

議論における発言間の階層関係に基づく情報構造化とその応用

Information Structuralization and Its Application
Based on Hierarchical Relationship of Utterances in Discussion

三浦 寛也 *¹

Hiroya Miura

竹川 佳成 *¹

Yoshinari Takegawa

平田 圭二 *¹

Keiji Hirata

*¹公立はこだて未来大学

Future University Hakodate

In this research, we developed a system that generates minutes with repeating Q&A between a user and the system. The aim of the research is to provide the interactive analysis of the meeting records with a specific purpose or viewpoint for a user. To realize the system, we proposed a novel tree structure dedicated to represent the discussion structure of a meeting based on the relative importance of utterances, by analyzing semantic structure of meeting records.

1. はじめに

近年、様々な分野において、電子化された膨大な情報が蓄積されている。知識管理の分野では、デジタル化されたマルチメディアデータを容易に取り扱うことが可能になり、音声・映像を議事録に組み込み可能なシステムも増えてきた [1]。これらに含まれる情報は膨大であるため、蓄積されたデータに対する多様な視点や、柔軟なものの見方への重要性が主張されてきた。しかし、こうした技術の確立は、マルチメディアに含まれるメタデータの考慮や、利用時の文脈に基づく動的な情報の再構成といった点において、計算理論による大量のデータと議論構造の理解との結びつけが一般に困難である [2]。また、徳永は、ユーザの持つ問題（情報要求）を解決するための情報を抽出する情報検索において、人間の認知過程の工学的な定式化の難しさを指摘した [3]。

上記で述べた問題を解決するためには、マルチメディアデータを含む会議コンテンツを対象とした、人間の意図や思考、認知過程を考慮した意味理解とそれに基づくパーサの開発が求められる。言語や音楽などの時系列メディアを介した意思疎通では、人は意図を込めて、記号や数学的実体（関数と集合）を記述するオブジェクト列を表出する。その意図は、Noam Chomsky を始祖とする階層的順序構造（木構造）として表現されることが多い。そのため、時系列メディアである会議コンテンツにおける意図の表現には、木構造を用いる手法が有望であろう。我々はこれまで、人間の認知の観点から楽曲構造を論じる理論である Generative Theory of Tonal Music [4] のタイムスパン木を用いて、議論の意図を込めた表出を計算論的な立場から定式化を行い、会議記録における意図や意味の表現方法として、階層的順序構造に基づく構造化を行ってきた。本稿では、この議論構造を用いて、意図の抽出、要約、観点の切り替えといった柔軟な情報抽出の方法について述べる。

2. 木構造に基づく議論における意図の表現

2.1 議論タイムスパン木の表現

本研究では、時系列メディアである会議コンテンツにおける意図の表現方法として、木構造を用いる。人は会議の各発言の

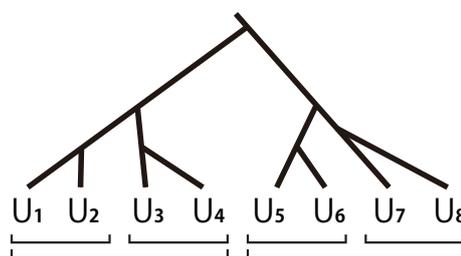


図 1: 議論タイムスパン木

間にある賛成と反対、質問と回答などの関係に基づき、隣接する発言間だけでなく離れた発言間の関係も考慮して、議論の構造を理解する。また、発言列には、ある観点から見て重要、あるいは瑣末なものがある。そのため、我々はその重要度に基づく発言列に関する簡約の概念を定義し、会議の発言列を分析して簡約の概念を内包する議論の木構造（議論タイムスパン木）を導入した（図 1）[5]。議論タイムスパン木によって、発言間の関係や集合に基づく会議内容や議論構造の理解や、段階を追った簡約化が可能となる。

議論タイムスパン木は、構造的に重要な発言が幹になるような 2 分木であり、以下のような内容を表示する：(i) 隣り合った関連のある発言を 1 つのグループにする、(ii) 各発言の重要度を階層的に表現する。表 1 は、8 件の発言から構成される話題のまとまりの例で、発言 U_1 から U_2 が生じ、 U_3, U_4, \dots, U_8 とさらなる発言が生じていることを表す。ここでは、開始発言から終了発言までの継続した発言群をセクションと呼ぶ。(i) に関して、発言 U_1 と U_2 、 U_3 と U_4 がそれぞれ 1 つのグループとなっており、さらに階層が上がると U_1-U_2 と U_3-U_4 がグループとして構成される。 $U_4 \sim U_8$ についても同様である。これは会議コンテンツにおける内容の近さを表現する。(ii) に関して、生成された木構造における枝と幹の関係から各発言の重要度が階層的に表現される。図 1 では、最も重要な発言が U_7 、その次に U_1 が重要であることが分かる。この重要度は半順序であり、どちらが重要なかわからないような発言の対が存在する場合は、2 つの発言間の関係はそれほど重要でないことを表す。

