

令和 4 年 6 月 29 日現在

機関番号：32689

研究種目：若手研究

研究期間：2019～2021

課題番号：19K20426

研究課題名（和文）ソーシャルメディアを分析ソースとした議論遂行能力育成のための有用事例抽出手法

研究課題名（英文）Methods for Extracting the Useful Cases from Social Media Aimed at Fostering the Discussion Skill

研究代表者

中山 祐貴（Nakayama, Hiroki）

早稲田大学・グローバルエデュケーションセンター・講師（任期付）

研究者番号：80761569

交付決定額（研究期間全体）：（直接経費） 3,300,000円

研究成果の概要（和文）：PBLなど主体性の高い学習活動における議論は、全体成果に影響を与える重要な存在である。その能力を育成するためには、未熟者に議論経験を積ませるべきであるが、参考になる事例の準備困難性などの問題から、十分に行えない実情がある。本研究では、学生の興味に応じてソーシャルメディア上の議論事例を抽出し、経験の浅い学生が自ら実際の議論の様子を観察できるように視覚的に提示する手法を開発した。これらの手法に基づく支援システムを実装することで、身近な潜在ソースを活かした議論遂行能力育成の新たな可能性を示した。

研究成果の学術的意義や社会的意義

議論経験の積み重ねの重要性が指摘されていたが、参考となり得る事例準備の負担から、これまで困難であった「議論経験を自ら醸成するきっかけとなる事例を示唆する育成方法」を可能とする教育支援充実の意義を有する。

また、学習者の議論経験の不足から生じる問題と、現実的な制約下で有用議論事例を提示する支援の困難さ、という二重構造の課題に対して、身近な潜在ソースを活用し得る形で議論遂行能力育成に挑戦する本研究は、教育工学研究の新たな可能性を提示する意義を有する。

研究成果の概要（英文）：In proactive learning activities such as PBL, discussions among students are important works that effect on overall outcome. Although students should gain experience in discussion to acquire the skills to carry out it, it is difficult for students to fully experience the discussions due to the difficulty of obtaining useful discussion cases. In this research, methods for extracting the discussion cases on social media in accordance with unskilled students' interests and visually presenting them so that students can observe discussion circumstances by themselves were developed. Moreover, a prototype was implemented based on these methods. Consequently, a novel way for fostering the skills to carry out discussions by utilizing familiar potential sources was shown.

研究分野：教育工学

キーワード：議論遂行能力育成 議論事例 ソーシャルメディア 情報視覚化 教育工学

様式 C - 19、F - 19 - 1、Z - 19 (共通)

1. 研究開始当初の背景

PBL(Problem-Based Learning)など主体的学習活動の中で、議論を実施する機会が増加している。この種の議論を上手く遂行するためには、各メンバの意見を的確に把握して、積極的に発言することが重要であるが、未熟者にとって容易ではない。議論遂行能力を育成するためには、未熟者に議論経験を積ませるべきであるが、参考になる事例の準備困難性などの問題から、十分に行えない実情がある。

これに対して、議論の内容要約や論点の抽出などの取り組みが報告されているが、議論の内容理解支援を目指したものが殆どである。また、議論の停滞時に発言を促すシステムが報告されているが、議論進行の円滑化に主眼を置いたもので、議論の遂行能力育成には及んでいない。

ゆえに、議論経験の獲得の支援として、未熟者の関心に応じて有用な過去の議論事例を抽出・提示する方法が必要と考えられる。

2. 研究の目的

本研究では、議論経験を積み重ねるきっかけとなる事例を未熟者に提示することで、議論遂行能力を育成する手法の開発を目指す。具体的には、情報の発信・共有に広く活用されているソーシャルメディア上の会話から、有用な議論事例を抽出し、観察し易い形で提示する手法を開発する。これにより、身近な潜在ソースを活かした議論遂行能力育成の新たな可能性を探る。

3. 研究の方法

本研究では、ソーシャルメディア上で実際に行われた議論の中から参考となり得るものを収集し、どのような話題が存在し、どのように発言が取り交わされているのか(議論様態)を視覚化する手法を開発する。

