

対話状況の可視化のための ヒューマン・コンピュータ協調モデル

高 悠史^{1,a)} 吉本 廣雅² 近藤 一晃² 中村 裕一²

概要：

本研究では、会議中の対話状況を可視化するためのユーザと計算機の協調モデルの設計を行う。その基本的な考え方は、人間が理解しやすく、操作しやすい表現を媒介として、ユーザの入力、計算機の自動処理、ユーザの修正を簡単に繰り返すことのできる計算・編集モデルとすることである。本稿ではその設計とプロトタイプシステムの試作、これを被験者に使ってもらって行った予備実験について報告する。

キーワード：会議支援、対話状況の可視化、ヒューマン・コンピュータ協調

Human-Computer Collaboration Model for Visualization of Conversation Situations

KO YUSA^{1,a)} YOSHIMOTO HIROMASA² KONDO KAZUAKI² NAKAMURA YUICHI²

Abstract:

In this paper, we discuss a new human-computer collaboration model for visualizing conversation situations in meetings. The key ideas are the representation of conversation situations which is easy for humans to understand and handle, and the model for calculation and editing which enables to repeat input by an user, automatic processing by computer, and modification by the user again. We first designed our model, prepared a prototype system, and investigated the performance of our model.

Keywords: meeting support, visualization of conversation situations, human-computer collaboration

1. はじめに

我々は、多様な状況下でも議論に参加しやすくするために、会議中の対話状況を逐一可視化し、議論の振り返りや点検の助けにするシステムの実現を目指している。

従来から、会議中に議論を記録したり、また、それを会議中に共有するための様々な工夫が行われている。

参加者が議論を逐一メモしたり、ホワイトボードや付箋でまとめたりするような、自然な会議を手作業で記録する

ものだけでなく、「議事録ドリブン」[1]のようにあらかじめ決まったフォーマットに従って議事録を作成しながら議論を進めていく提案もある。しかし、このような手作業はコストが大きく、議論に積極的に参加しない書記係を置く必要があったり、そうでない場合には、書いている間議論を中断したり、書く内容を大幅に取捨選択したりする必要があったりする。

このような背景から、音声認識や言語処理によって自動で議事録を生成するアプローチ [2] や、非言語的な特徴を使ってコミュニケーションの様子を可視化するアプローチ [3]、議事録の構造化を行いやすくするために補助的な合図や機器を用いる [4], [5] 等、様々な試みが行われてきた。これらの多くは議事録を残したり閲覧することを目的としたものであり、必ずしも会議中の支援を目指したものでは

¹ 京都大学工学研究科
Graduate School of Engineering, Kyoto University

² 京都大学学術情報メディアセンター
Academic Center for Computing and Media Studies, Kyoto University

^{a)} ko@ccm.media.kyoto-u.ac.jp

ないが、その考え方には本研究と共通するものが多い。

このような議事録作成の自動化には多くの課題がある。特に、音声認識や発言の意図の認識など、精度の面で困難な問題を抱えており、そのまま議事録として用いることのできるものとはなっていない。そのため、会議後の議事録作成であれば、自動的な記録の後に人手で編集操作を加えることが必要とされる。

本研究が目指しているような議論の可視化では、この問題がより深刻なものとなる。議論中に可視化された内容を利用するためには、分かりやすい内容が大きな手間をかけずに得られることが必要となるためである。これは難しい要求であるが、本研究では、そのためのアプローチの一つとして、ユーザと計算機の協調モデルの設計を行った。

その基本的な考え方は、人間が理解しやすく、操作しやすい表現を媒介として、ユーザの入力、計算機の自動処理、ユーザの修正を簡単に繰り返すことのできる計算・編集モデルとすることである。つまり、ユーザにとって認知と操作の両方の点で負担の少ない設計とし、計算機側も与えられるユーザの操作を柔軟に利用する計算モデルを用いることとした。

以上のような考えに基づき、本研究では協調モデルに基づいたプロトタイプシステムを試作した。これを被験者に使ってもらって予備実験を行い、本研究の有効性の確認と問題点の洗い出しなどを試みた。これらにより、協調が効果的に機能している事例や、設計の不備が現れている点等、多くの知見や今後の改良の指針が得られた。

以下，2章では対話状況を可視化する協調モデルの設計の概略を述べ，3，4章では今回試作したプロトタイプシステムの実装について述べる．5章では試作システムを用いた被験者実験をもとに，提案する協調モデルの現状評価と今後の改良指針の検討を行い，6章で結論を述べる．

2. 対話状況を可視化する協調モデル

2.1 会議の進行状況を把握するために必要な要素

会議の進行状況を可視化するうえで重要なことは、短時間で状況を把握できることである。

そのための一つの観点として、我々は対話の結束性に着目している [6] . 結束性とは、談話・対話研究の領域で知られている、対話のもつ一般的構造の一つである。対話は、会議参加者の発する一つ一つの発話の列によって構成されるが、それぞれの発話は互いに無関係ではなく、関連性に基づいて小さなまとまりを形成する。さらに、そのようなまとまり同士も互いに関連性をもち、階層的により大きなまとまりを形成する [7] .

