

書き起こし議事録を用いた懸案事項共有化のための 談話文脈可視化手法

Visualizing Discourse Context for Sharing Public Concerns Using Meeting Transcripts

白松 俊*¹ 高崎 隼*¹ タチアナ ジドラスコ*¹ 大園 忠親*¹ 新谷 虎松*¹
 Shun SHIRAMATSU Jun TAKASAKI Tatiana ZIDRASCO Tadachika OZONO Toramatsu SHINTANI

奥乃 博*²
 Hiroshi G. OKUNO

*¹名古屋工業大学 大学院工学研究科 Graduate School of Engineering, Nagoya Institute of Technology
 *²京都大学 大学院情報学研究科 Graduate School of Informatics, Kyoto University

Sharing public concerns between stakeholders is important to enhance public involvement. This paper describes a method for visualizing discourse context for sharing public concerns using meeting transcripts. Since a transcript of public meeting tends to be lengthy, visualizing topic transition is required to help users share their concerns. We defined a metric for salience of terms for each sentence to visualize topic transition. Since Web-mediated comments on the transcript tend to be confused and flaming, visualizing rhetorical structure is required to avoid misunderstanding. We proposed a method to suggest appropriate response to the comment based on the rhetorical relations manually specified by users.

1. はじめに

近年、行政への住民参画 (Public Involvement) の社会的な重要性が高まり、情報通信技術を活用した住民参画 (e-Participation) の研究が盛んになりつつある [Macintosh 04, 大塚 07]. 住民参画を促進するためには、住民公聴会などの公的討議を通じ、解決すべき懸案事項を集約し、住民・関係者間で共有化することが重要である。対面の討議には時間制限があり議論を尽くせない場合があるので、そのような場合は Web 等を用いて後から議論をフォローできる仕組みが必要となる。Web 上で意見を集約しようとする試み [Suga 10] も行われているが、デジタルデバイドに配慮した公平な意見集約のためには、Web だけでなく対面の公的討議との併用が不可避であると考えられる。

本研究では、図 1 に示すように、繰り返される対面の会議の間に (1) 議事録閲覧と (2) 議事録へのコメント付与を支援し、懸案事項を共有化するサイクルを考える。(1) の議事録閲覧時の課題は、長い議事録の閲覧労力の低減である。公的討議の議事録は、公平のため発言を取捨選択せず書き起こす場合が多いため、議論の流れの追跡が難しく、懸案事項の理解に多大な労力を要する。(2) のコメント付与時の課題は、コメントを通じた議論の円滑化である。意図の不明確さに起因する「噛み合わない議論」を防止することで、懸案事項の共有化を促進することができる。

これら 2 つの課題を解決するため、(1) 話題遷移と (2) 修辞構造という 2 つの観点から談話文脈を可視化する。議事録中の話題遷移は、センタリング理論 (centering theory) [Poesio 04] と整合する「注目されている談話実体ほど後続発話で参照され易い」という仮定に基づいて可視化する。コメント間の修辞構造については、修辞構造理論 (rhetorical structure theory) を拡張した修辞関係タグを定義し、ユーザが指定したタグを用いて可視化する。これらの可視化により、懸案事項共有化を促

進するシステムの開発を目指す。

2. 話題遷移の可視化

長い議事録の「流れ」を把握する労力を軽減するため、われわれは図 2 のように文ごとに变化する語の「目立ち具合」を定量化し、議事録中の話題遷移を可視化する。図 2 の横軸が表す量を、語の顕現性 (salience) と呼ぶ。われわれは過去の研究 [Shiramatsu 08a, Shiramatsu 08b] で「注目されている単語ほど、次の発話で参照される確率が高い」という仮定に基づき、顕現性を以下の参照確率 (reference probability) で定義した。

$$\begin{aligned} & \text{(発話 } U_i \text{ 時点における単語 } w \text{ の顕現性)} \\ & = p(\exists w' \xrightarrow{\text{coref}} w \text{ in } U_{i+1} | \text{pre}(U_i)) \\ & = p(w | \text{pre}(U_i)) \end{aligned}$$

