

# 対話モデルを用いた授業の発話分析とシーンの可視化

大西 朔永<sup>\*1</sup>, 椎名 広光<sup>\*2</sup>, 保森 智彦<sup>\*2</sup>

<sup>\*1</sup>岡山理科大学大学院, <sup>\*2</sup>岡山理科大学

## Utterance Analysis and Scene Visualization in Class using a Conversation Model

Sakuei Onishi <sup>\*1</sup>, Hiromitsu Shiina <sup>\*2</sup>, Tomohiko Yasumori <sup>\*2</sup>

<sup>\*1</sup> Graduate School of Informatics, Okayama University of Science

<sup>\*2</sup> Okayama University of Science

小学校等の授業において、教員が省察に使用できる時間は少なく、短時間での省察が必要になっており、システム化が望まれている。そこで本研究では、小学校の算数の授業の発話をテキスト化し、ニューラルネットワークを用いた対話モデルによって教員と児童の発話を分析している。また、発話の分類とシーンを分析・可視化するシステムの開発を行っている。

キーワード: 対話, 深い学び, 対話分析, 対話モデル, シーンの可視化

### 1. はじめに

文部科学省中央教育審議会<sup>(1)</sup>では教員に求める資質・能力として実践的指導力を挙げ、教員の探求力や学び続ける教員像の確立が必要であるとしている。続いて、平成 29 年公示の文部科学省小学校指導要領<sup>(2)</sup>で「主体的・対話的で深い学び」を実現する授業が目指されている。小学校の授業において様々な取り組みが行われてきており、教育学の実証研究における授業の分析では、教員や児童の発話の分析も行われている。その中で省察活動<sup>(3)</sup> <sup>(4)</sup>が注目されている。熟達化研究の観点では、教員は自立した学習者を育てるには、適応的熟達者であるべきであるとされている<sup>(5)</sup>。PCK (Pedagogical Content Knowledge) をベースとした省察による研究<sup>(6)</sup>では、教員の適応的熟達者は、教師中心と学習者中心の PCK の 2 面から省察を行っていることが示されており、授業中の発話について教師中心だけでなく、学習者中心の PCK に基づいた支援が求められている。教師中心や学習者中心の PCK に加えて協働的な学習を重視する省察マトリックス<sup>(6)</sup>が提案されている。「主体的・対話的で深い学び」の観点から省察方法のデジタル化を分析した研究<sup>(7)</sup>がなされており、コンピュータシステムによる発話分析<sup>(8)</sup>が試み

られている。その背景には日本の小中学校の教師は多忙で、研究授業や省察といった研修に使う時間が 48 か国で最も短いという調査結果<sup>(9)</sup>がある。多忙な教師にとって、いつでも一人で負担なく省察できる方法である機械的手法による分析手法の開発は急務である。

実際の小学校の授業では、教員は児童の状況を見ながら授業を進めており、教員が一方向的に授業を行っていることはまれであると考えられる。児童の場合は授業について自分の学びの状況や意見、感想を同級生間で発話することが多くあり、児童は対話をしながら授業を進めていると考えられる。また、大学においても反転授業が進められており、学生間の対話や教員との対話が理解を深めるという意味で重要である。教員と児童や児童間では、一種の対話が成り立っており、授業の理解を促したり示したりする発話や対話を自動的に分析することができれば、教員に対して多くのフィードバックが可能となる。

機械学習を用いた自然言語処理では、Transformer<sup>(10)</sup>を用いた BERT<sup>(11)</sup>等の文脈を考慮した言語処理が可能となっている。さらに、対話においてはチャットボット等の対話応答生成の研究が行われており、Transformer を用いた GVT (Global Variational

表 1 主体的・対話的で深い学びのラベル

小分類	大分類
主体的な学び	見通しを持たせる 主体的 児童の主体的な学びとは逆行している 振り返りに課題あり。教師が一方的に提示している。 正答を出すことが目的になっている
対話的な学び	対話させている 対話を促す
深い学び	深い学び 関数 数学的な見方

表 2 シーンの一覧

シーン
前回の振り返り
前回の振り返り_発表
作業の説明
作業させる
問題を解く
話し合い
答え合わせ
答え合わせ_発表
発表まとめ
授業まとめ
授業まとめ_発表
挨拶