この結束性を踏まえ、会議の進行状況を可視化するうえで重要な要素を挙げると以下になる。

● 発話内容

発話の書き起こしや要約、議論の意味的な内容をを知

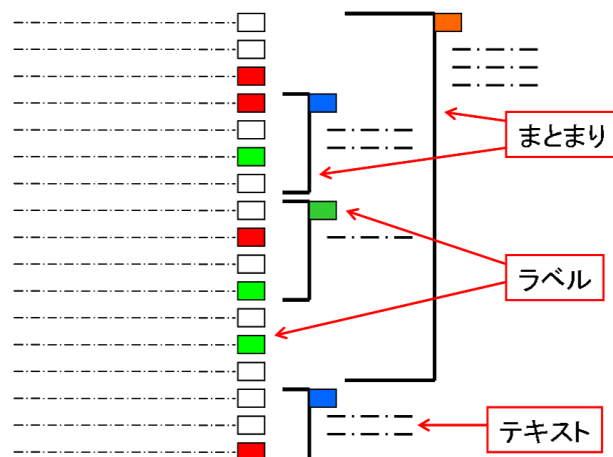


図 1 対話状況を把握するために必要な三つの要素

るために必須である。

- 発話の役割

テーマ、決定事項、補足的情報など、対話の文脈の中で個々の発話が果たしている役割、会議中に発話内容を全て読むような時間的な余裕は期待できない、そのため、重要な発話に効率良くアクセスするための補助として必要である。

- 発話のまとめり

話題を形成する発話のまとまりや、それら同士の包含関係、発話の役割とも重複するが、発話の流れや関係性を効率的に把握するのに必要となる。まとまりの情報がなくただ情報が羅列されているだけでは、どのような話題が話されたか分かりにくいからである。

これら以外にも、例えば話者の顔画像など、対話状況の把握に役立つ要素はいくつも考えられるが、本研究では、上記三つを短時間で議事を把握するための必要最小限の組み合わせとして扱う。

議論しながら参照するという、短時間での理解が求められる形態では、これらの三要素がそれぞれ適切に配置されていることが必須である。この三つを用いて対話状況を表現できるようにすることが本研究の目的である。例えば図1のような表現であれば、どのような話題についてどのくらい議論がなされたか、また、どの発話が重要であるかを簡単に把握することができる*1。

2.2 協調モデルの設計方針

前節で挙げた「発話内容」、「発話の役割」、「発話のまとまり」の三要素を発話から求めるための、ユーザと計算機の協調モデルを考える。

基本的考え方としては、ユーザと計算機は図 2 のような協調関係をもつ。計算機は三要素の自動推定を行う。ユー

*1 書記係がとる議事録などでは、発話内容を記録するだけで精一杯になりやすく、発話の役割やまとまりといった構造指標が不十分になりがちである。そのため、瞬時にできるものとはなりにくい。

ずは GUI を通じて三要素を直接入力したり、中間結果を修正する。ただし、三要素は独立でないため、各々の結果が互いに波及する。

次に、このような協調を可能にするための、ユーザ側、計算機側からの要件を検討する。
ユーザ側からの要件

上記三要素は、場面によっても、またユーザの好みなどによっても、考えやすさ、操作の難しさが異なる。例えば、議論の文脈上重要な役割を果たす発言が出た場面なら、当該発言にその果たした役割をラベリングする操作が重要かつ簡単であろう。それに対して、ある話題が一段落したときなどは、一まとまりの終りであることを記録するのが自然である。他にも、大局的な話題からトップダウンにまとめる方が作業しやすいと感じるユーザ、ボトムアップにまとめることを好むユーザ、発言の役割のラベリングを優先しようとするユーザなど、どの発言のどの要素に対して操作するかについては様々な要求があり得る。そのため、ユーザ側からの要件としては、三要素それぞれについて直観的に操作できる機能を用意し、どこからでも、どの操作からでも作業できることがあげられる。

計算機側の要件

まず、計算機側は発言の種々の特徴量を基に三要素の初期データを生成しなければならない。この段階で完全な精度は要求されないが、ユーザの入力や修正を妨げない程度の質は必要である。次に、ユーザからの入力・修正を受けて、三要素の更新ができなければならない。三要素は本来それぞれ独立に決まる性質のものではないため、どれか一つが決定すると他の二つも制約される。そのため、矛盾が起らないように再計算を行う必要があり、この再計算はユーザがどこから、どの操作を行ったとしても可能でなければならない。

その他の要件

計算機が自動推定を行うときは、ユーザからの入力・修正が自動推定のための事前知識として利用でき、推定精度が向上する。この入力・修正の個所によって向上の程度が異なることから、入力・修正が望まれる個所を計算機側からユーザへ知らせることができれば効率的である。ただし、必ずしもユーザはそれに従う必要はない。

以上の要件を基に、本研究における協調モデルを以下のように設定する。

- (1) 計算機は会議を観測し、三要素の自動推定を行い、それを提示する。
 - (2) 計算機は、可能な限り、ユーザからの入力・修正を促すヒントも提示する。
 - (3) ユーザは自分の取り扱いやすい箇所から、操作しやすい要素を選んで入力・修正する。
 - (4) 計算機はユーザの入力・修正に応じて再計算を行う。
- このような段階を繰り返していくことで、質の高い可視