具体的には、まず、ソーシャルメディア上の会話から、ユーザの関心に応じて議論を抽出し、それらを構成する発言群を収集する手法を開発する。その上で、収集した発言データの分析に基づいて、発言と応答の積み重ねによる議論の流れ(議論構造)を抽出する手法、議論されている話題の分布を視覚化する手法をそれぞれ開発する。

これらの手法に基づいて、目を向ける話題の選択を補助し、複雑な議論構造の中で注目する話題に関する議論が実際に行われている部分を視覚的に提示するシステムを開発する。

4. 研究成果

4.1 議論様態視覚化手法

4.1.1 議論を構成するツイート群の収集

代表的なソーシャルメディアである Twitter 上の発言を収集・分析することで、参考となり得る議論を抽出する。Twitter では、ある発言をきっかけとして、それに反応する発言、または、それをきっかけとした新たな発言など、一連の会話が形成される。これら、一連の会話を1つの議論と呼ぶ。

まず、未熟者が関心のある議論を表現したキーワードに応じて、Twitter 上から該当するツイートを抽出する。ツイートの抽出には Twitter API を用いる。その際、一定の注目を集めているツイートであること、投稿日時が最近のツイートであることを加味して収集する。さらに、収集したツイート群の1つ1つについて、オブジェクト情報を抽出する。ここで扱う主なオブジェクト情報を表1に示す。

このようにして、未熟者の関心に応じた議論を抽出し、それを構成するツイート群を収集する。

表1 TwitterAPI で取得する主なオブジェクト情報

オブジェクト名	意味
creates_at	ツイートの投稿日時
favorited_count	当該ツイートがお気に入り登録された数
tweet_id	ツイート ID (数値型)
id_str	ツイート ID (文字列型)
in_reply_to_status_id	返信先のツイート ID
text	ツイートの本文 (URL やハッシュタグは含まない)
screen_name	Twitter のユーザ ID

4.1.2 議論構造の抽出

収集したツイート群のオブジェクト情報の分析に基づいて、ツイート間の返信関係を特定する。その上で、特定した返信関係から、議論を構成する会話の流れを抽出する。具体的には、まず、収集したツイート群からツイートを1つピックアップし、分析の基点(基点ノード)とする。基点ノードのオブジェクト情報を分析して、`in_reply_to_status_id`を抽出する。さらに、他のツイートのオブジェクト情報の `tweet_id` が、基点ノードの `in_reply_to_status_id` と一致するものを検索する。該当するツイートがあれば、それを基点ノードに対する返信ツイートとして抽出する(図1)。返信ツイートを抽出した場合は、その返信ツイートを新たな基点ノードとして、同様の検索を行う。この処理を返信ツイートがなくなるまで再帰的に実施する。このとき、返信先のツイートが収集したツイート群中に存在しない場合には、TwitterAPIを用いて、該当するツイートを検索・取得する。

これにより、ツリー状のツイートの繋がりを得る。ノードはツイートを表し、エッジはツイート間の返信関係を表す。基点ノードは、ツリーの根として表現される。このツイートによる返信の繋がりを議論構造と呼ぶ(図2)。

