

複数人議論における取りまとめ役の分類と特徴の分析

塩田 宰[†] 山村 崇^{††} 嶋田 和孝^{††}

[†]九州工業大学 ^{††}九州工業大学大学院

{t.shiota[†], t.yamamura^{††}, shimada^{††}}@pluto.ai.kyutech.ac.jp

1 はじめに

議論は様々な考えや背景を持つ人が集まって意思決定や意見交換を行う対話行為である。参加者の文化的・社会的背景によって話者の発言力の関係に差が出る、あるいは特定の話者からの意見が得られなくなるなどして議論が停滞することがあり、議論を円滑に進行するためには取りまとめ役の存在が重要である。

本研究室では人間に代わって議論の円滑な進行やトピックの制御を行うデジタルファシリテータの構築を目指している。これによりファシリテーションの経験のある人間を必要とせず、誰でも建設的な議論を行えるようになる。その実現のためには人間の取りまとめ役の特性を分析することが重要である。

そこで本論文では、議論参加者から得られる特徴を基にした取りまとめ役の分類モデルを構築し、複数人議論の取りまとめ役の特徴について分析する。

2 関連研究

名古屋工業大学の COLLAGREE[1] は公共事業などのためのオープンな環境での多人数の意見を集約することを目的として開発されており、西田ら [2] はそれを用いた社会実験により得られたデータを対象としてファシリテータの発言の文脈特徴について分析している。また、大本ら [3] は非言語情報とパラ言語情報に着目して、ファシリテータがどのような情報に基づいてファシリテーション行動を行っているのかを分析している。それに対して本研究は、企業などで行われているミーティングのように複数人での対話環境における合意形成をサポートすることを目的として、取りまとめ役の発言の時間情報や談話行為など発言の持つ行為や傾向に特徴があるのかを分析している。

3 言語資源

本節では、取りまとめ役の分類および特徴分析を行うために対象とする2つの言語資源について述べる。

3.1 Kyutech コーパス

Kyutech コーパス¹[4] は、4人の参加者による複数人対話を対象としている。4人の参加者は、ある都市

¹<http://www.pluto.ai.kyutech.ac.jp/~shimada/resources.html>

のショッピングモールの経営者であるという設定のもと、そのショッピングモールのレストラン街にあるレストランの閉店後に来店させるレストランを3つの候補の中から1つ選ぶというタスクについて議論する。議論を行うために必要な来店候補のレストランの情報、閉店するレストランの情報と閉店理由、レストラン街の既存店の情報、ショッピングモールの立地や来客者の時間帯・性別分布、ショッピングモールのある都市の人口などの統計情報、さらに隣接する市町村の情報などが書かれた10ページほどの資料が準備されている。議論参加者はこの資料を10分間黙読した後に20分間の議論を行った後、来店レストランを1つ決定する。現在4つのシナリオが準備されており、計9対話収録されている。またKyutech コーパスの書き起こされた各発話には話者IDや時間情報に加えて、人手によりアノテーションされたトピックタグと談話行為タグが存在する。

トピックタグとは各発話が言及している話題を示すタグである。Kyutech コーパスには閉店したレストランについて言及している発話に付与される「Closed」や複数の候補店に関連する発話に付与される「Exists」など、計28種類のトピックタグが用意され、各発話に付与されている。

談話行為タグとは各発話がどのような機能を持つかを示すタグである。Kyutech コーパスには談話行為の国際標準化規格ISO24617-2 [5] に対応した計22種類の談話行為タグが用意され、各発話に付与されている。

本研究で扱う日本語の発話単位は発話単位ラベリングマニュアル [6] により定義されている長い発話単位とする。長い発話単位の定義は「話し手と聞き手が行為や情報を交換する際の基本単位に相当し、統語的・談話的・相互行為的な一まとまり」である。公開されている書き起こしデータをマニュアルに従って長い発話単位に変換したものを使用する。

3.2 AMI Corpus

AMI Corpus [7] はあらかじめ議題の与えられた状態で行われるシナリオ会議と、議題の与えられていない状態で行われるシナリオ無し会議の2種類の複数人対話を収録したコーパスで、それぞれ人手による書き起こしデータと機械による書き起こしデータが公開さ