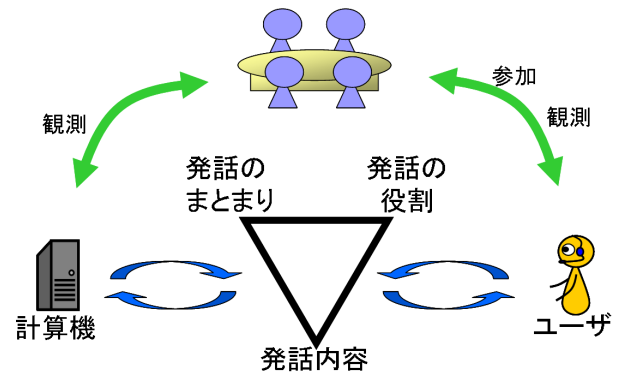


図2 ユーザと計算機の協調関係

化データが得られると期待される。

ただし次節で述べるように、今回試作したプロトタイプシステムでは三要素のうち、発言内容をユーザが操作する手段が未整備であり、発言の役割とまとめのみ操作可能となっている。

2.3 対話状況の表現

前節で議論した要件に基づき、対話状況の可視化インタフェースを次のように設計した。概観を図3に示す。

発言内容

発言の書き起こし、またはその要約を表示する。実際の会議に用いる際にはリアルタイム音声認識および自動要約の援用を想定している*2。本稿の実験では人手による書き起こしを用いた。

発言の役割

今回は「発議」と「それ以外」の二者択一とする。発議とは、ある話題のテーマを提示して意見交換の口火を切るような役割の発言を意味する。発議に相当する発言としては、疑問文が多く選ばれることを想定している。発議をハイライトすることで表現し、一連の議論のテーマにあたる発言が強調表示されたような可視化結果が得られる。他にも「決定事項」「補足事項」などの役割も有用だと考えられるが、これは今後の課題としたい。

発言のまとめ

本インタフェースでは、対応する発言列を枠で囲うことで表現する。入れ子構造をもつことも可能である。枠による表現は、話題の境界を決定することに基づいている。これに対して、坊農らによって境界よりもむしろ話題の核を重視するという考え方も提案されている[8]。今後の検討課題としたい。

以上のインタフェースに対し、3章で述べる計算機側の自動推定結果が入力され、得られた結果に対して、4章で述べるユーザの編集が行われることとする。

*2 ある程度の認識誤差を想定し、深い意味処理などは想定しない。

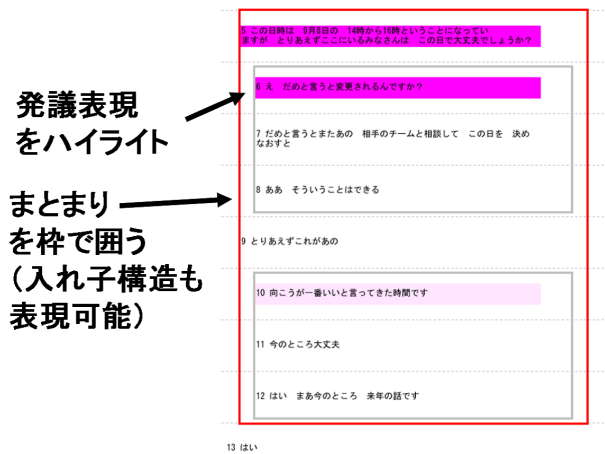


図 3 可視化インターフェース

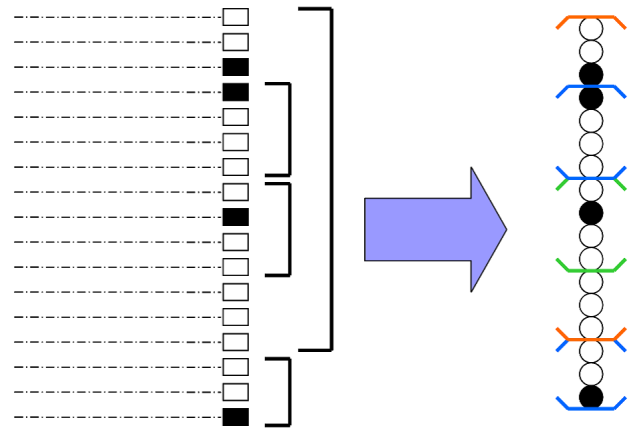


図 4 データ構造

3. 対話状況の計算モデル

3.1 発話系列とデータ構造

前章 2.1 節で議論した可視化のための三要素を推定するためのデータ構造を設定する。

まず、本モデルでは会議中の対話を一つ一つの発話の系列として扱う「発話内容」と「発話の役割」はこの一つ一つの発話に対応する属性として対応付ける「発話のまとめり」は、話題の始点と終点にあたる発話に「話題の始点」「話題の終点」という属性を対応付けることで表現する。例えば、図 4 では、括弧が話題の始点と終点を表し、黒丸が「発議」、白丸が「それ以外」を表す。また、話題の構造をスタックと考えることにより、このような表現だけで、話題の入れ子構造も表現できる