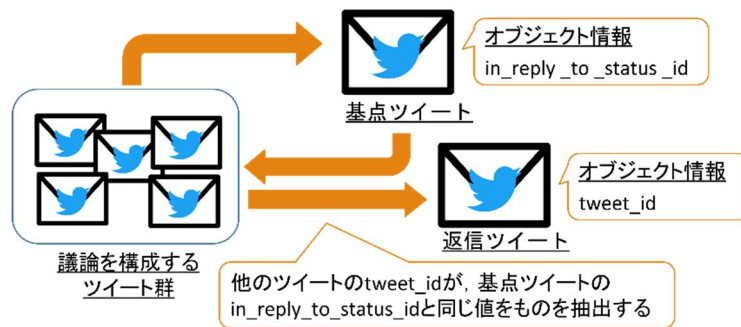


図1 返信ツイートの抽出

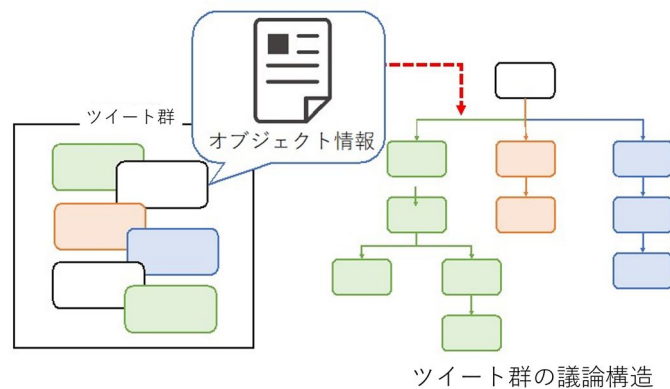


図2 議論構造の抽出

さらに、抽出した議論構造の中で、他者の注目を集めている発言を抽出する。ここでは、以下の3つの要素を考慮する。

- ・ Twitter 上で各ツイートに付与された「いいね」の数
- ・ Twitter 上で「リツイート」された数
- ・ Pagerank アルゴリズムに基づく重要度

具体的には、議論構造を構成する個々の発言について、これらの値を算出し、正規化する。その上で、議論構造中で、この3つの値の和が最も大きい発言を抽出する。ユーザに議論構造を提示する際に、他者の注目を集めている発言であることが分かるように強調表示する。これにより、多くの発言から構成される議論構造を観察する際に目を向けるきっかけを提供する。

4.1.3 話題を表す特徴フレーズの抽出手法

抽出した議論構造が、どのような話題についての発言のやり取りで構成されているのかを特定するために、ツイート本文を分析する。特に、ここでは、いくつかのフレーズが「議論されている話題」を体現する事象に注目する。

まず、議論構造を構成するツイートに含まれる特徴フレーズを抽出する。具体的には、Universal Dependencies に基づくオープンソース日本語自然言語処理ライブラリ GiNZA を用いて、ツイートに含まれる名詞句を抽出する。次に、ツイート中の1文ごとに文書ベクトルを生成し、ツイートに含まれる名詞句ごとに単語ベクトルを生成する。その上で、教師データ無しの重要語抽出アルゴリズムである EmbedRank を用いて、重要な名詞句を選出し、それを特徴フレーズとして採用する。ここでは、まず、文章ベクトルと単語ベクトルの類似度に基づいて、語の重みを算出する。その上で、重みのランキングを生成することで、上位N位の単語を重要語として抽出する。

また、似通った特徴フレーズを抽出し過ぎる問題を軽減するために、Maximal Marginal Relevance アルゴリズムを加味したフレーズの重みづけを採用している(式1)。Cは特徴フレーズ候補集合、Kは各文章に対する類似度が算出済みのフレーズの集合をそれぞれ表す。また、λはハイパーパラメータを表す。

$$MMR := \operatorname{argmax}_{C_i \in C \setminus K} [\lambda \operatorname{Sim}(C_i, \text{doc}) - (1 - \lambda) \max_{C_j \in K} \operatorname{Sim}(C_i, C_j)] \quad (1)$$

4.1.4 特徴フレーズの共起ネットワークの生成

ツイート本文中の特徴フレーズの共起状況を分析することで、話題を表す特徴フレーズのネットワークを生成する。

具体的には、まず、ツイートに含まれる特徴フレーズ群中の任意の2つについて、Jaccard 係数を算出する。特徴フレーズの全ての組について算出を実施した上で、Jaccard 係数が一定以上の特徴フレーズを結び付ける。その上で、特徴フレーズをノード、一定以上の共起度を持つ特徴フレーズ間の関係をエッジとして表すことで、一種のグラフを生成する(図3)。これを話題を表す特徴フレーズの共起ネットワークと呼ぶ。