連絡先: 三浦寛也, 公立はこだて未来大学, 〒 041-8655 北海道函館市亀田中野町 116-2, 0138-34-6462, murh1016@gmail.com

表 1: 議論の一例

ID	発言者	発言要旨
U_1	O	危険ではない状況というのは、お互いに目と目があっている状況のことではないのだろうか。
U_2	W	お互いに目が合っていないでも大丈夫。目が合うというか、認識できているかが問題である。
U_3	O	ずっと認識する必要はないが、一度は相手が何処にいてどの方向に動かかを知る必要がある。
U_4	W	認識して回避することは可能だが、双方が認識できていないというのが先ほどの主張である。
U_5	N	AT が人を認識していても、次にとる行動を予測しないと認識して回避すると言えないのでは。
U_6	W	相手がこちらを認識していない時は、行動を予測できないと思うが、そこは従来研究に譲る。
U_7	N	相手を認識したら、AT がやるべきは回避ではなくて人間に AT の存在を知らせることである。
U_8	W	仰る通り、安全走行のためには、そういうことに気をつけることも必要だと思う。

表 2: GPR 一覧

項目	パラメータ
2	発言間隔
3a	発言順序
3b	発話量変化
3c	発言時間長変化
4a	発表者の発言
4b	重要単語の初出箇所

表 3: SPR 一覧

項目	パラメータ
2a	導入/終了発言
2b	重要単語の初出箇所
3a	発言時間長
3b	発話量
3c	分岐数
4a	賛同数
4b	社会的ステータス
4c	発表者の発言
4d	重要単語を含む発言

2.2 議論タイムスパン木の生成ルール

議論タイムスパン木は、テキスト情報やそのメタデータを含む会議コンテンツの情報に基づき、以下 2 項目のルールに基づきボトムアップに生成される。(1) 議論に含まれるグループ(ゲシュタルト)を抽出するルール (Grouping Preference Rule: GPR), (2) あるグループ全体の時間幅(タイムスパン)を代表する重要な発言を同定するルール (Significance Preference Rule: SPR)。(1) では、「連続した 4 発言の間隔 (b_1, b_2, b_3) において、 b_2 が最も大きい値の場合、境界を付与する」といった 6 項目のルールから構成される。(2) では、「連続した 4 発言 (u_1, u_2, u_3, u_4) において、最も発話量の高い発言に重要度を付与する」といった 9 項目のルールから構成される。GPR および SPR の項目一覧を表 2・3 に示す。

議論タイムスパン木の生成に関して、各ルールの適用順序が決まっていないため、ルール間の競合が起きるという問題があった。本研究では、ルールの重み値を制御するためのパラメータを導入することで、これらの問題を解決する。各パラメータは、値域を 0.0~1.0 となるよう正規化し、重み値を調整している(初期値: 1.0)。選好ルールの適用には自由度があり、どのように適用するかはこのパラメータを用いて、分析者の判断に委ねられる。

3. 議論タイムスパン木による情報抽出

本章では、表 1 の議論例を題材とした議論タイムスパン木の生成による意図の抽出、要約、観点の切り替えといった柔軟な情報抽出を行う方法について述べる。

3.1 議論に含まれる意図の抽出

図 1 の議論タイムスパン木は、ルールの重みの値を初期値である 1.0 に統一したベースラインとなる木構造である。木構造の生成によって得られるグループ構造や木構造の枝の付く位置などを考慮することで、さまざまな議論情報が抽出できる。

例えば、生成された議論タイムスパン木をトップダウンに探索すると、2 分木である議論タイムスパン木から大域的なグループ構造 ($U_1 \sim U_4, U_5 \sim U_8$) が抽出され、議論を 2 つのまとまりに分割した意味のまとまりが同定される。また、ボトムアップに探索すると、局所的なグループ構造 ($U_1-U_2, U_3-U_4, U_5-U_6, U_7-U_8$) が抽出される。これにより、質問とその応答が隣接対の構成となった質問応答対の同定が期待できる。

さらに、議論タイムスパン木によって得られた大域的なグループでの初出發話は、グループ内での新しいトピックを導入した発話と解釈できる。そのため、導入的な発話 (U_1) や大域的なグループ $U_5 \sim U_8$ における初出發話 (U_5) は、質問や問題提起を表現する発話として同定できる。このように、各発話が内包している質問応答や問題提起といった発話意図が表出される。