れている。本研究は、この中でもシナリオ会議の人手による書き起こしデータ (135 対話) を対象とする。シナリオ会議は架空の家電企業に務める異なる役割の 4 人の従業員が、市場に出回っているユーザに不親切で魅力の無い古いデザインのテレビリモコンの代わりとなる新しいテレビリモコンを開発するという議題で、計 4 回行われる。この会議で 4 人の従業員に与えられる役割はプロジェクトマネージャー (以降, PM), マーケティングエキスパート (以降, ME), ユーザーインターフェースデザイナー (以降, UI), インダストリアルデザイナー (以降, ID) の 4 種類である。PM はプロジェクトが決められた時間と予算の間でのプロジェクトの実行, および会議の運営や議事録・報告書の作成の役割を果たす。ME はユーザの要求や市場動向を調査し, プロトタイプを評価する役割を果たす。UI は開発するリモコンの提供する機能やインターフェースについて責任を持つ役割を果たす。ID は開発するリモコンの構成要素の決定と設計と同時に, インターフェースについても考える役割を果たす。Kyutech コーパスと同様に, AMI Corpus にも人手によりアノテーションされたトピックタグと談話行為タグが存在する。

AMI Corpus には会議の開始周辺の発話に付与される「Opening」やプロトタイプの評価に関する発話に付与される「evaluation of prototype(s)」など, 計 24 種類のトピックタグが用意され, 対応する発話に付与されている。加えて, 何らかの情報を提供している発話に付与される「Inform」や相槌などの発話に付与される「Backchannel」など計 15 種類の談話行為タグが用意され, 対応する発話に付与されている。

本研究で扱う英語の発話単位は「。」や「?」といった文の終わりを表すパンクチュエーションを基準に区切った書き起こしデータとする。

4 分類モデルの構築

本節では, 機械学習による分類モデルの構築および使用する素性について述べる。

4.1 取りまとめ役の分類モデル

本研究では, 対話の中に現れる言語情報を素性とした決定木 (CART[8]) による取りまとめ役と非取りまとめ役を 2 値分類するモデルを構築する。構築したモデルは次節の分類実験および人手による取りまとめ役の特徴の分析に利用する。

4.2 素性

- 自身 (および他者) の発話の繰返しの相対的割合 (self_repetition_ratio, other_repetition_ratio)

杉山 [9] はテレビ番組内の進行役とゲストの発話に現れる繰返しについて考察している。繰返しとは先行する発話をそのまま再現した発話や, 意味を保持した状態での言い換えや要約の発話としている。ゲストの

繰返しは自身の発話を繰返すことが多く, それにより自身の情報をわかりやすく伝える働きを担っていると述べられている。一方で進行役の繰返しはゲストの発話を繰返すことでゲストから情報を引き出す, 対話のトピックのコントロールや円滑な進行を実現すると述べられている。そこで本研究ではこの「自身の発話の繰返しの相対的割合」および「他者の発話の繰返しの相対的割合」を抽出する。

まず対話中に出現する内容語を抽出し, 各発話を内容語の頻度ベクトルに変換する。変換した各発話を基準として後ろに続く 10 発話の中で \cos 類似度が 0.6 以上となる発話を繰返しとして抽出する。繰返し元の発話が同じ話者によるものなら「自身の発話の繰返し」, 異なる話者によるものなら「他者の発話の繰返し」として扱い, それぞれの総数に対する各話者の繰返し数の割合を算出し素性とする。

- 対話中のトピックを網羅した割合 (join_topic_ratio)

取りまとめ役は議論全体を円滑に進行する役割を担っていることから, 対話の様々な意見交換の場面において発言をすることが考えられる。つまり, 取りまとめ役は対話内に存在するそれぞれのトピックで発言をすることが考えられる。そこで対話の全てのトピックを抽出し, 各話者が各トピックにおいて発言したか否かを抽出する。そこから全トピック数に対する各話者の発言したトピック数の割合を算出し素性とする。