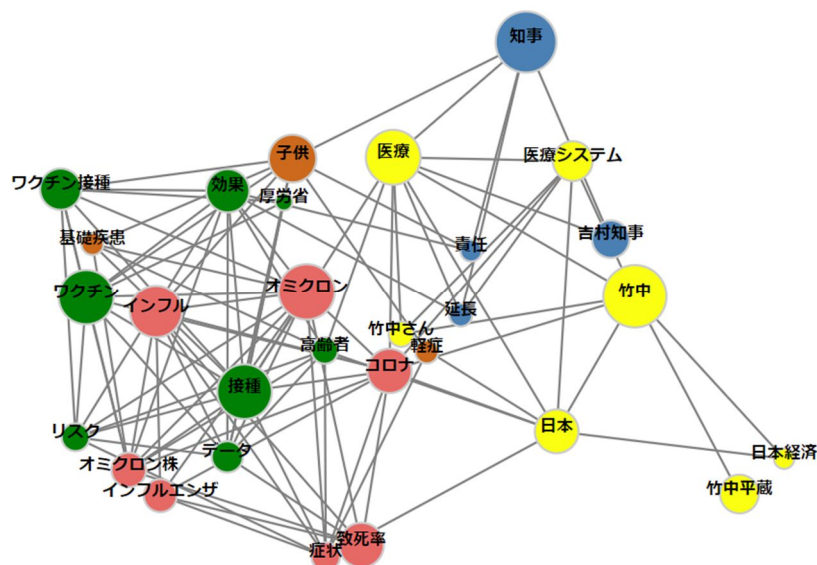


図3 話題を表す特徴フレーズの共起ネットワーク

次に、話題を表す特徴フレーズの組合せを抽出する。ここでは、共起ネットワークのノードを、モジュラリティに基づいてクラスタリングする。モジュラリティは、コミュニティ内部のエッジ密度は高く、コミュニティ間のエッジ密度は低くなるように、ノードをいくつかのコミュニティに分割する指標である(式2)。 e_{ii} は総エッジ数に対する「コミュニティ*i*内部における各ノードが有するエッジ数の総和」の割合である。 e_{ij} は総エッジ数に対する「コミュニティ*i*に属するノードからコミュニティ*j*に属するノードに張られているエッジ数の総和」の割合である。また、 a_i は総エッジ数に対する「コミュニティ*i*から他のコミュニティに向けて張られているエッジ数の総和」の割合を表す。これにより、各コミュニティに含まれる特徴フレーズの集合(話題を表す特徴フレーズの組合せ)を抽出することで、議論されている話題の分布を視覚化する。

$$Q = \sum_i (e_{ii} - (\sum_j e_{ij})^2) = \sum_i (e_{ii} - a_i^2) \quad (2)$$

4.2 議論事例理解支援システムの開発

開発した手法に基づいて、ソーシャルメディア上の議論データの収集から、議論構造抽出、話題分布の視覚化、これらの視覚的提示までを担うシステムを開発した。実装したシステムのイン