このようなデータ構造に曖昧性はなく、可視化表現がただ一つ対応するこれにより、対話状況を可視化する処理は、発話系列の各発話に対して上記の属性を付与する問題に帰着される

また、一つ一つの発話について会議を観測することによって得られた特徴量も属性として与えるが、それについては次節で述べる。

3.2 発話の特徴

各々の発話に対して、以下のような特徴を観測し、属性として与える。これらは音声や画像等のセンシングによって自動的に得られることを想定しているが、今回は全て人手で特徴量を抽出した。

- 発話文 (書き起こし)
- 発話の長さ (秒)
- 前後の発話との間や重なりの特徴 (秒)
 - (後の発話の開始時刻) - (前の発話の終了時刻)
 - (後の発話の終了時刻) - (前の発話の終了時刻)
- 話題転換に関わるキーワードの有無 (有: 1, 無: 0)

今回用いたキーワードは「で、では、じゃ、じゃあ、まず、さて、とりあえず、ところで、それと」

- あいづちとフィラーに関わるキーワードの有無 (有: 1, 無: 0)

今回用いたキーワードは「はい、あ、あの、ああ、え、ええ、ええと、えーと、うん、うーん、んー、そう、ふーん、まあ、いや」

- 疑問文かどうか (該当: 1, 非該当: 0)
- 発話が「～か」で終わっているかどうか (該当: 1, 非該当: 0)
- 話者交代が起きているかどうか (該当: 1, 非該当: 0)

これらは発話の役割推定に関する従来の研究 [9] で用いられている特徴であり、その有効性がある程度検証されているものである。

3.3 状態遷移モデルと推定

2.3 節で述べたように、今回は「発話内容」については書き起こしを用いるのみとし、操作を行わないので、「発話の役割」(「発議」か「それ以外」か)と「発話のまとめり」のみを推定する問題を扱う。ただし、この二つの要素は独立ではいため、同時に推定するために、発話の特徴に対応付けられた入力、「発議」、「それ以外」という属性に対応付けられた出力をもち、「話題の始点」、「話題の終点」という属性を内部状態にもつ状態遷移モデルを考え、各発話にこれらの属性 (以下、「発話役割」と呼ぶ) を割り当てる計算を行う。

本研究では、状態遷移モデルとして確率的プッシュダウン・オートマトンを用いた (図 5) このオートマトンは、「司会者等による議題の導入」→「論点の提起」→「意見交換」という流れを基本とし、その途中で参加者が発する質問や意見をきっかけにサブトピックが挟まれるようなパターンを表現している「発議」は「!」、「それ以外」は「-」という出力シンボルに対応し、話題の始点と終点はそれぞれスタックへの push と pop に対応する

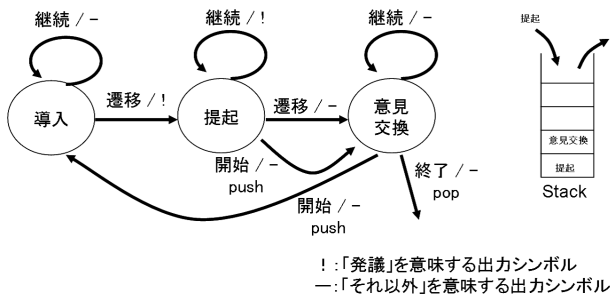


図 5 今回用いたプッシュダウン・オートマトン

入力シンボルは C (継続), T (遷移), S (開始), E (終了) などである。これらのシンボルは各発話の特微量を用いて、サポートベクタマシン (SVM) によって選択される。この入力に従って遷移先が決定する。なお、本稿の実験では、事前に収録して特微量を手で付けた 293 発話のデータを用いて SVM のパラメータを学習した。

このオートマトンを用いて、確率的に幅優先探索し、確率が閾値以上の系列 (発話役割の系列) を求める。各遷移先への遷移の尤度は SVM による入力シンボル識別の尤度 (0 ~ 1) を用い、これを一回の遷移のスコアとする。一連の状態遷移のスコアをかけた累積値を、対応する発話役割系列のスコアとする探索中、スコアを累積した値が閾値を下回った系列は棄却する。候補が複数ある場合には、スコアの低い系列から順に優先する。

なお今回は予め構造を想定したオートマトンを利用したが、将来的には、角 [10] が提案する N-gram モデルなどの援用が期待できる。

4. 可視化データの編集モデル

4.1 ユーザによる編集操作

「発話の役割」と「発話のまとまり」の二つの要素に対して、ユーザが直観的に操作を行えるよう、以下の機能を用意した。

- 発議のハイライト：該当発話をクリックする
- 発話のまとまりを枠で囲う：該当する発話列をドラッグする
- 誤ったハイライトの取り消し：該当発話を再度クリックする
- 誤った枠の取り消し：該当する枠をダブルクリックする