3.2 簡約による抜粋型要約の生成

議論タイムスパン木の簡約とは、重要な発言と装飾的な発言を分離するもので、構造的に重要な発言が幹になるような 2 分木を求める分析である。この機構によって、要約対象全体であるドキュメントを単語や文などの単位に分割し、それらの文を抽出する方法である抜粋型要約が実現できる。図??のように得られた議論タイムスパン木に対して簡約を行い、各発言のレベルを揃える。これにより、例えば、議論タイムスパン木の上層の断面 (Level a) を切ると、 U_1 と U_7 の 2 発言による要約が生成される。次に下層の断面 (Level b) を切ると、 U_1, U_3, U_5, U_7 の 4 発言による要約が生成される。このように、簡約レベルを任意に切り替えることができる要約が生成される。これは、段階を追った議論情報の理解に有効である。

3.3 観点を切り替えた異なる目的への対応

先述したルール群の重み値が調整可能なパラメータの設計は、異なる目的や観点による情報構造化や重要発言の抽出を実現する。図 1 の議論タイムスパン木は、ルールの重みの値を一律にしたベースラインとなるものであった。この議論タイムスパン木における要約は、質問 (U_1, U_5) や意見 (U_3, U_7) といった発言を満遍なく抑えた点において、事実重視の観点から生成された要約と言えよう。その一方で、例えば「発言時間間隔の長い箇所で境界が生じる」「プレゼンタ(表 1 では W)の発言は重要度が高い」といったルールを重視すると、話題の提起や展開性といった観点から木構造が生成される。このように、各ルールの重み値を調整し、観点を切り替えることで、異なる目的に即した要約を生成することができる。

表 4: 各機能とその実装方法

番号	機能項目	実装方法
F_1	2つのまとまりに分割	GPR 適用によるグルーピング構造分析の結果から、セクションにおける最も深い境界を抽出し、その前後で得られた大域的なグループ構造に基づき、議論を2つのまとまりに分割する。
F_2	質問応答対の同定	局所的なグループ構造に基づき、議論を最小単位のグループまで分割する。
F_3	主題を含む発話の同定	GPR 適用によるグルーピング構造分析の結果から得られる大域的なグループ構造における最初の発話を出力する。
F_4	要約	生成された議論タイムスパン木による内部情報に基づき、上位発話を抽出する。さらに、任意の発話数を選択し、簡約レベルの切り替えができる。
F_5	結論までのプロセス提示	分析結果から得られた大域的なグループ構造の先頭の発話を始点、最重要発話を終点とする一連の発話群を出力する。
F_6	観点の切り替え	表 2・3 の各ルールに関する重み値を値域内で任意に調整する。パラメータ調整によって、議論タイムスパン木を変形させた後、 $F_1 \sim F_5$ の機能を同様に用いることができる。

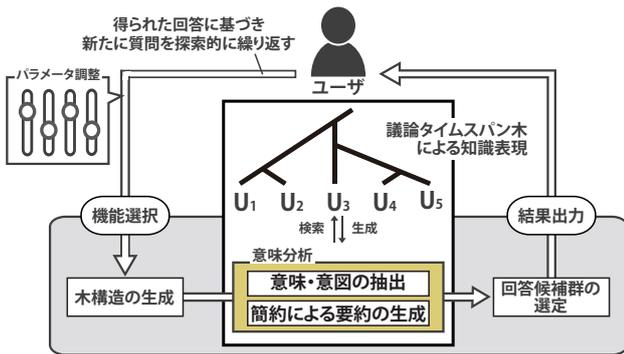


図 2: システム構成図

4. 対話的議事録生成システム

4.1 システム概要

本節では、議論タイムスパン木によって表現される議論構造に基づき、システムとユーザがインタラクティブに質問応答を繰り返しながら議事録を生成するシステムについて述べる。本システムの全体構成を図 2 に示す。ユーザは、会議コンテンツに対して様々な問い合わせを行い、その結果を受け、さらに新たな問い合わせを繰り返すことで、会議をパーソナルな視点から動的に振り返ることができる。探索的な分析においては、重要となる事柄や注目すべき事象が必ずしも明確ではない。そのため、様々なインタラクションを可能にし、ユーザが自由に試行錯誤できる環境が望まれる。本システムでは、目的や観点、議論における重要発言の簡約レベルを効果的に切り替えることで、ユーザの要求に柔軟に対応することを目指している。また、ユーザは情報探索過程において、議論タイムスパン木を生成する際に用いるパラメータ調整の操作によって、情報探索のための新たな視点を変更することで、情報要求を精緻化していくことができる。本システムでは、これまでに 6 項目の機能を実装した。以下に各機能の項目とその実装方法を示す(表 4)。