Transformer)<sup>(12)</sup>が提案されている。

本研究では、事前にクラスタリングを用いて発話者の特徴を抽象化する拡張 GVTSC モデルを提案すると共に、小学校の授業の対話に対して、拡張 GVTSC モデルによる発話の分析を行っている。具体的には、小学校の算数の授業を録画し、教員と児童の発話に対して文字起こしを行った対話形式のテキスト情報を分析している。分析においては、あらかじめ授業中の発話に対して、人手でアノテーションを 2 種類付している。一種類目は、「主体的・対話的で深い学び」に関連する発話に対して、学びの種類を示すラベルを付与している。もう一方のアノテーションは、授業における場面を表すシーンを付与している。これらのアノテーション後に、その発話と近い距離にある発話を提案する拡張 GVTSC を用いて抽出している。また、授業中の発話に対して、ラベルの推定及び、シーンの推定を行っている。発話に対して、推定したラベルやシーンを見るだけでは、省察の補助としての利用は簡単ではない。そこで、授業の動画と併せて、ラベルやシーンの分布を時系列で確認できるように可視化を行うシステムの開発を行っている。

## 2. 主体的・対話的で深い学びの分類

「主体的・対話的で深い学び」に関する定義や具体的な発話例の整理を行った研究<sup>(13)</sup>では、「主体的な学

び」、「対話的な学び」、「深い学び」に大きく分けられおり、複数の分類に重複する発話の存在が示されている。本研究で用いた「主体的・対話的で深い学び」の分類 10 種類を表 1 に示す。「主体的・対話的で深い学び」の定義には、各教科・領域に共通する一般化された定義と教科固有の定義がある。本研究では小学校の算数の授業を対象としているため、「数学的な見方」等の教科固有の分類が存在している。

## 3. 授業の場面の分類

授業の対話データに対して、授業における場面を表すシーンをアノテーションしている。シーンは計 12 種類となっており、シーンの一覧を表 2 に示す。

## 4. 利用した授業の対話データ

授業の対話データは、小学校の算数の授業（45 分）を録画した上で、教員と児童の発話に対して文字起こしを行い、対話形式のテキスト情報を作成している。本研究では、あらかじめ授業中の発話に対して、人手でアノテーションを 2 種類付している。一種類目は、「主体的・対話的で深い学び」に関連する発話に対して、表 1 の学びの種類を示すラベルを付与している。もう一方のアノテーションは、表 2 の授業における場面を表すシーンを付与している。授業のデータは、4 年生の算数（比例）の授業 1 と 6 年生の算数（比例）の授業 2 を得られている。授業 1 のデータは、発話数が

193, ラベルを付与した発話数は 13 である. 授業 2 のデータは, 発話数が 274, ラベルを付与した発話数は 27, 全ての発話を各シーンに分類している.

## 5. 発話者のクラスタリングを追加した拡張 GVTSC モデル

### 5.1 概要

対話応答生成には, 様々な対話に対しての応答と成り得る無難な応答を生成するために, 応答の多様性が低くなるという課題が存在する<sup>(14)(15)</sup>. GVT モデルは, Decoder の入力にサンプリングした潜在変数を利用する手法であるが, 発話者の特徴を潜在変数で表現し, サンプリングを行うことで, 応答の多様性をもたらしていると考えられる. しかし, 先行研究では潜在変数が生成された応答の一貫性を低下させる傾向があることが示されている<sup>(16)</sup>. そこで, 発話者ごとの特徴を考慮するために, クラスタリングを用いて発話者の特徴を抽象化し, Encoder においてその発話者の特徴を加味することで, 一貫性と多様性を向上させる. 本研究では, 話者の分離を行っている拡張 GVT モデルに対して, 事前に発話者を分類するクラスタリングを追加した拡張 GVTSC モデルを提案し, 評価を行っている.

### 5.2 クラスタリングによる発話者の特徴作成

拡張 GVTSC モデルの概要を図 1 に示す. GVT モデルに発話者ごとの Encoder を追加した拡張 GVT モデルに対して, クラスタリングを用いて発話者の特徴ベクトルを作成する部分(図 1 の点線部分)を追加し, コンテキストのエンコードにおいて利用している.

拡張 GVTSC モデルの処理について述べる. 最初に, クラスタリングを用いて発話者の特徴ベクトルを作成する. コンテキストは対話を行う 2 者の発話をまとめたものであり, 発話者ごとに分割することが可能である. そこで, 対話のコンテキストを発話者ごとに分割し, 発話者ごとに処理を行っている. 発話者ごとの処理は同様であるため, 発話者 1 の場合を述べる. Speaker1 TRS Encoder で発話者 1 のコンテキストをエンコードする. TRS Encoder では, 入力系列の先頭に CLS トークンを付加しており, Transformer によって出力ベクトルが計算される. 発話者 1 のコンテキストのベクトル (Speaker1 Vector) として, CLS ト