- Meeting タグの発話の相対的割合 (meeting_ratio)

取りまとめ役は対話の運営に責任を持つ役割があることから, 議論の進行に関連するトピックではより多くの発言を行うことが考えられる。そこで対話から Kyutech コーパスに存在するトピックタグの一つである Meeting タグの発話を抽出し, その総数に対する各話者の発話数の割合を算出し素性とする。Meeting タグとは話を進めるための議事提案や最終決定部に関連するトピックの発話に付与されるタグである。なお AMI Corpus に同様の意味を持つトピックタグは存在しないため, AMI Corpus からこの素性は抽出しない。

- 特定の談話行為の発話の相対的割合 ([DIALOGUE_ACT_NAME]_ratio)

取りまとめ役は他者から意見をくみ取る, 他者の意見を傾聴するなどの行為を非取りまとめ役より多く行うことが考えられる。そこで Kyutech コーパスおよび AMI Corpus の中で特徴が顕著に現れると考えられる談話行為タグについて発話を抽出し, 各談話行為の発話数に対する各話者の発話数の割合を算出し素性とする。本研究では Kyutech コーパスについては「Q(情報の要求)」「An(情報要求に対する返答)」「Inf(情報の提供)」「Su(行為の提案)」「PF(肯定)」の 5 つ, AMI Corpus については「Backchannel(相槌)」

表 1: Kyutech コーパスにおける分類の評価値

	精度	再現率	F 値
取りまとめ役	0.778	0.700	0.737
非取りまとめ役	0.889	0.923	0.906

「Inform(情報の提供)」「Elicit-Inform(情報の提供を引き出す)」「Assess(物事の評価)」「Elicit-Assess(物事の評価を引き出す)」「Elicit-Offer-Or-Suggestion(行為や提案を引き出す)」「Suggest(提案)」の7つについて算出し素性とする。

- 対話および4分割した中での発話の相対的割合 (utterance_num_ratio, t11, t12, t13, t14)

取りまとめ役は議論をコントロールするために様々な発話を行うことから、非取りまとめ役と比較して発話量が増えることが考えられる。また議論の発散や収束の場面など、状況によって発話するタイミングが決まっていることが考えられる。そこで対話全体の発話数に対する各話者の発話数の割合、および対話を時間で4つに分割したときのそれぞれの発話の総数に対する各話者の発話の割合を算出し素性とする。

- 平均発話文字数・平均発話時間 (ave_strlen, ave_time)

取りまとめ役は議論の中で出現した意見や情報をまとめるなど他の話者と比較した際に発話の長さが長くなることが考えられる。そこで各話者の発話から平均発話文字数および平均発話時間を算出し素性とする。

- 沈黙後に行われた発話の相対的割合 (break_silence)

取りまとめ役は新たな意見が出ないときや対立等で議論が停滞した場合に、再びその議論を活性化させることが考えられる。そこで一つ前の発話から10秒以上経過してから行われた発話を対話から抽出し、その総数に対する各話者の発話数の割合を算出し素性とする。

5 実験と分析

本節では、各言語資源に対する分類手法の適応、およびモデルから得られた取りまとめ役の特徴について述べる。

5.1 Kyutech コーパスによる実験

Kyutech コーパスの36話者分のデータを対象に9対話の交差検定による分類実験を行った。この実験での取りまとめ役はKyutech コーパスの被験者に実施されたアンケート「誰が議論をコントロールしていたと思いますか?」の回答で2人以上から指名された話者とした。(取りまとめ役10人、非取りまとめ役26人)

表1に9対話交差検定によって得られた分類の評価値を示す。

表 2: AMI Corpus における分類の評価値

	精度	再現率	F 値
取りまとめ役	0.654	0.630	0.642
非取りまとめ役	0.878	0.889	0.883

5.2 AMI Corpus による実験

AMI Corpus の540話者分のデータを対象に10分割交差検定による分類実験を行った。この実験での取りまとめ役はAMI Corpus に付与されている役割の一つであるPMとした。(取りまとめ役135人、非取りまとめ役405人)