ただし $\text{pre}(U_i)$ は発話 U_i までの先行文脈 $[U_1, \dots, U_i]$ を表し、 $\exists w' \xrightarrow{\text{coref}} w \text{ in } U_{i+1}$ は単語 w の指示対象を参照する単語が後

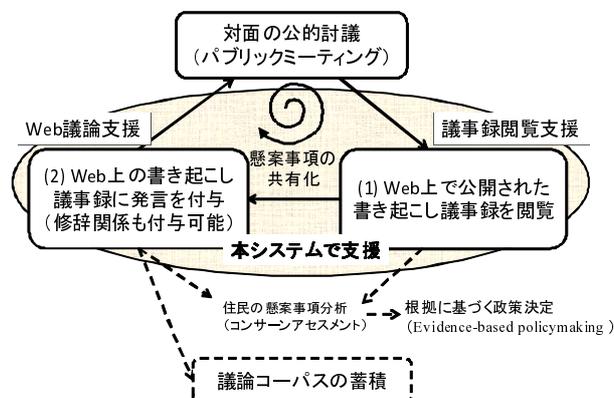


図 1: 懸案事項共有化サイクル

連絡先: 白松 俊, 名古屋工業大学 大学院工学研究科 情報工学専攻, 〒466-8555 名古屋市昭和区御器所町, E-mail: siramatu@nitech.ac.jp

続発話 U_{i+1} 中に存在することを表す $p(w|\text{pre}(U_i))$ は参照確率の簡易表記である。

この参照確率による顕現性定義は、センタリング理論 (注意中心の連続性と代名詞化についての理論) [Grosz 95] との整合性が実験により確認されている [Shiramatsu 08a]。また、以下の素性を用い、コーパスから統計的に計算可能である [Shiramatsu 08b]。

- 新近性効果: $w' \xrightarrow{\text{coref}} w$ から次発話 U_{i+1} への発話距離に応じた窓関数
- 出現頻度: $\text{pre}(U_i)$ における $w' \xrightarrow{\text{coref}} w$ の頻度
- 文法役割: $w' \xrightarrow{\text{coref}} w$ が係る助詞等の機能語
- 品詞: $w' \xrightarrow{\text{coref}} w$ の品詞

文法役割と品詞は、図 3 に示すように単語の各出現箇所ごとに抽出される。新近性効果は、1 発話毎にシフトする窓関数により表され、窓関数で重み付けされた各素性値の総和をとることで出現頻度が反映される。窓関数は、コーパス中の各単語 w が後続文脈で参照される確率から決定する。図 4, 5 は、淀川水系流域委員会の議事録をコーパスとして用いた場合の新近性効果を表している。

語 w と発話位置 U_i の組 $\langle w, U_i \rangle$ に対する素性値の重み付き和を $\text{feat}_f(\langle w, U_i \rangle)$ とすると、参照確率は以下のロジスティック回帰式で計算される。回帰重み b_k は、コーパス中のサンプル $\langle w, U_i \rangle$ を用いて求める。

$$p(w|\text{pre}(U_i)) = \frac{1}{1 + \exp(- (b_0 + \sum_{k=1}^n b_k \cdot \text{feat}_k(\langle w, U_i \rangle)))}$$

さらに、各単語の参照確率から成るベクトルを PLSA (Probabilistic Latent Semantic Analysis) を用いて次元圧縮すると、図 6 に示すような話題遷移の可視化が可能になる。このよう