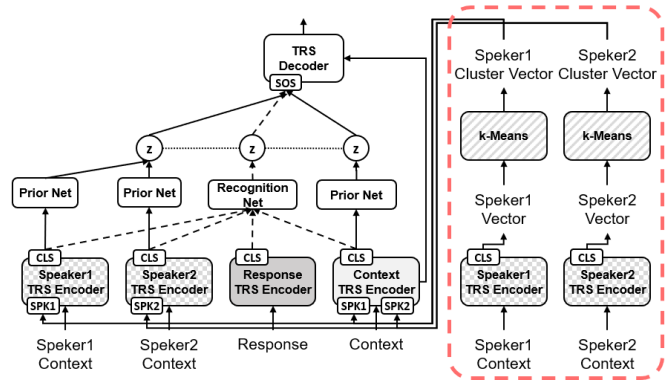


図 1 拡張 GVTSC モデル

ークンのベクトルを取得する. Speaker1 Vector に対して, クラスタリングを行う. 本研究では, クラスタリングに k-Means を使用している. クラスタ数  $k$  については, ハイパーパラメータ同様実験において決定する必要がある. クラスタリングの結果, Speaker1 Vector が属するクラスタを予測し, そのクラスタの中心ベクトル (Speaker1 Cluster Vector) を取得する. 発話者 2 に対して, 発話者 1 と同様の処理を行い, Speaker2 Cluster Vector を取得する. ここで, クラスタリングに使用している TRS Encoder は, 応答生成で訓練している TRS Encoder を共有している. ただし, クラスタリングの処理では, 誤差逆伝搬による訓練は行われない.

### 5.3 対話応答生成

対話のコンテキスト全体は, Context TRS Encoder に入力し, 出力ベクトルを得る. コンテキストのエンコードでは, 入力系列に発話者ごとのトークン (SPK1, SPK2) を追加し, SPK1 に Speaker1 Cluster Vector, SPK2 に Speaker2 Cluster Vector を入力する. また, 発話者ごとにコンテキストを分割し, それぞれを各 Speaker TRS Encoder に入力し, 出力ベクトルを得る. その際には, Speaker1 TRS Encoder には入力系列に発話者 1 のトークン (SPK1) を追加し, Speaker1 Cluster Vector を入力し, Speaker2 TRS Encoder には入力系列に発話者 2 のトークン (SPK2) を追加し, Speaker2 Cluster Vector を入力する. 発話者ごとのコンテキストのエンコードにおいて, 各発話者の特徴ベクトルを利用することで, より発話者の特徴を考慮したエンコードを図っている.

事前・事後分布を多層パーセプトロン (MLP) によって近似した Prior Net・Recognition Net から潜在変

数 $z$ をサンプリングする。Prior Net は、Speaker TRS Encoder または、Context TRS Encoder の CLS トークンの出力ベクトルを基に、MLP によってコンテキストのベクトルの平均と分散を推定する。その平均と分散に従う正規分布より潜在変数 $z$ をサンプリングする。Recognition Net では、Speaker TRS Encoder と Context TRS Encoder に加えて Response TRS Encoder の CLS トークンの出力ベクトルも用いて、MLP により対話全体のベクトルの平均と分散を推定する。Prior Net と同様に推定した平均と分散に従う正規分布から潜在変数 $z$ のサンプリングを行う。TRS Encoder の CLS トークンの出力ベクトルは入力全体の表現したベクトルとみなすことができるため、CLS トークンの出力ベクトルから事前・事後分布を生成し、潜在変数 $z$ をサンプリングしている。

TRS Decoder では、入力系列の先頭の SOS トークンに通常の潜在変数に加えて、応答の発話者の潜在変数を入力することで、潜在変数を応答の生成に利用している。この際、TRS Decoder は、学習時には Recognition Net からサンプリングした潜在変数を利用し、生成時には Prior Net からサンプリングした潜在変数を利用する。

#### 5.4 拡張 GVTSC モデルの定式化

拡張 GVTSC モデルは、 $c$ をコンテキスト、 $c_{s1}$ を発話者 1 のコンテキスト、 $c_{s2}$ を発話者 2 のコンテキスト、 $x$ を応答、 $z$ を潜在変数、 $v_{s1}$ を発話者 1 のクラスタベクトル、 $v_{s2}$ を発話者 2 のクラスタベクトルとして、下記の ELBO を最大化することでモデルの最適化を行う。

$$\begin{aligned} \mathcal{L}_{ELBO}(x, c) &= \log p(x|c) \\ &\geq E_q(z|x, c, v_{s1}, v_{s2})[\log p(x|z, c, v_{s1}, v_{s2})] \\ &\quad -KL(q(z|x, c, v_{s1}, v_{s2})|p(z|c, v_{s1}, v_{s2})) \\ &\quad -KL(s(z|x, c_{s1}, c, v_{s1}, v_{s2})|r(z|c_{s1}, v_{s1})) \\ &\quad -KL(s'(z|x, c_{s2}, c, v_{s1}, v_{s2})|r'(z|c_{s2}, v_{s2})) \end{aligned} \quad (1)$$

