

議論支援のためのマルチモーダル情報処理機構

庄司 祐希[†] 林 佑樹[†] 瀬田 和久[†]

大阪府立大学大学院人間社会システム科学研究科[†]

1. はじめに

議論の創造性を高める上で、参加者は互いの状態に気を配り、相互理解やコミットメントを引き出し合うことが重要である。一方でリアルタイムに進行する議論において、自他の意見を相対的に捉えながら、他者の思考や心情を汲み取りタイムリーに働きかけることは容易ではない。他者の様子がカメラ映像に限定される分散環境下での議論ではいっそう難しく、アイデア共創の契機を逸失してしまう可能性がある。

そこで我々は、議論中に交わされる参加者個々の視線や発話といった「マルチモーダル情報」に着目し、複雑な議論インタラクション状況の一端を計算機が捉え、状況に応じた適切な介入を実現する議論支援フレームワークの開発を進めている[1]。

本稿では、議論時に集積される各参加者の視線・発話情報に基づき、情報システムが多人数インタラクションを構造的に解釈し、それらを参照した適応的な助言を生成するマルチモーダル情報処理機構を提案する。

2. アプローチ

2.1 先行研究

先行研究[1]で提案した議論状況検出ルールと助言生成ルールを、議論時に交わされるマルチモーダル情報に適用することを考える。前者は角らのインタラクションの階層的解釈モデル[2]を参照モデルとし、発話や視線といった参加者個々の原始的なインタラクション要素を捉える初期設定ルールと、それらを統合した高次インタラクションを捉える統合ルールから構成される。後者は議論状況検出時に誰にどのような助言内容を生成・提示するかを規定している。これらのルールは様々な議論セッションの構成に適用可能で、共有・比較・洗練できるよう可読性の高い宣言的形式で規定されている。なお、各ルールの詳細については[2]を参照されたい。

2.2 マルチモーダル情報処理機構の設計指針

議論インタラクション状況をマルチモーダル

情報に基づき捉え、議論の状況に沿った助言生成を実現するために以下の機能要件を満足する必要がある。

[要件1]インタラクション要素のリアルタイム検出機能：リアルタイムの議論において逐次発現する視線行為や発話行為などの参加者毎のマルチモーダルデータ系列を検出する必要がある。視線情報においては、注視対象の意味内容（セマンティクス）を捉える仕組み[3]を活用し、議論状況を捉える手立てとする。

[要件2]議論状況の構造的解釈機能：様々な粒度の議論状況を計算機システムが同時並列的に解釈し、その粒度に応じて階層毎に管理できるとともに、検出した議論状況データの構造的関係性（解釈根拠）も蓄積可能な仕組みが必要である。

[要件3]議論状況に即した助言提示機能：要件1、要件2を充足する仕組みによって検出される議論状況データを参照して、適切な助言の生成および、働きかけの意図に沿った助言を提示する機能が必要である。

3. マルチモーダル情報処理機構の枠組み

本稿で提案する処理機構の概念図を図1に示す。本研究では、分散環境下で交わされる視線・発話情報を計測可能なCSCLシステム開発プラットフォーム[4]を利用することを前提に計測される断続的なデータ（Raw Data Layer: RD層）に対して、ボトムアップ的にインタラクションを解釈するために、Working Memory(WM)をインタラクションの階層（Interaction Primitive Layer: IP層、

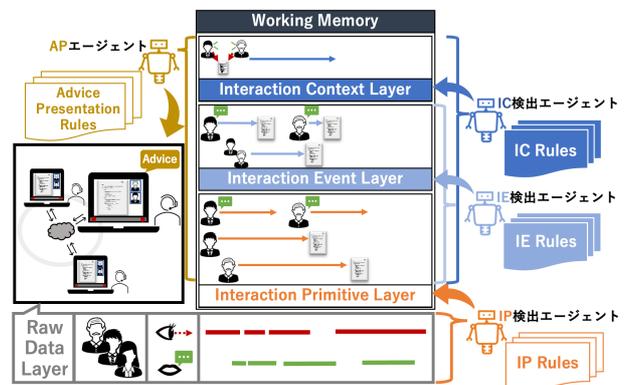


図1 マルチモーダル情報処理機構の概念図

Multimodal Information Processing Mechanism for Discussion Support

[†]Yuki Shoji, Yuki Hayashi, Kazuhisa Seta, Graduate School of Humanities and Sustainable System Sciences, Osaka Prefecture University

Interaction Event Layer: IE 層, Interaction Context Layer: IC 層) 毎に区分し, 検出されたインタラクションデータを管理している。