タフェースを図4に示す。本システムは、Python、JavaScriptを用いて実装した。また、データベースは、MySQLを用いて構築した。

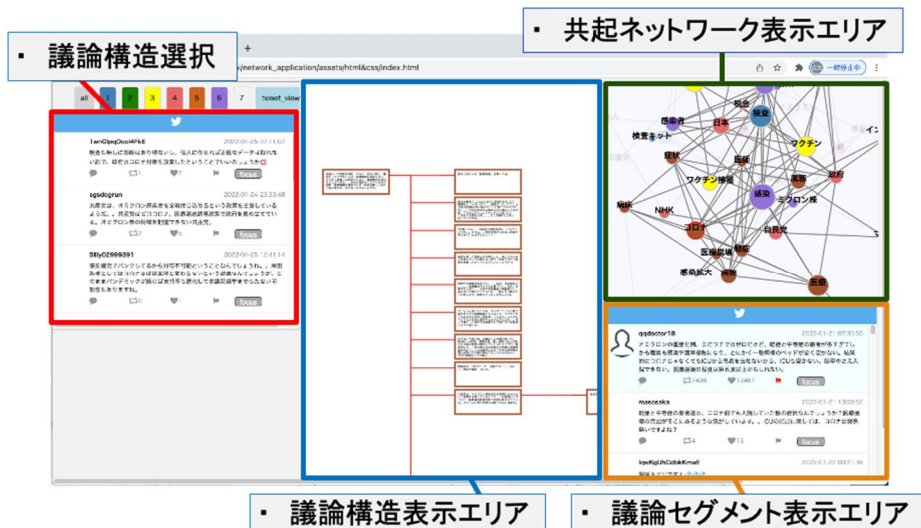


図4 議論事例理解支援システムのインタフェース

本システムは、ユーザが自身の関心に依じて入力するキーワードに基づいて、Twitter 上での発言のやり取り（ツイート群）を収集・蓄積する。収集したツイート群の本文・返信関係を分析することで、議論構造、話題を表す特徴フレーズの共起ネットワークを生成する。収集・蓄積した議論事例を観察できる環境をユーザに提供することで、実際の議論の様子の把握を支援する。ユーザが本システムを利用して議論事例を観察する際には、まず、複数の議論構造の中で、観察したい議論構造を選択する。システムは、ユーザが選択した議論構造を視覚的に提示する。ユーザは、提示された議論構造を観察することで、議論の概況を把握することができる。

また、本システムは、ユーザが注目する議論構造に含まれる話題を表す特徴フレーズの共起ネットワークを提示する。ユーザは、この特徴フレーズを観察することで、その議論構造における主な話題を把握することができる。特定の話題に関する議論を観察したい場合には、その話題を表す特徴フレーズを共起ネットワーク上で選択する。システムは、ユーザが選択した特徴フレーズを含む発言がやり取りされている箇所（議論セグメント）を提示する（図5）。