ここで、 $KL$ は分布間の KL divergence であり、事前分布 $p$ ,  $r$ ,  $r'$ は、下記の式で定義される。

$$p(z|c, v_{s1}, v_{s2}) \sim \mathcal{N}(\mu_p, \sigma_p^2) \quad (2)$$

$$r(z|c_{s1}, v_{s1}) \sim \mathcal{N}(\mu_r, \sigma_r^2) \quad (3)$$

$$r'(z|c_{s2}, v_{s2}) \sim \mathcal{N}(\mu_{r'}, \sigma_{r'}^2) \quad (4)$$

ここで、

表 3 自動評価指標による結果

モデル	多様性			類似性
	Dist-1	Dist-2	Dist-3	BERT
GVT	0.484	0.720	0.739	0.654
拡張 GTSC	<b>0.640</b>	<b>0.950</b>	<b>0.975</b>	<b>0.672</b>
実際の応答	0.647	0.947	0.963	-

$$[\mu_p, \log(\sigma_p^2)] = \text{MLP}_p(c, v_{s1}, v_{s2}) \quad (5)$$

$$[\mu_r, \log(\sigma_r^2)] = \text{MLP}_r(c_{s1}, v_{s1}) \quad (6)$$

$$[\mu_{r'}, \log(\sigma_{r'}^2)] = \text{MLP}_{r'}(c_{s2}, v_{s2}) \quad (7)$$

事後分布 $q$ ,  $s$ ,  $s'$ は、下記の式で定義される。

$$q(z|x, c, v_{s1}, v_{s2}) \sim \mathcal{N}(\mu_q, \sigma_q^2) \quad (8)$$

$$s(z|x, c_{s1}, c, v_{s1}, v_{s2}) \sim \mathcal{N}(\mu_s, \sigma_s^2) \quad (9)$$

$$s'(z|x, c_{s2}, c, v_{s1}, v_{s2}) \sim \mathcal{N}(\mu_{s'}, \sigma_{s'}^2) \quad (10)$$

ここで、

$$[\mu_q, \log(\sigma_q^2)] = \text{MLP}_q(x, c, v_{s1}, v_{s2}) \quad (11)$$

$$[\mu_s, \log(\sigma_s^2)] = \text{MLP}_s(x, c_{s1}, c, v_{s1}, v_{s2}) \quad (12)$$

$$[\mu_{s'}, \log(\sigma_{s'}^2)] = \text{MLP}_{s'}(x, c_{s2}, c, v_{s1}, v_{s2}) \quad (13)$$

また、学習が進むにつれて Decoder が潜在変数 $z$ の情報を考慮しなくなる KL vanishing 問題のため、KL アニーリング<sup>(17)</sup>、BoW (Bag-of-Words) loss<sup>(18)(19)</sup>の手法を取り入れている。KL アニーリングは式1の KL divergence の値について、学習が進むに連れて 0 から 1 に線形に増加する重みをつける手法である。BoW loss は応答に含まれる単語の集合を潜在変数から推定するサブタスクを追加する手法であり、潜在変数と応答中の単語の関連性を強くすることを目的としている。

## 6. 対話モデルの評価

データセットには小学校の授業対話データを用いた。前処理として SentencePiece を用いてサブワードへ分割している。コンテキストの長さについては、3-turn までの対話応答を評価している。自動評価指標として、Dist-n<sup>(20)</sup>と BERT Score<sup>(21)</sup>を用いている。Dist-N は、N-gram の総数に対して N-gram の種類数が占める割合を算出し、この比率が高いほど、多様性が高いことを示す指標である。BERT Score は、事前学習した BERT の埋め込みを使用し、モデルが生成した応答と参照応答の類似性を評価する手法である。

各モデルが生成した応答の自動評価指標による結果を表 3 に示す。なお、拡張 GVTSC モデルのクラスタ

表 4 対話応答の生成例

コンテキスト
発話 1: 12。
発話 2: 4 段の時は?
発話 3: 16。
発話 4: 5 段の時は?
発話 5: 24。24,24,24,28。えー。
応答
<b>GVT:</b> そうそうそう。はい,ここまでのいいかな?同じです。
<b>拡張 GVTSC:</b> この式,合ってる?違ってる?
<b>参照:</b> これで合ってる?