表2に10分割交差検定によって得られた分類の評価値を示す。

5.3 特徴の分析

図1にKyutech コーパスを用いた実験で得られた木の例、図2にAMI Corpus を用いた実験で得られた木の例を示す。それぞれのノードには分類に用いられる素性と閾値、ノードに所属する話者のデータ数 (speaker_num), ノードに所属する話者の役割の内訳²(role_num) が示されている。

図1の木では9人いる取りまとめ役の内8人が、情報要求の発話の割合 (Q_ratio) が0.3095より大きく、平均発話時間 (ave_time) が1.957より長いという条件で最右下のノードに分類されている。この結果から取りまとめ役は「情報の要求に関する発話の割合が相対的に多く、平均発話時間が長い」という特徴を持っていることがわかる。また、図2の木において同様の分析をすると取りまとめ役は「対話全体での発話の割合 (utterance_num_ratio) が相対的に多く、行為や提案を引き出す発話の割合 (elicit-offer-or-suggestion_ratio) が相対的に多く、対話中のトピックの網羅率 (join_topic_ratio) が高い」という特徴を持っていることがわかる。

本研究で対象としたKyutech コーパスとAMI Corpus の収録条件には、予め誰が取りまとめ役かが明示的であるか否かという大きな違いがある。その違いにより対話の取りまとめ役に異なった特徴があるかを分析する。図3および図4にそれぞれの実験で構築された分類モデルに使用されている素性の深さごとの分布を示す。

図3より、深さが0のところでは「情報要求の発話の割合 (Q_ratio)」が分類する際の大きな特徴となっていることがわかる。また図4より、深さが1のところでは「行為や提案を引き出す発話の割合 (elicit-offer-or-suggestion_ratio)」や「情報を引き出す発話の割合 (elicit-inform_ratio)」が分類する際の大きな特徴となっていることがわかる。以上の2つの結果から、

²左が非取りまとめ役、右が取りまとめ役の数を示している。例えば、図1の最右下の[0,8]は8人中8人が取りまとめ役であることを示している。

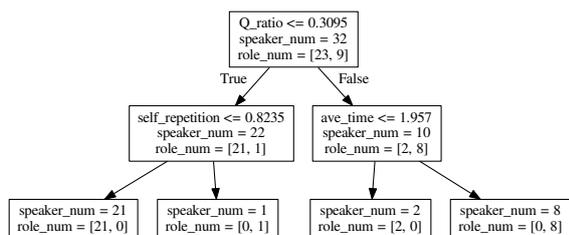


図 1: Kyutech コーパスで構築された木の例

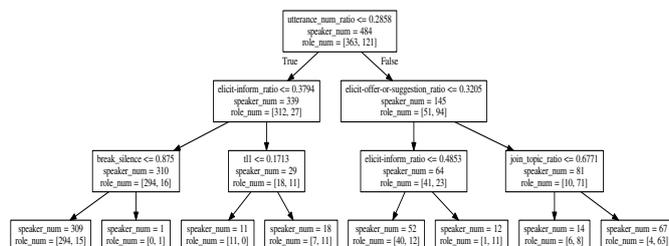


図 2: AMI Corpus で構築された木の例

取りまとめ役が明示的であるかどうか関係なく共通して取りまとめ役は他の話者から情報や行為、提案など言動を引き出す行為を相対的に多く行うことが示唆される。

一方、図 3 の Kyutech コーパスによる実験で得られた分布のみに着目すると、自身の発話の繰返しの相対的割合 (self_repetition) および平均発話時間 (ave_time) の 2 つが取りまとめ役を分類する際の大きな特徴となっている。このことから、対話に参加している話者が議論をコントロールする役割を担うためには比較的長い時間発言権を獲得すること、自身の発言を対話に反映させるために自身の行った発話を多く繰り返すことが必要であると示唆される。