4.2 ユーザの入力・修正に基づく再計算

ユーザが行った入力・修正に応じて以下のような再計算を行う。

- (1) ユーザが実行したハイライト操作や枠で囲う操作は、該当する発話に発議や話題の始点・終点の属性を確定

表 1 実験で使った模擬会議映像

	会議のテーマ	発話数	参加者数
被験者 A	サッカーの試合について	200	4 人
被験者 B	PC の購入について	233	4 人

的に書き込む操作とみなす。

- (2) この操作により、オートマトンの現状態から次状態への状態遷移先が指定されたことになる。

- (3) そこで、指定された状態遷移には最高のスコア 1 を、それ以外の遷移にはスコア 0 を与えるものとする。

ユーザが入力・修正操作を行うと、それを受けて計算機が発話役割の再計算を行うこれにより、スコア 0 が与えられた候補は棄却され、ユーザの指定する条件に整合する候補だけが残る。こうして残った候補系列から、改めて累積スコアの低い順に推定結果が構成される。

このような入力がどの時点でどの発話に対して与えられても、発話役割の推定処理を行うことによって再計算が可能であり、このような編集処理をユーザが満足するまで繰り返すことにより、質の良い結果を得ることができる。

4.3 ヒントを利用した編集

発話役割の自動推定結果を編集処理のヒントとして利用するため、以下の機能を用意した。図 3 がその可視化例である。

まず、ユーザが編集した結果は発議は濃いハイライト発話のまとまりは濃い枠と表現するこれに併せて、計算機が推定した結果を発議は薄いハイライト、発話のまとまりは薄い枠で囲うことで表現するこれにより、ユーザは自分が入力した情報の他に、計算機が推定した結果をヒントとして利用することが可能になる。

5. 被験者実験

5.1 概要

試作したプロトタイプシステムを用いた予備実験により、会議の対話状況がどの程度可視化できるのか、またその際にユーザと計算機の間でどのような協調が生じるか、分析を行った。

5.2 方法

実験は 2 人の被験者に対して行った。各被験者はプロトタイプシステムの機能・操作方法について十分な説明を受けており、対話状況の可視化作業の予行練習をすませた上で、以下の手順で実験を行った。

まず被験者は模擬的な会議対話の映像を閲覧し、閲覧後にプロトタイプシステムを用いて会議の対話状況を可視化作業を行う。その際、同時進行で実験者は被験者へアンケートを行い被験者はその質問へ答えながら各種操作を行う。各実験で使った模擬会議映像を表 1 に、アンケート

・発議表現をハイライトしてください。 ・トピックのまとまりを持ってください。 ・正確さにこだわりすぎず、速さ重視で作業してください。 ・作業次の3つの段階に区切り、そのつど手を止めてください。 ①次を行う操作を決定する。 ②操作を実行する。 ③表示の変化を読み取る。 自由に行っていたり構わない操作は、マウスのホイールや画面のスクロールによってスクロールすること。 それ以外(クリック、ドラッグなど)は実行前に必ず手を止めてください。 各段階で以下の質問に答えていただきます。あまり深く考えず、ハッと答えてください。 ①次に何を行う操作を決定した。手を止めて以下の質問に答えてください。 (a)今から何を行う操作を決定したか。 ア. ○をハイライトする イ. ○を△を括弧する ウ. 誤りを取り消す (b)すぐ決められましたか? 1. すぐ決められた 2. 少し時間がかかった 3. すぐ時間がかかった (c)決定に迷いはありませんか? 1. 全くない 2. 少し迷いがある 3. すぐ迷ったが、とりあえず →どう迷いましたか? (複数回答可) ア. 他に迷うような箇所がある イ. まとまりの位置に自信がない ウ. ハイライトの位置に自信がない エ. その他(自由回答) (d)自分が既に書き込んだ情報は役に立ちましたか? 1. とても役に立った 2. 少し役に立った 3. 役に立たなかった 4. 無視した →どう役に立ちましたか? (複数回答可) ア. どこに注目するかの目安になった イ. 既に書き込んだ枠やハイライトが、次へすべき操作の根拠になった ウ. その他(自由回答)	(e)(計算機による自動推定がある場合のみ)計算機が出てきたヒントは参考になりましたか? 1. とても参考になった 2. 少し参考になった 3. 参考にならなかった 4. 無視した →どう参考になりましたか? (複数回答可) ア. どこに注目するかの目安になった イ. ヒントをそのまま採用できる ウ. まとまりの位置を修正すれば済む エ. ハイライトの位置を修正すれば済む オ. その他(自由回答) (f)それでは操作を実行してください。操作が済んだら、手を止めて以下の質問に答えてください。 (f)操作は簡単でしたか? それとも面倒くさかったですか? 1. 簡単だった 2. 少し面倒だった 3. とても面倒だった →どう面倒くさかったですか? (複数回答可) ア. まとまりが大きすぎて面倒に イ. 操作したい位置を見失いやすい ウ. その他(自由回答) (g)表示の変化について、思うことがあれば教えてください。 (g)操作は成功しましたか? 1. 成功した 2. 失敗した 3. どちらともいえない (h)(計算機による自動推定がある場合のみ)計算機が出てきたヒントは変化しましたか? 1. 望ましい変化をした箇所がある 2. 望ましくない変化をした箇所がある 3. 意味のある変化が起きたと思わない →どう変化しましたか? (複数回答可) ア. 枠が 適切になった / 不適切になった イ. ハイライトが 適切になった / 不適切になった ウ. その他(自由回答)
--	--

