

議論事例理解支援のためのソーシャルメディア上の議論様態の抽出

Extraction of Discussion Circumstances to Support Understanding Discussion Cases on Social Media

中山 祐貴^{*1}, 千葉 広汰^{*2}, 大沼 亮^{*3}, 神長 裕明^{*2}, 宮寺 庸造^{*4}, 中村 勝一^{*2}
Hiroki NAKAYAMA^{*1}, Kota CHIBA^{*2}, Ryo ONUMA^{*3}, Hiroaki KAMINAGA^{*2}, Youzou MIYADERA^{*4}, Shoichi NAKAMURA^{*2}

^{*1}山形大学 地域教育文化学部

^{*1}Faculty of Education, Art and Science, Yamagata University

^{*2}福島大学 共生システム理工学類

^{*2}Faculty of Symbiotic Systems Science, Fukushima University

^{*3}津田塾大学 学芸学部 情報科学科

^{*3}Department of Computer Science, College of Liberal Arts, Tsuda University

^{*4}東京学芸大学 教育学部

^{*4}Faculty of Education, Tokyo Gakugei University

Email: nakayama@e.yamagata-u.ac.jp, chiba@cs.sss.fukushima-u.ac.jp, r.onuma@tsuda.ac.jp, {kami, nakamura}@sss.fukushima-u.ac.jp, miyadera@u-gakugei.ac.jp

あらまし：PBLなどで学生が議論に取り組む機会が増加しているが、実際の議論事例を垣間見る経験が不足しているケースが少なくない。本研究では、ソーシャルメディア上の議論事例理解支援を目的として、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か（議論様態）を抽出する手法を開発する。本稿では、主に、議論様態抽出手法の概要について述べる。また、実際の発言データを用いた実験について報告する。

キーワード：議論遂行能力育成、議論様態、話題分布、ソーシャルメディア、情報視覚化

1. はじめに

PBLなど多くの主体的学習活動において、学生に議論を経験させる機会は増加している。このような議論を上手く遂行するためには、各メンバー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。しかし、他者との意見の交換を的確に行うスキルは、経験を通して学ぶところが大きく、その習得は未熟者にとって容易ではない。これに対して、議事録の分析に基づいて過去の議論の活発さを視覚化する試み⁽¹⁾、エージェントによるオンライン議論のファシリテーション支援に関する研究⁽²⁾などが報告されている。議論の活性度合の把握・向上、円滑な合意形成を目指した興味深い取り組みであるが、未熟者に議論経験のきっかけを与えるような支援は十分ではない。特に、「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か（議論様態）」を現実的に観察できる仕組みが必要である。

本研究では、ソーシャルメディア上の議論様態抽出・視覚化手法を開発する。これにより、未熟者の議論事例理解支援の実現を目指す。

2. 問題点と支援方針

2.1 未熟者の議論経験獲得に関する問題点

まず、ソーシャルメディア上の無数の会話の中から、参考となる会話を未熟者が発見することは難しい（問題点1）。また、ある議論における多様な意見を把握することは容易ではない（問題点2）。さらに、複数の発言の積み重ねによって構成される議論の流れを理解することが困難である（問題点3）。

2.2 支援方針

本研究では、まず、ユーザの関心に応じた議論を構成する会話（ツイート）を収集する手法（問題点1に対応）を開発する。次に、議論されている話題の分布を視覚化する手法（問題点2に対応）を開発する。さらに、発言と応答の積み重ねによる議論の流れ（議論構造）を抽出する手法を開発する（問題点3に対応）。

その上で、これらの手法を導入した支援システムを開発する。本研究の議論事例理解支援の概要を図1に示す。

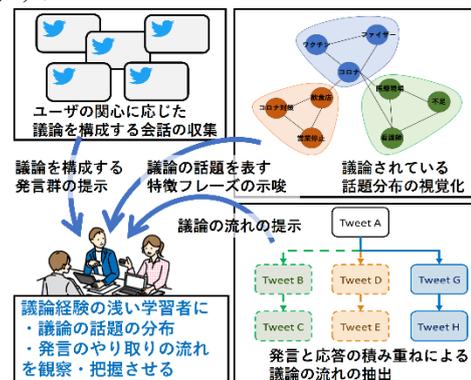


図1 議論事例理解支援の概要

3. 議論様態抽出手法

3.1 話題を表すフレーズの共起ネットワークの生成

本手法では、ソーシャルメディア上の議論事例の理解を支援するために、議論様態を抽出する。まず、関心のある話題についてのキーワードをユーザに捻

出してもらい、そのキーワードを基に TwitterAPI によるツイート検索を行うことで、キーワードと関連するツイートを収集する。また、収集したツイートデータのオブジェクト情報を分析して、リプライツイートを収集する。収集したツイートを基に議論を構成する発言群を抽出する。

