

ユーザの注視行動に基づく会話参加態度の推定 -会話エージェントにおける適応的会話制御に向けて-

石井 亮[†] 中野 有紀子[†]

東京農工大学 工学府 情報工学専攻[†]

1 はじめに

対面会話において、聞き手が関心を持って会話に参加していることを、話し手は聞き手の動作や視線から察知し、積極的に参加していない様子であれば、話題を変えるなど、会話の内容や方略を調整している。人対会話エージェントのインタラクションにおいて、このような柔軟な会話を可能にするには、システムがユーザの会話参加態度を認識・推定する機能を有する必要がある。

ユーザの参加態度推定方法として、[1]では、ユーザの頭部運動の情報から、ロボットとの会話への engagement を判断している。また、情報提示システムの研究では、より詳細な視線データを用い、視線移行パターンからユーザの迷いや[2]、興味対象を推定する方法も検討されている[3]。しかし、会話中のユーザの詳細な視線行動から会話参加態度を推定する方法はまだ確立されていない。

そこで本研究では、ユーザの注視対象遷移パターンに着目し、会話エージェントとのインタラクションにおける、ユーザの会話参加態度の推定方法を提案する。まず、Wizard-of-Oz システムによる実験から得られた視線データと会話参加態度に関する主観的評価データを分析し、理想的な会話参加態度から逸脱しているとみなされる視線遷移パターンを同定する。次に、エージェントとの会話中にリアルタイムに取得される視線データから、個人差を考慮しながら逸脱度の高い状態を発見することができる個人適応型会話参加態度推定アルゴリズムを提案する。

2 会話実験による収集データ

会話エージェントとの会話におけるユーザの参加態度を分析・推定する上で有用な言語・非言語行動を収集するために、Wizard-of-Oz システムを用いて対話収録実験をおこない、平均約 16 分の会話を 10 会話収録した[4]。ここでは、その中から以下のデータについて分析を行った。

- ユーザならびにエージェントの発話テキスト(全データの発話総数はエージェントが 951 回、ユーザが 61 回)

- エージェントのジェスチャならびに視線行動

- ボタン押下方式を用いた、ユーザ自身と観察者による、会話への飽きに関する主観的評価(会話に飽きている場合に主観ボタンを押下)

尚、ユーザの視線データはアイトラッカ(Tobii-X50)によって取得された 50fps のデータのうち、半径 20pixel 以内の円内に 20ms 以上視線が停留した区間のみを採用した。最終的に、これら全ての情報を統合し、言語・非言語情報の共起関係を視覚的にとらえられるコーパスデータを作成した。

3 分析：会話参加態度と視線遷移パターン

収集した会話の観察結果[4]に基づき、ユーザが適切なタイミングで、エージェントからのフィードバック要求に答えて相互注視をしたり、エージェントの視線方向に従って説明対象に共同注視を行うといった行為を行っている場合を「積極的に望ましい会話参加態度」と定義した。逆に、エージェントの会話に飽きたり興味がなくなることにより、これらの行為が行われない状態を「望ましい会話参加態度からの逸脱」とみなす。例えば頻繁によそ見をすることは望ましくない会話参加態度であろう。そこで、会話参加態度に応じて頻出する視線パターンが異なると仮定し、会話への飽きに関する主観的評価と視線パターンの関係について分析を行った。

エージェントが説明対象を注視しながら発話している時のユーザの注視行動データとして、時間的に連続した注視対象遷移の 3-Gram を用いた。より詳細には、注視終了から 200ms 以内に同じ対象が再び注視されたときは、注視行動が継続しているとみなし、注視データが 1 秒以上存在しない場合は、未完の 3-Gram として破棄した。また、視線行動の変化が主観ボタン押下よりも先に表出するという仮定の下、3-Gram 開始から終了後 5 秒後までの間に主観ボタンが押下されたか否かをデータとした。

注視対象 3-Gram と主観ボタン押下率を図 1 に示す。図中の記号はそれぞれユーザの注視対象が、T はエージェントの説明対象に(共同注視状態)、

Estimating the Degree of Conversational Engagement based on User's Gaze Behaviors in Human-Agent Interaction: Towards Adaptive Dialogue Management in Conversational Agents
Ryo ISHII[†] and Yukiko NAKANO[†]

[†]Graduate School of Tokyo University of Agriculture and Technology

AH はエージェントの頭部, AB はエージェントの頭部以外の上半身, F はそれ以外の対象物 (F1≠F2≠F3) にあることを表す. 図 1 に示すように, 注視対象 3-Gram によって, 主観ボタン押下率に大きな差異が見られることから, ボタン押下率の高い 3-Gram は理想の会話参加態度からの逸脱行為であると言える. この結果から, 注視対象 3-Gram による会話参加態度推定の可能性が示された. また, 押下率上位の 3-Gram の要素に T が含まれていないことから, 説明対象に共同注視をおこなわない行為が会話参加態度として望ましくないことも実証された.