3.1 インタラクション要素検出機能

RD 層に蓄積されるデータとドキュメントセマンティクス情報[3]から, 参加者個々の原始的な振舞いを初期設定ルールに基づき検出する(要件 1)。初期設定ルールには, 例えば「提案者が 30 秒間に 40%以上資料領域 X を注視している」, 「参加者が 20 秒間に 50%以上発話している」といった検出条件が規定されている。IP 検出エージェントはタイムリーに蓄積される RD 層のデータを監視しており, 設定時間(例: 1 秒)毎に各初期設定ルールが対象とする発話/注視割合を対象時間分遡って算出する。初期設定ルールと合致する場合には, 「提案者が 50%以上資料エリア X (“実践目的”) を注視している」といった実データが検出され, WM の IP 層に書き込まれる仕様となっている。

3.2 議論状況の構造的解釈機能

WM に書き込まれたインタラクションデータに統合ルールを適用し, より高次のインタラクション状況を検出する(要件 2)。統合ルールには, 例えば「(a) 提案者が資料領域 X を見ながら説明している」といった IE や, 「(b) 指導者と提案者が資料領域 X について活発に議論したのち, 資料領域 Y についての議論に移った」などの IC を検出する条件を, 複数の状況の重複区間の積集合(Overlap), 和集合(Union), 前後関係(Before/After)といった関数を用いて規定できる。IE 検出エージェントは WM の IP 層と IE 層を, IC 検出エージェントは WM の全層を監視対象とし, 新たなインタラクションデータが各層に書き込まれることをトリガとして, 各エージェントが対象とする統合ルールを自律並列的に照合する仕組みを備えている。例えば上述の(a)の検出には, このルールの前件部で Overlap 関数の引数として指定された「提案者が話している」「提案者が資料領域 X を見ている」という実データ(IP 層)が生じた場合に, その重複区間が WM 上の IE 層に蓄積されることになる。

3.3 助言提示機能

助言提示ルールに基づき, 統合ルールで検出されたインタラクション状況に即した助言を生成し, インタフェースに提示する(要件 3)。助言提示ルールでは, 対象とする議論状況が検出された際に, 誰に, いつ(タイミング), どのような側面(認知/行動/情動)を対象とする助言を提示するかを規定できる。例えば「指導

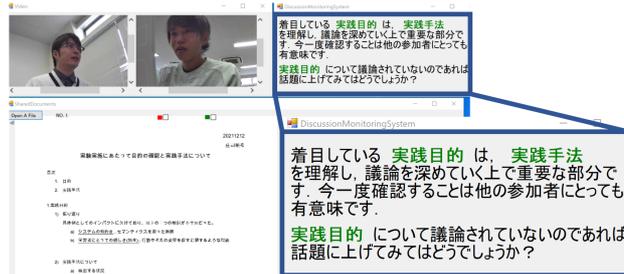


図 2 助言提示画面に表示される助言例

者と提案者が資料領域 Y について議論している時, 他の参加者が資料領域 X に注視しているという IC 層に該当するインタラクション状況が生じた際に, 「他の参加者に対して即時に注視部分(資料領域 X)の発言行動を促す」ルールを設定できる。助言提示ルールを対象とする助言提示(AP)エージェントは, 統合ルールで検出されたデータが WM に書き込まれることをトリガとして, 議論セッションに適用する助言提示ルール群を照合する。ここで, 例えば上述の助言提示ルールが合致した場合, 「実践目的について議論されていないのであれば, 話題に上げてみてはどうでしょうか?」といった行動を促す助言がテンプレートに基づき生成され, この助言が対象者(他の参加者)の助言提示画面に(議論が一段落したタイミングではなく)即時に表示されることになる(図 2)。

4. まとめと今後の課題

議論インタラクション状況を情報システムが捉え, 適応的な介入(助言提示)を実現するためのマルチモーダル情報処理機構を提案した。今後の課題として, システムの実践利用を通じた助言提示の有効性を検証する必要がある。

参考文献

- [1] 庄司祐希, 林佑樹, 瀬田和久: “ドキュメントセマンティクスとマルチモーダル情報に基づいた議論支援フレームワークの検討”, 第 46 回 JSiSE 全国大会, D3-4, pp. 155-156, 2021.
- [2] 角康之, 矢野正治, 西田豊明: “マルチモーダルデータに基づいた多人数会話の構造理解”, 社会言語科学会誌, Vol.14, No. 1, pp. 82-96, 2011.
- [3] 正野敦也, 林佑樹, 瀬田和久: “議論の内省を促すリフレクション支援環境—ドキュメントセマンティクスとマルチモーダル情報を活用して—”, JSiSE 2020 年度特集論文研究会, Vol. 35, No. 7, pp. 13-20, 2021.
- [4] 杉本葵, 林佑樹, 瀬田和久: “言語・非言語アウェアな CSCL システム開発プラットフォーム”, 電子情報通信学会論文誌, Vol. J101-D, No. 4, pp. 713-724, 2018.