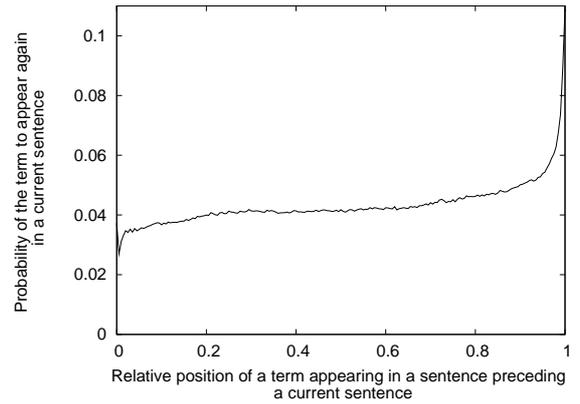


図 4: 議事録中の系列位置曲線

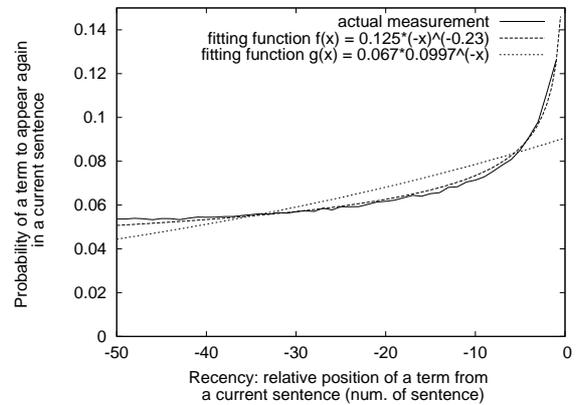


図 5: 議事録中の新近性効果

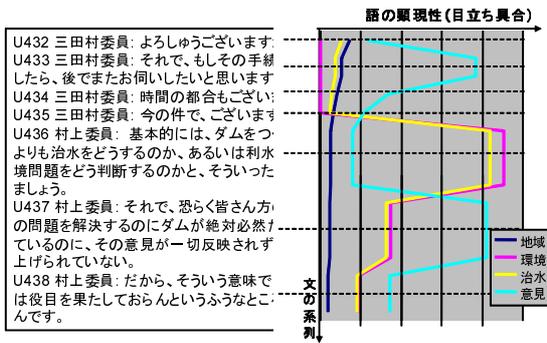


図 2: 文ごとに動的に変化する語の顕現性 (salience)

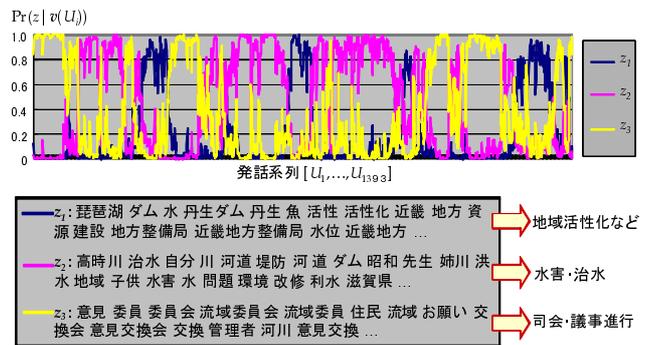


図 6: PLSA の潜在トピックに基づく SaliencyGraph

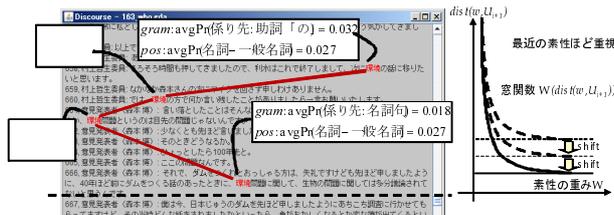


図 3: 参照確率算出のための素性抽出とシフトする窓関数

な潜在トピックの遷移図は、単語毎の遷移図よりも全体の傾向を表していると予想され、懸案事項の理解に有効であると考えられる。

この定式化により、図 7 のような議事録閲覧支援システム SaliencyGraph を試作した。所望の議題語の顕現性が高い位置にスライダーを移動すれば、連動して所望の議論が表示される。これにより、議題遷移の把握と所望の議論の発見が容易になり、懸案事項の共有化に繋がると期待される。