リング (k-Means) では, 予備実験における結果からクラスタ数を 8 としている。

拡張 GVTSC は, 多様性の評価では全ての N-gram において GVT モデルより高い評価を得ている。また, 実際の応答の多様性に近い評価となっている。類似性の評価においても, GVT モデルから約 0.018 向上している。Encoder におけるエンコード過程で発話者の特徴を考慮することで, 潜在変数のサンプリングや Decoder での Attention に利用する Encoder の各トークンの出力ベクトルに影響を与えていると考えられる。

小学校の授業対話データの評価において生成された応答の例を表 4 に示す。GVT モデルと比較すると, 拡張 GVTSC モデルはコンテキストに関連した応答を生成することができており, 参照応答と意味的に近い応答でもある。また, 多様性がある応答となっている。

## 7. 拡張 GVTSC による発話ベクトルの生成

本研究では拡張 GVTSC モデルによる対話応答生成が目的ではなく, 対話のベクトル化が目的である。最初に, 対話データを用いて拡張 GVTSC モデルの訓練を行う。次に, 訓練したモデルを用いて, 対話のベクトル化を行う。対話データを用いて, 拡張 GVTSC モデルを訓練することで, モデルは対話応答を生成するために必要なコンテキストの発話をベクトル化する能力を得ている。対話のベクトルは, 拡張 GVTSC モデルにコンテキストとして対話を入力し, 拡張 GVTSC モデルの Context TRS Encoder と Speaker TRS Encoder が出力した CLS トークンのベクトルの和を

計算することで作成している。処理手順を次に示す。

- (1) 対話を発話者ごとの発話に分割し, 発話者ごとのコンテキストと対話全体のコンテキストを作成する。
- (2) Tokenizer で各発話をトークン (単語, 単語をさらに分けたサブワード) に分割する。
- (3) 文全体を表すベクトルを生成するために, トークン系列の先頭に特殊なトークン ([CLS]) を追加する。
- (4) 拡張 GVTSC モデルにトークン系列を入力し, トークンの情報とトークンの文中の位置情報を埋め込んだベクトルに変換する。
- (5) Transformer によって他のトークンの情報を取り込み, 文脈を考慮したベクトルを生成する。
- (6) 発話文全体のベクトルとして, Context TRS Encoder と 2 個の Speaker TRS Encoder が出力した CLS トークンのベクトルの和を出力する。

## 8. 発話ベクトルによる発話分析

拡張 GVTSC を用いて生成した発話ベクトルによる分析では, 授業の対話データについて, 1 発話ごとと, 対話として扱うために複数の発話ごとのデータを作成し, 発話について, アノテーション付き発話との距離を求めている。距離については, ベクトルで表現できていることから Cos 類似度を用いた。比較のために, 文字の一致に関する距離であるジャロ・ウィンクラー距離 (JW 距離, 0-1 の値を取る) も求めている。

授業 1 のデータのラベルを教師データとして, 授業 2 のデータから類似発話を抽出した例を表 5 に示す。1 行目は, 1 発話単位で類似性を図っているが, 「数学的な見方」(比例の関係) や「深い学び」(他の意見を児童に聞く) のラベルに合致した類似発話を抽出できている。また, 類似発話に人手で付与されていたラベルは「数学的な見方」, 「深い学び」であり, 人間による分類とも一致する結果となっている。2 行目は, 2 発話単位の場合であり, 児童が比例する値の求め方を発表する発話を抽出しており, ラベルあり発話との表層的な類似性は 0.111 と非常に低い発話である。この類似発話には人手で付与されたラベルはなく, ラベルを付与する発話の候補を機械的に得ることができること

表 5 授業 1 のデータのラベルを教師データとした授業 2 のデータからの類似発話の抽出例

ラベルあり発話	ラベル	類似発話	Cos 類似度	JW 距離
段の数を、ありがとう。はい、書くよ。段の数を 4 倍すると、周りの長さになる。 <u>ちなみにさあ、今のは縦に見とるん？横に見とるん？</u>	数学的な見方, 関数, 深い学び	〇〇さんはこの考え方でやってやったんだよね。 <u>1 と 1 5 に目をつけて 4 倍 4 倍って考えたんですね。正しいですね。</u> ちなみに〇〇さんさ、ここに目をつけたんだけど <u>他のところに目をつけた人いる？</u> うん。はいチャレンジャー。	0.953	0.450
うん。 <u>[SEP]じゃあ、どんな決まり見つけたん？みんな。</u> 見つけた <u>決まりを教えてください。</u> 〇〇君。	数学的な見方, 関数, 深い学び	<u>表には表には入っていない値を求めるには、1 0 のような切りの良い数字で数字から倍すればいい</u> と思いました。 <u>[SEP]どれでもいいんだけど切りの良い</u> この時に、よし。じゃああと 2 分しかなくなっちゃいました。	0.962	0.111