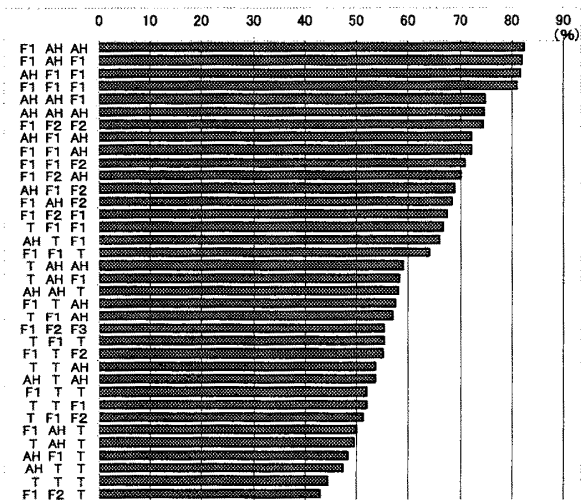


図 1 注視対象 3-Gram と主観ボタン押下率

4 会話参加態度の推定アルゴリズム

以上の分析結果をふまえて, ユーザの飽きと相関の高い注視対象遷移パターンから, 理想の会話参加態度からの逸脱を推定する方法を提案する. 各ユーザの 3-Gram の分布を調べたところ, 主観ボタン押下率の高いパターンから低いものまで分布が広がるユーザと, 押下率が中から低いパターンを中心に分布しているユーザがいた. このため, このようなパターン分布の個人差を考慮する推定モデルが必要であると考えた. そこで, 初めてシステムを使うユーザに対して, この分布の違いを考慮した個人適応が迅速に行われるシステムを実現するため, エージェントの説明開始から最初の 120 秒間のユーザの注視遷移パターンから, 逸脱度の閾値を算出し, それ以降は, その閾値を越える注視遷移パターンが表出したときに理想の会話参加態度から逸脱していると判定する事とした.

個人ごとの閾値を算出するために, まず, はじめの 120 秒間に現れた 3-Gram を k-means 法を用いて 4 つにクラスタリングした. この時, 各 3-Gram の継続時間長の合計による重み付けを行った. そして, 最も逸脱度の高いクラスと 2 番目に高い

クラスターの重心の midpoint を閾値とした. 図 2 に実際の 3-Gram のプロットとクラスタリングの結果を示す. このとき, 閾値は 69.1 である.

この推定アルゴリズムを用いて, 会話開始後 120 秒以降の, 会話参加態度 (主観ボタンが押された状態) の予測を行った. 10 秒間刻みでの推定の結果, 適合率 79.4%, 再現率 71.2%, F-measure 75.1% という結果が得られ, 推定アルゴリズムの有用性が示された.

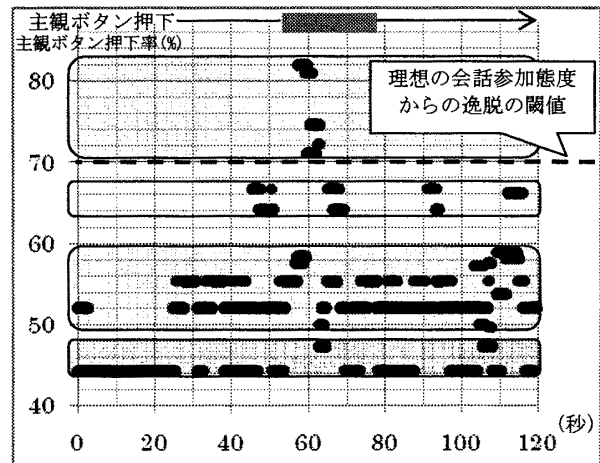


図 2 3-Gram のプロットとクラスタリング結果

5 まとめ

本稿では, ユーザの会話参加態度に応じて適応的に振舞う会話エージェントを目指し, ユーザの会話参加態度をユーザの視線行動, 特に注視対象 3-Gram から推定するアルゴリズムを提案した. 今後は, 人対会話エージェント会話において, ユーザが理想の会話参加態度から逸脱した時に, エージェントがどのように振る舞えば良いかを検討するとともに, 推定アルゴリズムを会話機構に組み込み, 対話実験をおこなう予定である.

参考文献

- [1] Sidner, C.L., et al.: Where to Look: A Study of Human-Robot Engagement; In Proceedings of *ACM International Conference on Intelligent User Interfaces (IUI)*, pp. 78-84, 2004.
- [2] 高木: 視線の移動パターンに基づくユーザの迷いの検出-効果的な作業支援を目指して; 情報処理学会論文誌, Vol.41, No.5, 2000.
- [3] 水口 他: Mind Probing: システムからの積極的な働きかけによる視線パターンからの興味推定; 情報処理学会ヒューマンコンピュータインタラクション研究会(SIGHCI), 2007.
- [4] 石井 他: ヒューマン・エージェントインタラクションにおける会話関心度推定のためのユーザ視線パターンの分析; HAI シンポジウム 2007, 2007.