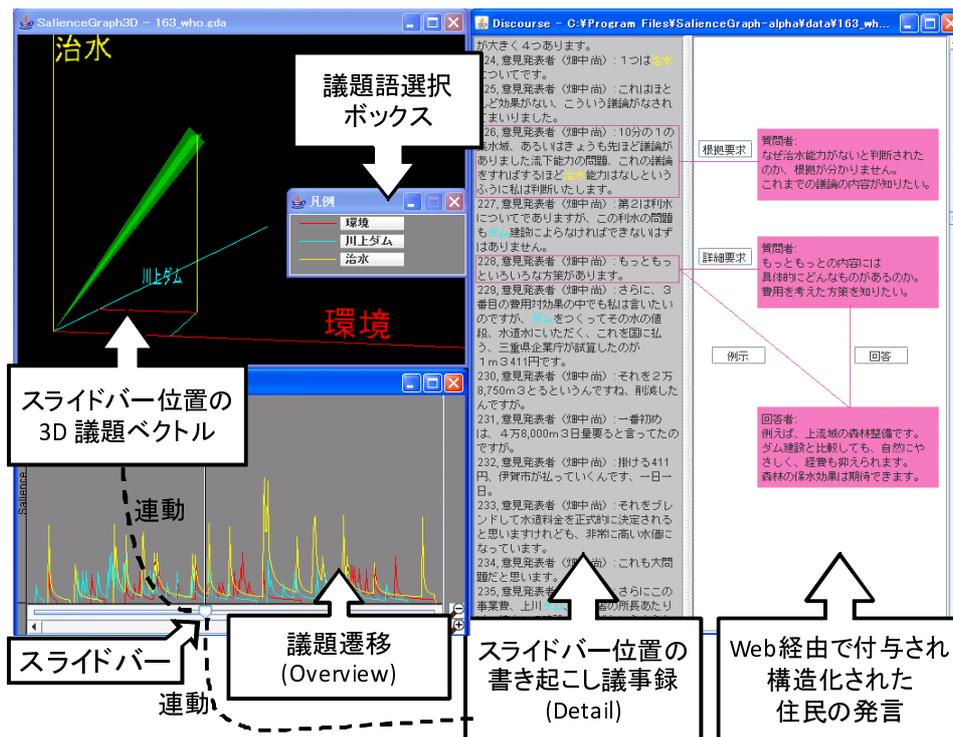


図 7: SaliencyGraph: 話題遷移の可視化による議事録閲覧インターフェース

SaliencyGraph で閲覧する議事録に対しては, CaboCha [Kudo 02] による係り受け解析と, 先行文脈で出現した全名詞句 w の顕現性 $p(w|pre(U_i))$ の計算を, あらかじめ前処理として行う. 言語構造を付与するための XML ボキャブラリである GDA (Global Document Annotation) [Hasida 04] によりこれらの情報を議事録に自動タグ付けしておけば, SaliencyGraph で話題遷移図を可視化することができる. SaliencyGraph は, Shneiderman が提唱する *Visual Information-Seeking Mantra*, すなわち “Overview first, zoom and filter, then details on demand” [Shneiderman 98] に基づいて設計されている.

3. 修辞構造の可視化

対面の会議の時間的制約や, 出席できない関係者の存在により, 会議終了後から次の会議の間までに書き起こし議事録を用いて懸案事項の共有化を促進する必要がある. そのためには閲覧労力の低減のみではなく, 議事録中の発言に対する質問・回答や補足発言の支援が必要であるので, 図 7 の右端に示す発言付与インターフェースを試作した. これは, 議事録閲覧支援システム SaliencyGraph とセマンティックオーサリング [Kamimaeda 07] を統合した議論支援システムのプロトタイプである. ユーザは Web 経由で発言ノードを付与し, 議事録中の発言あるいは Web 経由で付与された発言ノードにリンクを張る. リンクは, どの発言に関する発言であるかを表す. さらに, リンクに修辞関係レベルを指定できる (必須ではない).

本研究では, 付与された修辞構造を用い, 懸案事項共有化や合意形成を阻害する「噛み合わない議論構造」の検出・抑止手法を検討中である. そのために, 実験的なルール設計 [高崎 10] と定量的コーパス分析 [Zidrasco 10] という 2 つのアプローチを併用している. 噛み合わない議論検出のためのルールとし

て, 主張に根拠が無い, 訊かれたことに答えていない, 等の状態を修辞関係レベルを用いて定式化した [高崎 10]. これに基づいて図 8 のように適切な応答を示唆する機能を開発することで, 懸案事項の共有化促進に繋がると考える. また小規模な議論コーパスを作成し, 合意を導く確率の高い修辞構造パターンを調査中である [Zidrasco 10]. 今後は, 試作したシステムを運用して議論コーパスを蓄積し, これら 2 つのアプローチを同時並行的に推進する.

