽出時のLLMの思考プロセスと出力を用いる。比較を用いた発話置き換えは、教育的特徴の教員間比較の出力結果と、若手教員の授業対話データの一場面を用いる。

本研究では、教員の思考プロセスの抽出およびアドバイス生成を行うLLMには、Googleが開発したGemini 3 Pro[7]を採用した。

3.1 教育的特徴の抽出

授業対話データを入力とし、教員の発話意図と教育的特徴を抽出する。LLMを活用し、単に授業対話データの教員の発話内容の特徴を出力させるだけでなく、実際の教員の発話に対して「なぜその発話を選択したのか」という潜在的な意図や、「なぜこの授業の流れに持っていくのか」という背景にある教育的戦略をプロンプトを用いて出力させる。表面的に出力されるテキスト情報だけでは、捉えきれない教員の思考プロセスを含めた特徴抽出を行う。

3.2 教育的特徴の教員間比較

抽出された若手教員と経験度が高い教員の教育的特徴を比較し、両者の差異を明確化する。システム

表1 教育的特徴の比較（一部抜粋）

項目	若手教員	経験度が高い教員
基本スタンス	「学習支援（Scaffolding）」重視。児童が自ら道具（表・式）を使いこなせるよう足場を組む。	「法則の発見・納得」重視。児童が自ら概念を構築し、数学的な見方・考え方を深めるよう導く。
対話のコントロール	短時間の頻繁な交流。「10秒間で相談」など、テンポ良く全員を思考に巻き込む。	「付け加え」と「バトンタッチ」。他者の意見を練り上げ、クラス全体の理解へと統合する。
思考の価値付け	主体性とプロセスを称賛。一生懸命な姿や、児童が自分で決めた戦略を高く評価する。	「泥臭い努力」と「誤答」を肯定。数えまくる努力や間違いを、深い理解への重要な鍵として活用する。

は教育的特徴の抽出の際のLLMの思考プロセスと出力の両方を用いることで、抽出された意図や戦略の差に基づき、若手教員と経験度が高い教員の授業中の発話が教育的効果の観点で当てている焦点の違いや構造化して提示する。単なる優劣の判定ではなく、教育的特徴の違いを生成する。

3.3 比較を用いた発話置き換え

特定された教育的特徴の差異に基づき、発話置き換えを生成する。単に「より良い発話」を生成するのではなく、児童と若手教員の授業中の対話において、経験度が高い教員の特徴に近づける発話を生成するシミュレーションを行う。LLMに教育的特徴の教員間比較の思考プロセスと出力の情報と若手教員の発話から経験度が高い教員の教育的特徴へと発話を近づけるプロンプトで出力させる。若手教員が単なる修正案としてではなく、自身の現在のスキルと経験度が高い教員のスキルとの差を埋めるための具体的なフィードバックとして、代替発話を生成することが可能とすることが目的である。

4 教育的特徴の教員間比較例

本システムを用いて抽出された、若手教員と経験度が高い教員の教育的特徴の比較結果について述

表2 若手教員の発話と経験度が高い教員の特徴に基づく置き換え発話の比較例

場面	若手教員の実発話	経験度が高い教員の特徴に基づく生成発話
誤答・ズレへの対応	1, 2だから, 2センチなんよ。なんか教え方ちょっと。間違ってるから, もう一回教え直してごらん。ちょっとさ, 何個か, みんなでやってみようか?	おっ、『違う』か! その『モヤモヤ』こそ大事なヒントよ。いきなり正解じゃなくていいけえ, どうしてそう見えたんか教えてくれる? みんなも, ○○さんの見方と自分の見方, どこに『ズレ』があるか一緒に考えてみよう。
問題提示	周りの長さを求めよう。このあと, 問題配るので, のりを出してください。ノートに貼りましょう。	その通り, 周りの長さだね。でも, さっき見たあの長い20段を, 全部描いて数えるのは……正直『無理』だよな? だからこそ, みんなが提案してくれた『決まり』を見つけることに価値があるんだ。描かなくても計算一発で求められる『法則』を, 表を使って解き明かしてみよう。よし, 道具(のり)を出して。
思考を深める発問	じゃあ, 1つは頑張ろうか。ねえねえ。みんな。ここって何センチなん? ここって1センチなん?	ちょっとストップ。今, 答えが分かれそうな面白いポイントがあったよ。2段目のこのギザギザした角の部分, ここをどう数えるかで答えが変わるんじゃない? 『6 cm』派と『8 cm』派, それぞれの理由を出し合ってみよう。