4.2 情報探索過程の具体例

本節では、提案システムを用いて、表 1 の議論例によって生成される議論タイムスパン木(図 1)の情報に基づき、情報探索を行う過程を示す。まず、8 発言の中から概括的にどの部分に着目すべきかを判断するため、「この議論を 2 つのトピックに分割したい」という問い合わせを行うことが予想される。この問い合わせは F_1 と対応づけられ、その分析結果から前半のトピックが $U_1 \sim U_4$ 、後半のトピックが $U_4 \sim U_8$ ということが分かる。さらに細かく分割し、質問応答の対を把握したい場

合は、 F_2 を用いて局所的な構造を確認することができる。分割された各トピックにおいて、 F_3 を用いることで、「何を主題としているか」といった問い合わせに対処することができる。この問い合わせから、 U_1 と U_4 が、それぞれ質問と問題提起を表現する発言という分析結果が得られ、実際の発言を確認することができる。

この議論における全体の流れを把握するためには、要約を生成し、閲覧することが効果的であろう。 F_4 を用いることで、「この話題の要約を 4 発言程度でまとめてほしい」といった問い合わせや、簡約レベルの切り替えにより「さらに 2 発言に絞るとどうか」といったものに回答できる。この結果から、4 発言による要約では U_1, U_3, U_5, U_7 、2 発言による要約では、 U_1, U_7 の発言群が抽出され、本議論の大きな流れを段階を追って理解することができる。さらに、「最も盛り上がった箇所を教えてほしい」という質問は、 F_5 によって解決され、分析結果から本議論における総論とその総論を引き起こした発言まで遡った U_5, U_6, U_7 という一連の発言群を把握できる。ユーザはこれらの機能を用いることで、異なる視点から議論の大域的な展開を捉えることに役立つ。ここで、ある特定の人物の立場から議論を構造化する状況を仮定する。例えば、「発表者視点の要約を知りたい」という問い合わせがあった場合、 F_6 を用いて、ルールの重み値に関するパラメータを調整し、得られた木構造から機能 4 を生成することができる。このように、各機能を用いて質問応答を繰り返すことで、会議コンテンツに対する探索的な分析が可能となる。

5. 実験結果

情報検索システムの性能評価には、有効性と効率性の 2 つの側面がある [6]。有効性とは、システム利用によって必要な情報を完全かつ正確性に検索できるかどうかを観測する指標であり、これらは、正解を判定した評価用データを用いて、再現率と適合率によって計算される。そのため、本研究では、情報検索システムの有効性の評価に関して、議論タイムスパン木の自動生成に関する再現率と適合率を評価した。

本論文では、ゼミを対象とした会議風景のマルチメディアデータを含む議論内容のコンテンツによる 10 件分の議論を対象に、提案システムに対する有効性を検討した。各議論は、マルチモーダルインターフェース、セマンティック Web、自然言語処理に関する内容で、名古屋大学長尾研究室において実際に実施されたものである*1。本議論は、総時間が 22 時間 15 分の長さで、平均発言者数は 9.5 名、セクション総数は 169 件、総発言数は 758 件である。

*1 長尾研究室:ディスカッションマイニングプロジェクト, 入手先 | <http://dm.nagao.nuie.nagoya-u.ac.jp/>.