4. 関連研究

議論を構造化, 可視化することで議論分析を支援する研究は, 数多く存在する. ArguMed [Verheij 03], Araucaria [Reed 04] 等の議論可視化研究では, 発話同士をリンクで結び, 局所的な議論のグラフ構造を可視化していた [Kirschner 03, Braak 06]. 上前田ら [Kamimaeda 07] は, 議事録の修辞構造から議論参加者の貢献度を計算する手法を提案した. 土田ら [Tsuchida 09] は, 対面の会議における議論を構造化し, 可視化するシステムを開発した. 本研究における修辞関係の可視化もこれらの流れを汲む手法であるが, 図 8 に示すような議論の円滑化機能によって懸案事項の共有化促進を目指している点がこれらの研究との違いである. さらに, 長い書き起こし議事録の閲覧支援のためには局所的なグラフ構造だけでなく議事録の全体像を可視化する必要があるため, 文毎に遷移する話題の「流れ」を扱った.

5. おわりに

本稿では, 対面の公的討議と Web 上のインタラクションを組み合わせたサイクルにより関係者間の懸案事項共有化を促進するため, 議論の談話文脈を可視化する手法を提案した. 具

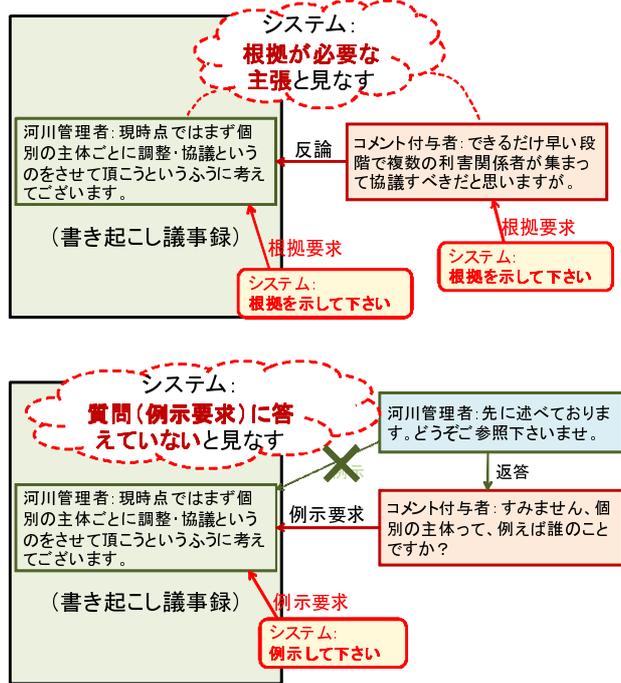


図 8: システムによる適切な修辞関係のサジェスチョン例

体的には、長い議事録における議題の「流れ」を分かり易くするための話題遷移の可視化手法と、Web 上のコメントによる議論の混乱を防止するための修辞構造の可視化手法について述べ、それに基づく議事録閲覧インタフェース SaliencyGraph を開発した。話題遷移については、参照確率と PLSA に基づく可視化手法を提案し、修辞構造については、ユーザが指定した修辞関係を用いて適切な応答を示唆する機能を提案した。SaliencyGraph は、Visual Information-Seeking Mantra に基づいて議事録閲覧を支援できるように設計した。今後は、試作したシステムを運用して議論コーパスを蓄積し、これらの可視化手法による懸案事項共有化効果の評価手法を検討する予定である。

謝辞 本研究は科研費補助金 (No. 09152946, No. 19100003) の助成を受けたものです。

参考文献

[Braak 06] Braak, van den S. W., Oostendorp, van H., Prakken, H., and Vreeswijk, G. A. W.: A critical review of argument visualization tools: Do users become better reasoners?, in *Workshop Notes of the ECAI-2006 Workshop on CMNA*, pp. 67–75 (2006)