図 6 アンケートの質問項目

の質問項目を図 6 に示す。

作業を始めるにあたって被験者には操作方法の説明とともに、以下の内容を指示した。

- 議論の口火を切るような発議表現をハイライトすること。発議表現を探すには、疑問文を目安にすることを勧めること。
- 話題がひとまとまりになっている箇所を探して枠で囲うこと。
- 正確さにこだわりすぎず、速さを重視して作業すること。

なお 3 点目の指示は、提案システムが会議進行中での作業を想定していることからである。

作業中、被験者は 1 回操作を行う毎に手を止め、実験者からの質問に答える。この作業を、被験者が納得できる可視化結果が得られるまで繰り返す行い、その過程でのユーザの操作、可視化の進行状況を逐次記録した。

5.3 実験結果

アンケートに対する 2 人の被験者の回答をそれぞれ表 2、表 3 に示す。被験者の操作回数は、2 人とも計 23 回であった。

また「発議表現をハイライトする」(以下、ハイライト操作と呼ぶ)と「話題のまとまりを枠で囲う」(以下、枠操作と呼ぶ)の 2 種類の操作別に、アンケートの回答を集計したものを表 4 に示す。

5.4 考察

5.4.1 計算機の自動推定精度はどの程度か

e からは計算機による自動推定を、ユーザが操作の参考にしたかどうか分かる。実験では、計 46 回の自動推定が

表 2 被験者 A の実験結果

順序	操作内容 (a)	b	c	d	e	h
1	3 番ハイライト	1	1		2 イ	3
2	5 番ハイライト	1	1	3	3	3
3	5-14 番囲う	2	2 イ	2 イ	2 ア	3
4	21-25 番囲う	2	2 イ	3	4	3
5	28 番ハイライト	1	1	3	1 イ	1 ア
6	32 番ハイライト	1	1	3	3	1 イ
7	32-44 番囲う	2	2 イ	3	2 ア	3
8	48 番ハイライト	2	1	3	3	3
9	58-64 番囲う	2	1	3	2 ア	3
10	65-76 番囲う	1	1	3	2 ア	3
11	87 番ハイライト	1	1	3	3	1 イ
12	88-94 番囲う	1	1	3	2 ア	3
13	150-158 番囲う	3	1	3	1 ア	3
14	159 番ハイライト	1	1	3	1 イ	3
15	159-166 番囲う	1	1	1 イ	2 ア	3
16	172 番ハイライト	1	1	3	1 イ	3
17	171-179 番囲う	1	1	1 イ	2 ア	3
18	188 番ハイライト	1	1	3	1 イ	3
19	186-199 番囲う	1	1	2 イ	2 ア	3
20	96-99 番囲う	1	1	3	4	3
21	100-111 番囲う	2	2 イ	3	3	3
22	124-129 番囲う	1	1	3	3	3
23	133-144 番囲う	2	2 イ	3	2 ア	3

表 3 被験者 B の実験結果

順序	操作内容 (a)	b	c	d	e	h
1	0 番ハイライト	1	1		2 ア	1 ア
2	5 番ハイライト	1	1	3	1 アイ	3
3	38 番ハイライト	2	1	3	1 アイ	3
4	46 番ハイライト	1	1	3	3	1 ア
5	66 番ハイライト	1	1	3	1 アイ	3
6	96 番ハイライト	1	1	3	4	2 ア
7	127 番ハイライト	1	1	3	3	1 ア
8	135 番ハイライト	1	1	3	3	1 ア
9	154 番ハイライト	1	1	3	1 アイ	3
10	173 番ハイライト	1	1	3	3	3
11	193 番ハイライト	1	1	3	1 アイ	3
12	0-3 番囲う	1	1	1 アイ	2 ウ	1 ア
13	4-8 番囲う	1	1	1 アイ	4	3
14	11-18 番囲う	2	2 ア	3	2 アウ	3
15	25-37 番囲う	2	1	1 イ	2 ア	3
16	38-45 番囲う	1	1	1 アイ	4	3
17	49-65 番囲う	2	2 イ	1 アイ	4	3
18	66-88 番囲う	1	1	1 アイ	4	3
19	127-134 番囲う	1	1	1 アイ	2 ウ	3
20	135-153 番囲う	1	1	1 アイ	2 オ	3
21	154-172 番囲う	1	1	1 アイ	3	3
22	173-192 番囲う	1	1	1 アイ	4	3
23	193-232 番囲う	1	1	1 アイ	3	3

行われているが、表 4 を見ると、その推定結果に対して、1「とても参考になった」が 10 件、2「少し参考になった」が 16 件見られる。これらについて、どう参考になったかの内