べる。抽出された特徴の比較結果の一部を表1に示す。なお, すべての観点における詳細な比較結果については, 巻末の付録Aに示す。

分析の結果, 両者の間には授業の進行や児童への関わり方において, 主に以下の3点における顕著な差異が確認された。

• **基本スタンスの差異:**

若手教員は学習支援を重視し, 児童が躓かないように問題を細分化して提示するアプローチをとる傾向が見られた。経験度が高い教員は法則の発見や納得を優先し, 児童が自ら思考し, 数学的な構造を見つけ出すプロセスを重視している。実際の授業対話データ上でも, 若手教員は児童に対して問題を段階に分けて発問する場面があり, 経験度が高い教員は, あえて児童の直感と反する事例を提示して「認知的な葛藤」を引き起こす場面や, 誤答をすぐに訂正せずクラス全体で議論させる戦略をとっている。

• **認知的葛藤の活用:**

若手教員が正解への誘導を優先する場面が多いのに対し, 経験度が高い教員は, 児童の直感と反する事例を提示して認知的葛藤を引き起こす戦略をとっていた。また, 誤答が生じた際も即座に訂正せず, それをクラス全体での議論の材料として活用し, 深い理解へと導く場面があった。

• **対話の手法**

若手教員が, 教員対指名した児童という「1対1」のやり取りを繰り返す傾向にあるのに対し, 経験度が高い教員は, ある児童の発言を「○○さんはどう思う?」と別の児童に投げかける場面が頻繁に見られた。

5 発話置き換え例

提案手法により生成された発話置き換えの例について述べる。若手教員の授業対話データを基に経験度が高い教員の特徴を反映した置き換え発話を生成した。ここでは, 「誤答・ズレへの対応」「問題提示」「発問」の3つの場面についての例を表2に示す。生成された発話と若手教員の実発話の比較分析結果は以下の通りである。

• **誤答・ズレへの対応**

若手教員は「間違ってるから, もう一回教え直して」と, 誤りを即座に指摘し, 行動の修正を求めている。これに対し, 生成された発話は「その『モヤモヤ』こそ大事なヒントよ」と, 誤答や違和感を肯定的に受け止めている。なぜそう考えたのかを言語化させることで, 経験度が高い教員の教育的特徴である「メタ認知の促進」の意図が反映されていると考えられる。

• **問題提示**

若手教員が「周りの長さを求めよう」と学習課題を直接的に提示しているのに対し, 生成発話では「20段を全部描いて数えるのは『無理』だよな?」と, 解決困難な状況を提示している。これは, あえて困難な状況を示すことで「法則を見つけたい」という学習への「切実感」を創出する意図が反映されたものと考えられる。

• **思考を深める発問**

若手教員が「ここって何センチなん?」と, 事実確認のための発問を行っているのに対し, 生成された発話は「それぞれの理由を出し合ってみよう」と議論を促している。多様な考えを比較検討させることで, 数学的な「構造理解」を深めようとする意図が反映されていると考えら

表3 教育的特徴を満たす発話数の比較

評価項目	元の発話	生成発話
法則の発見・納得重視	5	62
切実感の創出	6	22
付け加えとバトンタッチ	3	19
泥臭い努力と誤答の肯定	5	15
構造的理解のサポート	6	50
自治とメタ認知の促進	2	3