授業対話の分析

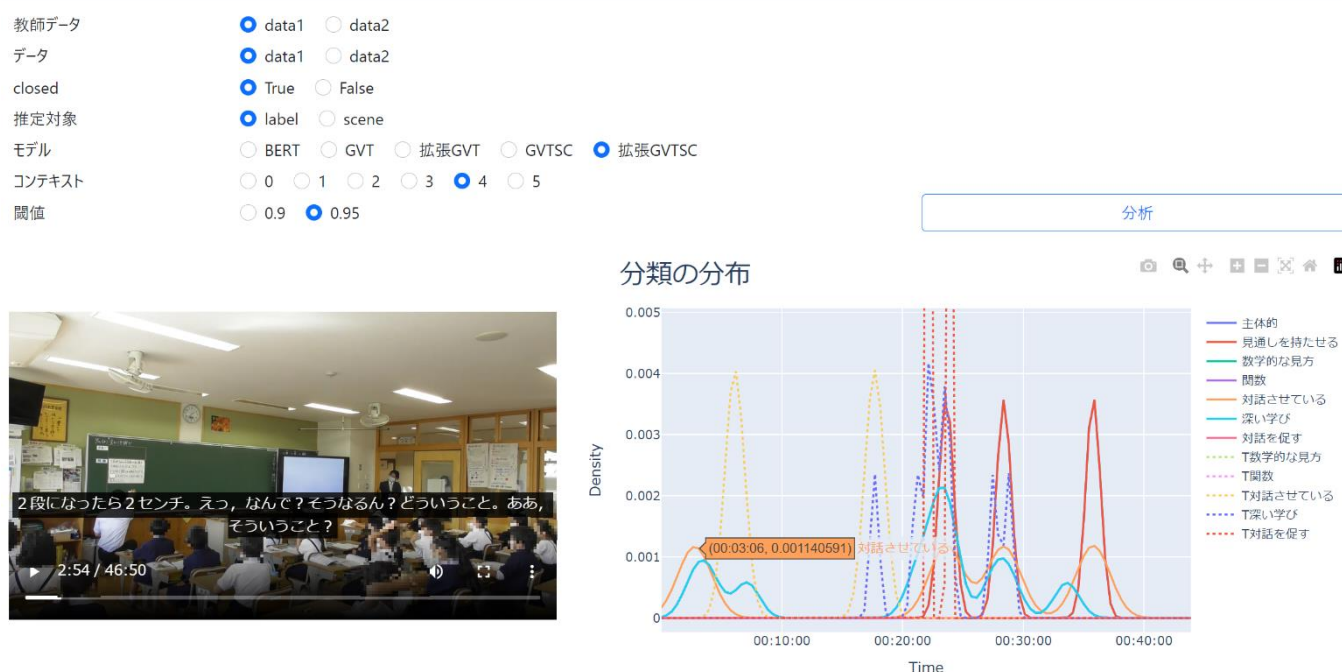


図 2 授業の発話分析の可視化システム

を示している。

### 9. 授業の発話分析の可視化システム

授業中の発話の種類や傾向を見ることで、教員により良いフィードバックができると考えられる。フィードバックについては、授業の動画と発話の分類や分布を可視化することで、より分かり易くなると考えられる。本研究では、発話の字幕を追加した授業の動画と発話の分類の分布を表示するシステムを開発している。特に、発話の分類の分布を推定するとともに、グラフとして表示している。分類分布の推定手法及び、可視化

方法を次に示す。

- (1) モデルを用いて、発話ごとに分類とその類似度を推定する。
- (2) 閾値 (0.9, 0.95) を設定し、閾値以上の類似度の分類のみを抽出する。
- (3) 類似度の平均を重みとして各分類の確率密度関数を推定する。確率密度関数の推定<sup>(22)</sup> <sup>(23)</sup>には、Python ライブラリの `scipy1.8.1`<sup>(24)</sup>を用いている。
- (4) 分類ごとの確率密度関数を分布として、全ての授業時間における分布をグラフとして可視化する。
- (5) 授業動画と分類分布のグラフを並べて表示する。