このように、本システムは、ユーザの興味関心に依じて話題を選択しながら、実際の議論の様子を対話的に観察することを支援する。

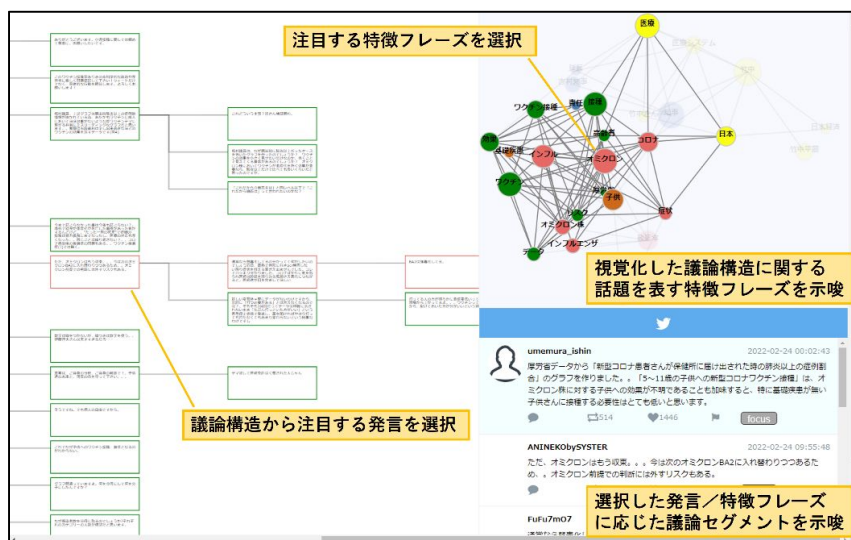


図5 ユーザの興味関心に依じた議論セグメントの提示

5. 主な発表論文等

〔雑誌論文〕 計6件（うち査読付論文 4件 / うち国際共著 0件 / うちオープンアクセス 0件）

1. 著者名 Hiroki Nakayama, Kota Chiba, Ryo Onuma, Hiroaki Kaminaga, Youzou Miyadera, Shoichi Nakamura	4. 巻 -
2. 論文標題 Visualization of Topic Distribution of Discussion on Social Media to Facilitate Experience of Novices	5. 発行年 2021年
3. 雑誌名 Proc. 2021 IEEE International Conference on Computing	6. 最初と最後の頁 366-371
掲載論文のDOI（デジタルオブジェクト識別子） 10.1109/ICOC053166.2021.9673540	査読の有無 有
オープンアクセス オープンアクセスではない、又はオープンアクセスが困難	国際共著 -
1. 著者名 Kota Chiba, Hiroki Nakayama, Ryo Onuma, Hiroaki Kaminaga, Youzou Miyadera, Shoichi Nakamura	4. 巻 -
2. 論文標題 Visualization of Diversified Opinions in Discussions on Social Media	5. 発行年 2021年
3. 雑誌名 Proc. 11th International Conference on Internet Technologies & Society	6. 最初と最後の頁 61-68
掲載論文のDOI（デジタルオブジェクト識別子） なし	査読の有無 有
オープンアクセス オープンアクセスではない、又はオープンアクセスが困難	国際共著 -
1. 著者名 千葉広汰, 中山祐貴, 大沼亮, 神長裕明, 宮寺庸造, 中村勝一	4. 巻 121(294)
2. 論文標題 議論事例理解のためのソーシャルメディア中の議論構造と話題分布の視覚化	5. 発行年 2021年
3. 雑誌名 電子情報通信学会 技術研究報告	6. 最初と最後の頁 41-46
掲載論文のDOI（デジタルオブジェクト識別子） なし	査読の有無 無
オープンアクセス オープンアクセスではない、又はオープンアクセスが困難	国際共著 -
1. 著者名 Kota Chiba, Hiroki Nakayama, Ryo Onuma, Hiroaki Kaminaga, Youzou Miyadera, Shoichi Nakamura	4. 巻 -
2. 論文標題 Methods of Extracting Useful Discussion Cases on Social Media for Promoting Novices' Experience	5. 発行年 2020年
3. 雑誌名 Proc. 2020 IEEE Conference on Big Data and Analytics	6. 最初と最後の頁 7-11
掲載論文のDOI（デジタルオブジェクト識別子） 10.1109/ICBDA50157.2020.9289709	査読の有無 有
オープンアクセス オープンアクセスではない、又はオープンアクセスが困難	国際共著 -

1. 著者名 千葉広汰, 中山祐貴, 大沼亮, 神長裕明, 宮寺庸造, 中村勝一	4. 巻 120(167)
2. 論文標題 発話間関係の分析に基づくTwitter上の有用議論事例抽出	5. 発行年 2020年
3. 雑誌名 電子情報通信学会 技術研究報告	6. 最初と最後の頁 57-62
掲載論文のDOI (デジタルオブジェクト識別子) なし	査読の有無 無
オープンアクセス オープンアクセスではない、又はオープンアクセスが困難	国際共著 -

〔学会発表〕 計4件 (うち招待講演 0件 / うち国際学会 0件)

1. 発表者名 中山祐貴, 千葉広汰, 大沼亮, 神長裕明, 宮寺庸造, 中村勝一
2. 発表標題 ソーシャルメディア議論における話題分布視覚化方法に関する検討
3. 学会等名 第46回教育システム情報学会全国大会
4. 発表年 2021年

1. 発表者名 千葉広汰, 中山祐貴, 大沼亮, 神長裕明, 宮寺庸造, 中村勝一
2. 発表標題 Twitter上の有用議論事例抽出と話題分布の視覚化
3. 学会等名 電子情報通信学会総合大会
4. 発表年 2021年

〔図書〕 計0件

〔産業財産権〕

〔その他〕

-

6. 研究組織	氏名 (ローマ字氏名) (研究者番号)	所属研究機関・部局・職 (機関番号)	備考

7. 科研費を使用して開催した国際研究集会

〔国際研究集会〕 計0件

8 . 本研究に関連して実施した国際共同研究の実施状況

共同研究相手国	相手方研究機関
---------	---------