表 4 各選択肢が選ばれた回数

質問項目		被験者 A		被験者 B	
		a: ハイライト	a: 囲う	a: ハイライト	a: 囲う
b	1	8	7	10	9
	2	1	6	1	3
	3	0	1	0	0
c	1	9	9	11	10
	2	0	5	0	2
	3	0	0	0	0
	ア	0	0	0	1
	イ	0	5	0	1
	ウ	0	0	0	0
	エ	0	0	0	0
d	1	2	0	0	11
	2	0	2	0	0
	3	8	10	10	1
	ア	0	0	0	10
	イ	0	4	0	11
	ウ	0	0	0	0
e	1	4	1	5	0
	2	1	9	1	5
	3	4	2	4	2
	4	0	2	1	5
	ア	0	10	6	2
	イ	5	0	5	0
	ウ	0	0	0	3
	エ	0	0	0	0
	オ	0	0	0	1
h	1	3	0	4	1
	2	0	0	1	0
	3	6	14	6	11
	ア	1	0	5	1
	イ	2	0	0	0
	ウ	0	0	0	0

容を見てみる。イ「ヒントをそのまま採用できる」が合わせて 10 件あり、これらは推定成功とみなせる。ウ、エの「修正すれば済む」という回答は合わせて 3 件で、これらは成功とは言えないものの、ユーザの思う正解に近い推定ができたと言えさせる。ア「どこに注目するかを目安になった」は合わせて 18 件あるが、その中でイやウとの複数回答ではなくア単独で選ばれている例が 12 件ある。このア単独選択の例については必ずしも推定成功とは言えない以上から今回の実験では全操作 46 件中、イ、ウ、エのどれかが選ばれた 13 件について、正解ないしは概ね良好な解を推定できたと言える。反面、それ以外の 33 件については推定成功とは言えず、現状の計算モデルでは対応できない事例がまだ多く残っている。

操作の種類別に見ると、イ、ウ、エのどれかが選ばれたのはハイライト操作では 2 人合わせて 10 件なのに対し、枠操作では 3 件であり、発話のまとまりの境界を正確に推定することの難しさが見て取れる。一方ハイライト操作につ

いての 10 件は全てイ「ヒントをそのまま採用できる」であり、エ「ハイライトの位置を修正すれば済む」という回答は 0 件である。

5.4.2 計算機の自動推定はユーザの操作のヒントとして利用されるか

d と e の回答を比較すると、ユーザが操作のためにどんな情報を利用したかが分かる。

前項で見た e の回答から、計算機が提示する推定結果が操作のヒントとして利用されていたことが分かる。特にハイライト操作について多く見られた、イ「ヒントをそのまま採用できる」という例では、自動推定がユーザの操作の代りに匹敵する効果を生んだと見る事ができる。

枠操作について、ウ「まとまりの位置を修正すれば済む」という回答が数例見られ、正解ではないまでも、自動推定がヒントとしてユーザの操作を助けたと言える。

他方、最も多かったのはア「どこに注目するかを目安になった」の回答であり、そのうちアが単独で選択された例が 12 件あった。この 12 件については次のような可能性が考えられる。つまり、推定の失敗によって偶然作られたハイライトや枠であっても、それが表示されることにより、何の目安もない発話の羅列を見るよりも閲覧しやすくなる効果を生んだ可能性である。この効果をどう理解すべきで、今後何らかの形で利用を検討すべきかどうか、本稿では評価を保留したい。

次に、d と e の回答を比較してみると、ユーザが利用した情報に優先順位があることが分かる。つまり、自分の入力した情報が役に立つ (d の回答が 1) 場合には、計算機の提示するヒントは無視されるか、少し参考になったとしても目安にする (ア) 程度にしか扱われないことが多い。ここから、正確性の低い情報でも参考になることがあるが、正確性が高い情報ほど優先的に活用されるという傾向が読み取れる。

以上の検討をまとめると、計算機の自動推定はユーザの操作のヒントとして利用されるかについては、利用された例が少なからず認められた。

5.4.3 ユーザは自分の取り扱いやすい箇所から、操作しやすい要素を選んで入力・修正することができるか

a は被験者が行った操作であり、b と c からは操作の難しさに対するユーザの印象が分かる。表 4 を見ると、二人の被験者のいずれもハイライト操作よりも枠操作の時に、時間がかかったり決定に迷いを残したりすることが多いようである。また表 2、表 3 からは、同じ操作でも取り扱う箇所によって難しさに対する印象が変化していることが分かる。このように、取り扱う箇所や操作する要素によって、難しさに違いがある。

次に 2 人の被験者が行った操作の順番を見ると、両者の間に違いがあることが分かる。被験者 A は、上から順にハイライト操作と枠操作を行なうという順番をとってい

る。もう少し細かく見ると、順序 2-3, 6-7, 14-15, 16-17, 18-19 のように、ある発話系列に対してハイライトと枠を両方入力する場合、先にハイライト操作を行なって、次にそのハイライト箇所を含むように枠操作を行なっているという特徴が見られる。これに対し被験者 B は、全てのハイライト操作を先に行なってしまい、それが済んでから枠操作を行なうという順番をとっている。