次に、日本語自然言語処理オープンソースライブラリを用いて、抽出した発言群から、特徴フレーズを抽出する。その上で、議論を構成する発言中の特徴フレーズの共起状況を分析して、特徴フレーズのネットワークを生成する。さらに、モジュラリティに基づいて、共起ネットワークから特徴フレーズをクラスタリングすることで、特徴フレーズ群の組合せを抽出する。

3.2 議論構造の抽出

さらに、議論を構成する発言群（ツイート）のオブジェクト情報を分析して、返信関係による発言間のつながり（議論構造）を抽出する。抽出した議論構造の1つ1つはツリー状であり、最も古い日時の発言がルートノードである。また、各発言のノード同士は返信関係に基づくエッジでつながっている。

ここで、ルートノードからある発言のノードまでのつながり（パス）に注目する。このパスを構成するルートとエッジのまとまりを議論セグメントと呼ぶ。この議論セグメント中の特徴フレーズの出現状況を分析することで、議論セグメントと話題を表す特徴フレーズの組合せを対応付ける。

その上で、対応付けた議論セグメントと特徴フレーズの組合せ（議論様態）をユーザに示唆することで、話題に応じた発言のやりとりの観察を支援する。

4. 実験と考察

4.1 実験概要

議論様態抽出手法の有効性の検証を目的として実験を行った。本実験では、本手法の抽出工程ごとに検証を実施した。まず、想定話題として「コロナ問題」を設定した上で、Twitter上の会話から、議論がある程度盛り上がり、想定話題に関する議論事例17件を収集し、そこからツイート群1478件を抽出した。

これらのツイート群を対象として、話題を表す特徴フレーズ、特徴フレーズ群の組合せ、議論構造の各抽出、および、議論セグメントと特徴フレーズの組合せの対応付けを、提案手法を用いてそれぞれ実施した。

一方、ツイート群を目視で精査し、各処理の正解データを作成した。最後に、提案手法による抽出結果と正解データを比較し、一致状況を確認した。

4.2 結果と考察

話題を表す特徴フレーズの抽出では、抽出結果から再現率、適合率、F値を算出した。それぞれ、0.54, 0.60, 0.57 という結果となった。抽出した特徴フレーズの抽出回数を精査したところ、「感染」、「ワクチン」、

「コロナ」、「医療」、「検査」等、概ね、設定した話題である「コロナ問題」と関わりの強いフレーズを抽出できていたことを確認した。

特徴フレーズ群の組合せ抽出では、抽出結果に対して、クラスタリングの評価尺度（Global Purity, Inverse Purity, F 値）を用いた。それぞれ、0.81, 0.97, 0.88 という結果となった。例えば、「病院, 医療崩壊, 病床, ...」, 「感染, 濃厚接触者, オミクロン, ...」など、1つの話題において出現して然りと考えられる特徴フレーズの組合せを抽出できたことを確認した。一方で、特徴フレーズの数が少ない組合せも存在しており、組合せを構成する特徴フレーズの最低件数を検討する必要がある。

議論構造の抽出では、抽出したツイート群、ツイート間の返信関係によるつながりについて、適合率、再現率、F 値を算出した（表1）。抽出結果から、概ね意図通り議論構造を抽出できていた。しかしながら、TwitterAPI では一部収集できないようなツイートが存在したことが確認できた。

表1 議論構造の抽出結果

	ツイート抽出	つながり抽出
適合率	0.62	0.61
再現率	1.00	1.00
F 値	0.76	0.75

最後に、議論セグメントと特徴フレーズの組合せの対応付けでは、組合せごとの適合率、再現率、F 値を求めた上で平均値を算出した。それぞれ、0.87, 0.83, 0.85 という値となった。概ね、意図通りの対応付けがなされていることを確認した。一方で、ツイートの文章の少なさや返信先のツイートの影響を過剰に受けたことで、想定する特徴フレーズの組合せとは異なる対応付けがなされた議論セグメントも存在した。これについては、当該ツイートの対応付けの際のルールなどを検討する必要がある。

今回の実験は、限定的なデータ数での実施であるため、多様な想定話題を用意した上で、実践的な検証を継続的に実施する必要がある。

5. おわりに

本稿では、議論様態抽出手法について述べた。また、実際の議論データを用いた実験に基づいて、提案手法の有効性について考察した。今後は、支援システムの実践評価の検討を進めたい。

参考文献

- (1) 村岡泰成, 石川誠彬, 尾澤重知, 江木啓訓: “議論の活発さを反映した共起ネットワーク図の提示による効果の検討”, 情報処理学会研究報告, Vol. 2020-GN-110, No.8, pp.1-8 (2020)
- (2) Rafik Hadfi, Jawad Haqbeen, Sofia Sahab, and Takayuki Ito: “Argumentative conversational agents for online discussions”, *Journal of Systems Science and Systems Engineering*, Vol.30, pp.450-464 (2021)