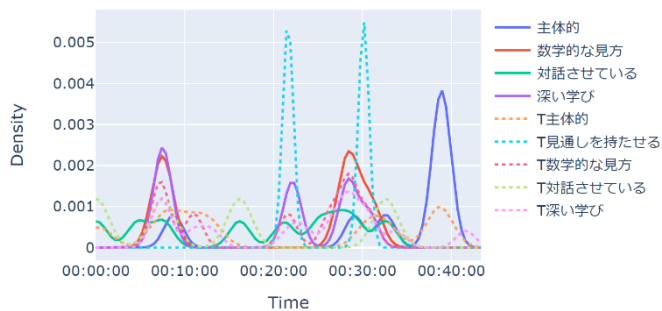


図 3 ラベルの推定と可視化（授業 2）

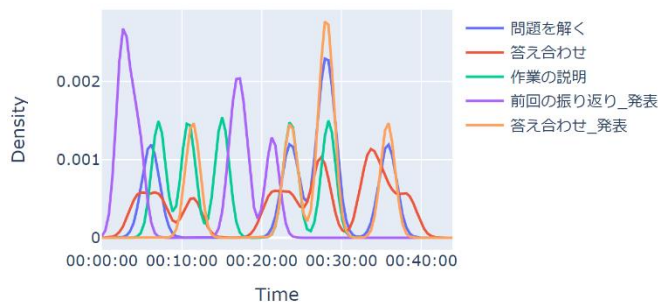


図 4 シーンの推定と可視化（授業 1）

授業動画と発話の字幕，分類分布のグラフの可視化システムを図 2 に示す。システムでは，画面の左側に授業動画を字幕付きで表示し，右側に分類分布のグラフを表示している。ただし，グラフの点線の分布は教師ラベルとしてあらかじめ分類した分布を示している。

### 9.1 学びに関するラベルの可視化

ラベルの可視化の例を図 3 に示す。このグラフは授業 2 にアノテーションしたラベルを用いて，授業 2 のラベルを推定・可視化している。推定においては，6 発話単位で閾値を 0.95 としている。アノテーションした発話は推定対象に含まないが，基本的に人手でアノテーションしたラベルと同様の分布となっている。人手との差異は，ラベル「対話させている」が 20 分から 30 分の間でも推定された点であり，実際の授業では児童が発表等で発話している部分と一致している。

### 9.2 シーンの可視化

シーンの可視化例として，授業 2 にアノテーションしたシーンを用いて，授業 1 のシーンを推定・可視化したグラフを図 4 に示す。推定では，5 発話単位で閾値を 0.95 としている。このグラフでは，全体的に「答え合わせ」関連のシーンが多く表れる結果となっているが，実際の授業においても細かく児童と対話をしながら解答と発表が行われていた。しかし，授業 2 では「前回の振り返り」は序盤に少し行われた程度であり，

その後は今回の内容の発表が行われているため，「答え合わせ\_発表」と推定するべきであり，推定精度の改善が必要である。

### 9.3 可視化システムに対するアンケート評価

学びのラベルを時系列的分布として可視化するシステムについて，小学校の教員経験者に対する自由記述によるアンケート評価を行った。小学校の教員経験者の評価を以下に示す。

- 「深い学び」や「対話的」のラベル分布が授業全体において見られ，「主体的・対話的で深い学び」を意識した授業であることと一致している。
- 同じラベルは同じ色で教師ラベルも推定ラベルも表示した方が分かり易い。
- 可視化システムとしては，分布上のある時点における発話の内容を確認できるとより使い易い。動画上に字幕として表示しているが，前後の発話も併せて確認できるように，アンケートの改善点である発話内容の確認機能を実装することで，分布のラベルと発話の関連性から要因の分析が可能になる。

## 10. おわりに

本研究では，小学校の授業を対象に，対話モデルを用いて発話の学びの分類や授業における場面を推定し，可視化を行った。省察の補助として授業の分析結果の可視化システムを開発し，小学校の教員経験者による評価では授業全体の傾向を知ることができるとされた。また，提案した発話者情報を組み込む拡張 GVTSC モデルは，対話応答生成において多様性及び類似性で改善が見られた。今後の課題は，アンケート評価で得られた可視化システムの課題の改善及び，学びの分類や授業の場面の推定精度の向上である。

### 参 考 文 献

- (1) 中央教育審議会：“教職生活の全体を通じた教員の資質能力の総合的な向上方策について（答申）”，[https://www.mext.go.jp/b\\_menu/shingi/chukyo/chukyo00/toushin/1325092.htm](https://www.mext.go.jp/b_menu/shingi/chukyo/chukyo00/toushin/1325092.htm) (2022 年 10 月 3 日確認)
- (2) 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](https://www.mext.go.jp/content/1413522_001.pdf) (2022 年 10 月 3 日確認)