この順序の違いが生じた一因は、提案モデルが「ユーザは自分の取り扱いやすい箇所から、操作しやすい要素を選んで入力・修正することができる」ような設計になっており、自然な形でユーザが対話状況を可視化できたためであると言える。

5.4.4 ユーザの入力・修正に応じて計算機が行う再計算では、推定精度が向上するか

h の回答から、ユーザの入力に伴う再計算がユーザにどのように受け止められたかが分かる。表 4 を見ると、全 46 回意味のある変化が起きたとは思わないという回答が二人合わせて 37 件と多い。これはヒントが変化しなかったという解釈だけでなく、ユーザがヒントの変化に注意を向けていなかったという解釈も可能であるが、いずれにせよ作業の助けになるような計算結果の向上があったとはみなされない。一方、二人合わせて 8 件と多くはないが、望ましい変化をしたと回答された例がある。これらの例については、ユーザの入力を利用した再計算が推定精度の向上につながっていると言える。変化の種類を見ると、ハイライト操作の結果に伴って枠が適切になった例が 8 件中 7 件とほとんどである。

他方、望ましくない変化をしたという回答が 1 件ではあるが見られる。この例はハイライト操作に伴って枠の推定がかえって不適切になったと判断されたものである。再計算によってユーザの入力に整合する候補に絞りこまれても、その中からより適切な候補が選ばれなければ、かえってユーザの求める結果から遠ざかってしまうことになると思われる。

5.5 実験のまとめ

ここまでの考察をまとめる。計算機の自動推定精度については、成功した例が一部見られたものの、現状の計算モデルでは対応できない事例がまだ多く残っている。計算機の自動推定はユーザの操作のヒントとして利用された例が少なからず認められる一方で、それらはユーザが自ら書き込んだ情報ほどには役立っていない。ユーザは自分の取り扱いやすい箇所から、操作しやすい要素を選んで入力・修正することができるかという観点については、効果的に機能している事例が確認された。ユーザの入力・修正に応じて計算機が行う再計算では、推定精度が向上するかという観点については、向上した例が少なからず見られた一方で、再計算がかえって推定精度を下げてしまう例も一部見

られた。

以上、本稿で設計した協調モデルが効果的に機能している事例が少なからず認められたとともに、現状の設計ではまだうまく解決できない事例が多く残っていることを確認し、本実験のまとめとする。

6. おわりに

本稿では、会議中の対話状況を可視化するヒューマン・コンピュータ協調モデルの設計を行った。そして設計した協調モデルに基づいたプロトタイプシステムを試作し、予備実験により提案モデルの評価を行った。実験の結果、提案モデルにより効果的な協調が生じた事例が観測でき、実際に提案モデルが有効に機能する場面があることが確認できた。

同時に、実験により設計の不備が現れている点等、今後の課題となる多くの知見も得られた。また今回行った検証は専ら被験者の主観に頼ったものであり、見出された傾向も統計的に結論付けられたものとは言えないことから、今後客観的・定量的な検証に取り組みたい。

謝辞 本研究の一部は、独立行政法人科学技術振興機構 (JST) 戦略的創造研究推進事業 (CREST) 「マルチモーダルな場の認識に基づくセミナー・会議の多層的支援環境」の助成を受けて行った。

参考文献

- [1] 鈴木健：究極の会議，ソフトバンククリエイティブ (2007)。
- [2] 秋田祐哉，河原達也：会議音声の自動アーカイブ化システム，情報処理学会研究報告，SLP，音声言語情報処理，2000(119)，pp. 61-66，(2000)。
- [3] 横江優貴，伊藤義道，馬場口登：参加者のインタラクションを可視化したマルチメディア議事録の作成，情報処理学会研究報告，CVIM，[コンピュータビジョンとイメージメディア]，2008(27)，pp. 121-126，(2008)。
- [4] 久保田秀和，齊藤憲，角康之，西田 豊明：会話量子化器を用いた会話場面の記録，情報処理学会論文誌，48(12)，pp. 3703-3714，(2007)。
- [5] Nagao, K., Kaji, K., Yamamoto, D. and Tomobe H: Discussion Mining: Annotation-Based Knowledge Discovery from Real World Activities, PCM 2004, pp. 522-531, (2004)。
- [6] 高悠史，吉本廣雅，近藤一晃，中村裕一：遠隔会議の実時間支援に向けた対話状況の可視化～対話の結束性に基づく表現の有効性～，電子情報通信学会技術研究報告，MVE，マルチメディア・仮想環境基礎，111(38)，pp. 13-18，(2011)。
- [7] 石崎雅人，伝康晴：談話と対話，東京大学出版会 (2001)。
- [8] 坊農真弓，高梨克也編：多人数インタラクションの分析手法，オーム社 (2009)。
- [9] Yu, Z. and Yu, Z. and Ko Y. and Zhou X. and Nakamura Y.: Inferring Human Interactions in Meetings: A Multimodal Approach, UIC 2009, pp.14-24, (2009)。
- [10] 角康之：マルチモーダルデータに基づいた多人数インタラクションの構理解，情報処理学会研究報告，HCI，ヒューマンコンピュータインタラクション研究会報告，2011-HCI-145(7)，1，(2011)。