続いて、図 4 の AMI Corpus による実験で得られた分布のみに着目すると、議論序盤や終盤、対話全体での発話の相対的割合 (tl1, tl4, utterance_num_ratio)、対話中のトピックの網羅率 (join_topic_ratio)、沈黙後に発言を行った相対的割合 (break_silence) が取りまとめ役を分類する際の大きな特徴となっている。このことから、最初から役割が定まった状態の議論では、取りまとめ役の役割の話者は議論序盤の発散場面や終盤の収束場面で多くの発話を行い、様々なトピックに介入することで議論をコントロールしていること、沈黙が発生した際には議論を活性化させるように発言することが示唆される。

6 おわりに

本論文では、複数人対話における取りまとめ役の分類および分類モデルから取りまとめ役の特徴の分析を行った。今後は、現在の素性に加えて視線や姿勢、音声などの非言語情報をから得られる特徴を素性として導入し、素性のマルチモーダル化を図る。また、本研究は対話というマクロな観点での特徴の分析を行ったが、デジタルファシリテータ実現のためにどのような

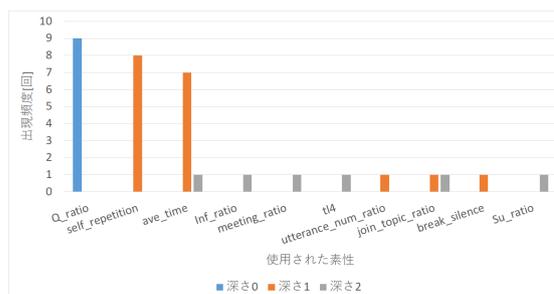


図 3: Kyutech コーパスの実験での素性の分布

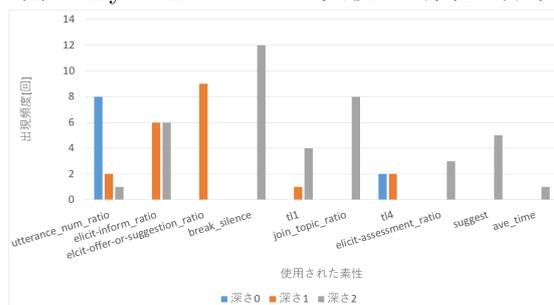


図 4: AMI Corpus の実験での素性の分布

タイミングでどのような言動を行うかなどミクロな観点での取りまとめ役の特徴の分析も行っていきたい。

謝辞

本研究は科研費 17H01840 の助成を受けたものです。

参考文献

- [1] 伊美裕麻, 伊藤孝行, 伊藤孝紀, and 秀島栄三. 大規模意見集約システム COLLAGREE の開発と名古屋市次期総合計画に関する社会実験. 人工知能学会全国大会論文集, 28:1-4, 2014.
- [2] 西田拓哉, 白松俊, 伊藤孝行, and 藤田桂英. Web 議論システムにおけるファシリテータ発言の分析のための文脈特徴. 情報処理学会第 78 回全国大会講演論文集, 2016(1):555-556, 2016.
- [3] 大本義正, 戸田泰史, 植田一博, and 西田豊明. 議論への参加態度と非言語情報に基づくファシリテーションの分析. 情報処理学会論文誌, 52(12):3659-3670, 2011.
- [4] T. Yamamura, K. Shimada, and S. Kawahara. The Kyutech corpus and topic segmentation using a combined method. In *Proceedings of the 12th Workshop on Asian Language Resources*, 2016.
- [5] H. Bunt, J. Alexandersson, J. W. Choe, A. C. Fang, K. Hasida, V. Petukhova, A. Popescu-Belis, and D. R. Traum. ISO 24617-2: A semantically-based standard for dialogue annotation. In *Proceedings of the 8th International Conference on Language Resources and Evaluation*, pages 430-437, 2012.
- [6] Japanese Discourse Research Initiative. 発話単位ラベリングマニュアル version 2.1, 2017.
- [7] J. Carletta. Unleashing the killer corpus: experiences in creating the multi-everything ami meeting corpus. *Language Resources and Evaluation Journal*, 41(2):181-190, 2007.
- [8] L. Breiman, J. Friedman, C. J. Stone, and R. A. Olshen. *Classification and regression trees*. CRC press, 1984.
- [9] 杉山ますよ. 進行役とゲストの発話にみられる繰り返し. 言語文化と日本語教育, 16:46-57, 1998.