- (3) 秋田喜代美: “教師教育から教師の学習過程研究への転回—マイクロ教育実践研究への変貌—”, 変貌する教育学, 世織書房, pp.45-75 (2009)
- (4) 坂本篤史: “授業研究の事後協議会における教師の省察過程の検討—授業者と非授業者の省察過程の特徴に着目して—”, 教師学研究, Vol.9, No.8, pp.27-37 (2010)
- (5) 坂本篤史: “現職教師は授業経験から如何に学ぶか”, 教育心理学研究, Vol.55, No.4, pp.584-596 (2007)
- (6) 保森智彦: “算数の授業中と省察の発話プロトコル分析をととした教師の PCK の検討”, 日本教科教育学会誌, Vol.41, No.1, pp.59-71 (2018)
- (7) 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”, 学習開発学研究, 広島大学大学院人間社会科学研究所科学習開発学領域, No.14, pp.3-11 (2022)
- (8) Wang, Y., 大井翔, 松村耕平, et al.: “新任教員の授業力向上のための授業振り返りシステムに関する研究”, 情報処理学会インタラクシオン, pp.753-757 (2021)
- (9) 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”, 国立教育政策研究所 (編), ぎょうせい (2018)
- (10) Vaswani, A., Shazeer, N., Parmar, N., et al.: “Attention is all you need”, Advances in neural information processing systems, pp.5998-6008 (2017)
- (11) Devlin, J., Chang, M.-W., Lee, K., et al., “Bert: Pre-training of deep bidirectional transformers for language understanding”, arXiv preprint arXiv:1810.04805 (2018)
- (12) Lin, Z., Indra, W. G., Xu P., et al.: “Variational Transformers for Diverse Response Generation”, arXiv preprint arXiv:2003.12738 (2020)
- (13) 保森智彦: “「主体的・対話的で深い学び」を実現するための教師の発話の検討”, pp.45-52, No.57, B, 人文・社会科学, 岡山理科大学紀要 (2021)
- (14) Sordani, A., Galley, M., Auli, M., et al.: “A Neural Network Approach to Context-Sensitive Generation of Conversational Responses”, In Proceedings of the 2015 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, pp.196-205 (2015)
- (15) Csaky, R., Purgai, P. and Recski, C.: “Improving Neural Conversational Models with Entropy-Based Data Filtering”, In Proceedings of the 57th Annual Meeting of the Association for Computational Linguistics, pp.5650-5669 (2019)
- (16) Sun, B., Feng, S., Li, Y., et al.: “Generating Relevant and Coherent Dialogue Responses using Self-Separated Conditional Variational AutoEncoders”, In Proceedings of the 59th Annual Meeting of the Association for Computational Linguistics and the 11th International Joint Conference on Natural Language Processing (Volume 1: Long Papers), pp.5624-5637 (2021)
- (17) Bowman, R. S., Vilnis, L., Vinyals, O., et al.: “Generating Sentences from a Continuous Space”, In Proceedings of the 20th SIGNLL Conference on Computational Natural Language Learning, pp.10-21 (2016)
- (18) Zhao, T., Zhao, R. and Eskenazi, M.: “Learning Discourse-Level Diversity for Neural Dialog Models using Conditional Variational AutoEncoders”, In Proceedings of the 55th Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers), pp.654-664 (2017)
- (19) Zhou, X. and Wang, Y. W.: “Mojitalk: Generating Emotional Responses at Scale”, In Proceedings of the 56th Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers), pp.1128-1137 (2018)
- (20) Li, J., Galley, M., Brockett, C., et al.: “A Diversity-Promoting Objective Function for Neural Conversation Models”, In Proceedings of the 2016 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, pp.110-119 (2016)
- (21) Zhang, T., Kishore, V., Wu, F., et al.: “Bertscore: Evaluating Text Generation with BERT”, In International Conference on Learning Representation (2019)
- (22) Scott, D. W.: “Multivariate Density Estimation: Theory, Practice, and Visualization”, Wiley series in probability and mathematical statistics, Wiley, New York (1992)
- (23) Silverman, B. W.: “Density Estimation for Statistics and Data Analysis”, Chapman & Hall/CRC monographs on statistics and applied probability, Chapman and Hall, London (1986)
- (24) Virtanen, P., Gommers, R., Oliphant, T. E., et al.: “SciPy 1.0: Fundamental Algorithms for Scientific Computing in Python”, Nature Methods, Vol.17, pp.261-272 (2020)