れる。

このように、提案システムは単に表現を丁寧にする等の表面的な修正ではなく、若手教員の発話を経験度が高い教員の教育的戦略に基づいた発話へと転換していることが分かる。

6 特徴に応じた発話生成の評価

本節では、提案システムによって生成された置き換え発話が、経験度が高い教員の教育的特徴をどの程度反映しているかを検証する。評価対象は、若手教員の授業における全発話と生成した置き換え発話とした。これらすべての発話に対して、提案システムを用いて経験度が高い教員の特徴を反映した置き換え発話を生成した。評価の手順として、オリジナルである若手教員の、システムにより生成された置き換え発話のそれぞれについて、評価項目（経験度が高い教員の特徴）を満たしているか否かを判定した。判定には、客観性と再現性を担保するため、評価基準をプロンプトとして与えた LLM (Gemini 3 Pro) による自動判定 (LLM-as-a-Judge) を用いて、評価項目に該当する発話数を集計した。

6.1 発話生成の評価項目

判定の基準となる評価項目は、第4章（および付録）の比較分析により抽出された、経験度が高い教員の教育的特徴とした。LLM には、入力された発話が以下の基準を満たしている場合を「適合」と判定させた。

(1)「法則の発見・納得」重視:

既習事項との対比や矛盾を提示し、「ずれ」や「葛藤」を意図的に作り出すことで、思考せざるを得ない状況を生み出し、数学的な見方・考え方を深めるよう導く。

(2)「切実感」の創出:

100 番目など、さらに膨大な数を設定し、「法則を見つけたい」という強い動機付けを行うこと

で、探究への没入感を高める。

(3)「付け加え」と「バトンタッチ」:

一人の発言に対して他の児童に関連付けさせたり、説明を引き継がせたりすることで、個人の意見をクラス全体の理解へと統合する。

(4)「泥臭い努力」と「誤答」の肯定:

数えまくるような泥臭い努力や、間違いを否定せず、深い理解へ到達するための重要な鍵 (リソース) として活用する。

(5)構造的理解のサポート:

図と式をリンクさせ、「なぜその数で割る/掛けるのか」という式の意味 (本質) を問い、構造的な理解を促す。

(6)自治とメタ認知の促進:

児童自身に時間配分を相談するなど、学習の進め方を委ねることで、学びに向かう集団としての自律を促す。

6.2 発話生成の評価結果

各評価項目について、LLM が「適合する」と判定した発話数を集計した結果を表3に示す。集計結果に基づく考察は以下の通りである。

・数学的本質の重視に関する項目の増加

「法則の発見・納得」重視 (5 件 → 62 件) や、「構造的理解のサポート」(6 件 → 50 件) といった項目において、該当数が大幅に増加した。これは、提案システムが単なる手続き的な処理に終始しがちな若手教員の発話を、数学的な本質理解を問う発話へと再構成できたことを示している。

・自治とメタ認知の促進に関する課題

一方で、「自治とメタ認知の促進」(2 件 → 3 件) については、大きな変化が見られなかった。学習の進め方を児童に委ねるような発話は、単発の発話修正だけでは実現が難しく、授業全体の時間の流れや文脈に強く依存するためであると考えられる。

7 おわりに

本稿では、LLM の思考プロセスを活用した授業発話の生成手法を提案した。今後は、生成されたを実際の教員が評価する被験者実験を行い、システムの有用性を検証する予定である。