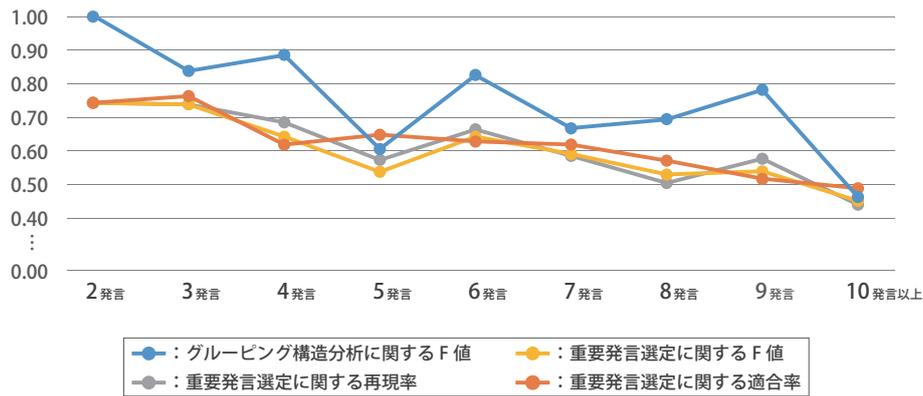


図 3: 発言数に対するシステム出力のグルーピング構造獲得分析と重要発言選定の精度

議論タイムスパン木の精度は、ルールの重みの値を一律にしたベースラインとなる木構造を対象とし、議論に含まれるグループの抽出(グルーピング構造分析)と、あるグループを代表する重要な発言の同定(重要発言選定)の性能に関する評価を適合率 P と再現率 R を組み合わせさせた F 値で評価した。グルーピング構造分析ではグループが所属する階層に関係なく、システム出力と評価用データの両方に同じグループがある場合を適合とした。一方、重要発言選定では、システムの出力と評価用データにおける枝と幹の接点と同じレベルかつ同じ位置にある場合を適合とした。

このような F 値による評価を行うため、対象とする全 10 件の議論データに対して手作業でグルーピング構造分析および重要発言選定を行った評価用データを用意した。評価用データは、議論タイムスパン木をよく理解している本研究報告の筆者を含む 2 名が作成した。セクションに含まれる発言数に対するシステム出力のグルーピング構造獲得分析と重要発言選定の精度をまとめたものを表 3 に示す。総セクション数 169 件のデータに対するグルーピング構造分析の F 値は 0.83、重要発言選定の F 値は 0.65 であった。なお、グルーピング構造分析では、適合率と再現率の値が同様であるため、 F 値の結果のみを記述している。また、発言数における該当するセクション件数の割合は、概ね一律である。

6. 考察

本章では、議論タイムスパン木の表現における正確性について、前章で述べた実験結果から考察する。議論タイムスパン木の自動生成に関する精度は、セクションにおける発言数が増加するに連れて低下するが、大域的なグループ構造の獲得や上位における重要発言の選定は、概ね評価用データと一致していることが分かった。不正解が頻発した代表例として、発言数や発言人物が多い場合が挙げられる。これらの議論の特徴は、前半と後半で議論の中心人物と話題が大きく変化することである。また、セクションに含まれる発言数が奇数である場合、精度が低下する傾向があった。この原因として、質問応答が入れ子構造である話題に対して、木構造を正確に生成できていないことが考えられる。現在のルール実行管理では、各ルールを一律全範囲に適用しているため、各ルールの重み付けを局部的に変更することが今後の課題となる。また、不正解データの追試として、ルールを管理するパラメータの調整による評価用データの再現を試行した。その結果、セクションに含まれる発言数が 4 件以下のものは完全に再現可能であった。

7. おわりに

本稿では、マルチメディアデータを含む会議コンテンツに対する探索的な分析の支援を目的とし、発言間の関係や集合に基づき議論の意図を表現する議論タイムスパン木による構造化手法を提案し、その有効性を検証した。また、議論タイムスパン木の機構を用いて、ユーザに情報探索のための新たな視点を与え、探索的に分析するためのインタラクションデザインの枠組みを構築した。

本システムの機能向上に関する課題としては、正確な議論の木構造を生成する方法、ルールを管理する複数のパラメータから特定の目的に即した要約を実装するための観点の指定法・制御法、議論の木構造を変形するための基本演算子の定義によるインタラクション手法の多様化・高度化、ユーザビリティ向上のため議論の木構造の可視化法などが挙げられる。

参考文献

- [1] Schultz, T., Waibel, A., Bett, M., Metze, F., Pan, Y., Ries, K., Schaaf, T., Soltau, H., Martin, W., Yu, H. and Zechner, K., The ISL Meeting Room System, Proceedings of the Workshop on Hands-Free Speech Communication (2001).
- [2] Rosenfeld, B., Feldman, R., and Aumann, Y., Structural Extraction from Visual Layout of Documents, Proc.11th International Conference on Information and Knowledge Management, pp.203-210 (2002).
- [3] 徳永健伸, 情報検索と言語処理, 東京大学出版会 (1999).
- [4] Lerdahl, F. and Jackendoff, R., A Generative Theory of Tonal Music, The MIT Press (1983).
- [5] 三浦寛也, 長尾確, 平田圭二, 音楽理論 GTTM に基づく議論タイムスパン木の生成方式とその評価, 情報処理学会論文誌, Vol.56 No.3, pp.942-950 (2015).
- [6] 徳永健伸, 言語処理を利用した知的情報アクセス - 検索, 抽出, 要約, 分類, QA -, 日本オペレーションズ・リサーチ学会誌, Vol.52. No.11. pp.713-718 (2007).