[Grosz 95] Grosz, B., Joshi, A., and Weinstein, S.: Centering: A Framework for Modeling the Local Coherence of Discourse, *Computational Linguistics*, Vol. 21, No. 2, pp. 203–226 (1995)

[Hasida 04] Hasida, K.: Global Document Annotation (GDA), <http://i-content.org/GDA/> (2004)

[Kamimaeda 07] Kamimaeda, N., Izumi, N., and Hasida, K.: Evaluation of Participants' Contributions in Knowledge Creation Based on Semantic Authoring, *The Learning Organization*, Vol. 14, No. 3, pp. 263–280 (2007)

[Kirschner 03] Kirschner, P., Shum, S., and Carr, C.: *Visualizing Argumentation: Software Tools for Collaborative and Educational Sense-Making*, Springer (2003)

[Kudo 02] Kudo, T. and Matsumoto, Y.: Japanese dependency analysis using cascaded chunking, in *Proc. of CoNLL-2002, COLING 2002 Post-Conference Workshops*, pp. 1–7 (2002)

[Macintosh 04] Macintosh, A.: Characterizing E-Participation in Policy-Making, in *Proceedings of the 37th Annual Hawaii International Conference on System Sciences*, pp. 117–126 (2004)

[Poesio 04] Poesio, M., Stevenson, R., Eugenio, B. D., and Vieira, R.: Centering: A parametric theory and its instantiations, *Computational Linguistics*, Vol. 30, No. 3, pp. 309–363 (2004)

[Reed 04] Reed, C. and Rowe, G.: Araucaria: Software for Argument Analysis, Diagramming and Representation, *International Journal of AI Tools*, Vol. 13, No. 4, pp. 961–980 (2004)

[Shiramatsu 08a] Shiramatsu, S., Komatani, K., Hasida, K., Ogata, T., and Okuno, H. G.: A Game-Theoretic Model of Referential Coherence and Its Empirical Verification Using Large Japanese and English Corpora, *ACM Transactions on Speech and Language Processing (ACM-TSLP)*, Vol. 5, No. 3, p. Article 6 (2008)

[Shiramatsu 08b] Shiramatsu, S., Komatani, K., Ogata, T., and Okuno, H. G.: SaliencyGraph: Visualizing Saliency Dynamics of Written Discourse by Using Reference Probability and PLSA, in *Proceedings of the Tenth Pacific Rim International Conference on Artificial Intelligence (PRICAI-08)*, pp. 890–902, Springer (2008)

[Shneiderman 98] Shneiderman, B.: *Designing the User Interface: Strategies for Effective Human-Computer Interaction*, Pearson Addison Wesley (1998)

[Suga 10] SugarForum.jp, 経済産業省: アイデアボックス「その後」, <http://after-ideabox.net/> (2010)

[Tsuchida 09] Tsuchida, T., Kikuchi, K., Ohira, S., and Nagao, K.: Visualization of Discussions in Face-to-Face Meetings, in *Proceedings of the 5th International Conference on Collaboration Technologies* (2009)

[Verheij 03] Verheij, B.: Artificial argument assistants for defeasible argumentation, *Artificial Intelligence*, Vol. 150, No. 1–2, pp. 291–324 (2003)

[Zidrasco 10] Zidrasco, T., Takasaki, J., Shiramatsu, S., Ozono, T., and Shintani, T.: Analyzing an Argumentative Discourse Structure for Supporting Argumentation, *情報処理学会創立 50 周年記念 (第 72 回) 全国大会論文集*, 5D-3 (2010)

[高崎 10] 高崎, Zidrasco, 白松, 大園, 新谷: 議論支援システムのための対立と質問に着目した適切性ルール設計, *情報処理学会創立 50 周年記念 (第 72 回) 全国大会論文集*, 4V-4 (2010)

[大塚 07] 大塚, 乾, 奥村: 意見分析エンジン-計算言語学と社会学の接点-, コロナ社 (2007)