謝辞

本研究は JSPS 科研費 JP23K11378 の助成を受けたものです。

参考文献

- [1] 文部科学省. 小学校学習指導要領 (平成 29 年告示) . https://www.mext.go.jp/content/1413522_001.pdf, 2016. 参照 2025-12-23.
- [2] 国立教育政策研究所. 教員環境の国際比較 OECD 国際教員指導環境調査 (TALIS)2018 調査報告書. ぎょうせい, 2018.
- [3] 秋田喜代美. 変貌する教育学, 教師教育から教師の学習過程研究への転回—ミクロ教育実践研究への変貌—, pp. 45–75. 世織書房, 2009.
- [4] 保森智彦. 省察方法のデジタル化に関する一考察: 「主体的・対話的で深い学び」の観点から, pp. 3–11. 学習開発学研究, No. 14. 広島大学大学院人間社会科学研究所科学学習開発学領域, 2022.
- [5] Jason Wei, Xuezhi Wang, Dale Schuurmans, Maarten Bosma, Brian Ichter, Fei Xia, Ed Chi, Quoc Le, and Denny Zhou. Chain-of-thought prompting elicits reasoning in large language models, 2023.
- [6] Dawei Li, Bohan Jiang, Liangjie Huang, Alimohammad Beigi, Chengshuai Zhao, Zhen Tan, Amrita Bhattacharjee, Yuxuan Jiang, Canyu Chen, Tianhao Wu, Kai Shu, Lu Cheng, and Huan Liu. From generation to judgment: Opportunities and challenges of llm-as-a-judge, 2025.
- [7] Google. Gemini 3 pro model card. <https://storage.googleapis.com/deepmind-media/Model-Cards/Gemini-3-Pro-Model-Card.pdf>, 2025. 参照 2026-01-08.

A 教育的特徴の比較結果（詳細）

表4に、本研究で抽出された若手教員と経験度が高い教員の教育的特徴の全比較項目を示す。

表4 若手教員と経験度が高い教員の教育的特徴の比較分析（詳細）

比較項目	若手教員（学習者中心のファシリテーター）	経験度が高い教員（概念構築の伴走者）
基本スタンス	「学習支援(Scaffolding)」重視 児童が自ら道具（表・式）を使いこなせるよう、丁寧な足場かけを行う。具体から抽象へのステップを重視し、手作業（具体）から表（半具体）、式（抽象）へと確実に繋げる。	「法則の発見・納得」重視 既習事項との対比や矛盾を提示し、「ずれ」や「葛藤」を意図的に作り出すことで、思考せざるを得ない状況を生み出し、数学的な見方・考え方を深めるよう導く。
問題提示の工夫	「ジャンプ」の活用 20段、50段など、描くのが困難な数を提示することで、数えることの限界を感じさせ、式の便利さを実感させるアプローチをとる。	「切実感」の創出 100番目など、さらに膨大な数を設定し、「法則を見つけたい」という強い動機付けを行うことで、探究への没入感を高める。
対話のコントロール	短時間の頻繁な交流 「10秒間で相談」など、時間を区切ってペアでの相談を促し、テンポ良く全員を思考に巻き込む工夫が見られる。	「付け加え」と「バトンタッチ」 一人の発言に対して他の児童に関連付けさせたり、説明を引き継がせたりすることで、個人の意見をクラス全体の理解へと統合する。
思考の価値付け	主体性とプロセスの称賛 児童の一生懸命な姿や、自分で決めた戦略を高く評価する。高エネルギーかつ柔軟に、ユーモアを交えて自治を促す。	「泥臭い努力」と「誤答」の肯定 数えまくるような泥臭い努力や、間違いを否定せず、深い理解へ到達するための重要な鍵（リソース）として活用する。
数学的概念への接続	関数的な見方の育成 表を縦・横に分析させ、二つの数量の関係を式に表すプロセスを丁寧に追うことで、関数的な見方を育成する。	構造的理解のサポート 図と式をリンクさせ、「なぜその数で割る/掛けるのか」という式の意味（本質）を問い、構造的な理解を促す。
教室マネジメント	心理的安全性の確保 親しみやすい言葉（方言等）を用い、全員参加を促しつつ発言への心理的ハードルを下げる工夫を行っている。	自治とメタ認知の促進 児童自身に時間配分を相談するなど、学習の進め方を委ねることで、学びに向かう集団としての自律を促す。