

対話知能学

ニュースレター (Vol. 16)



1. 研究紹介

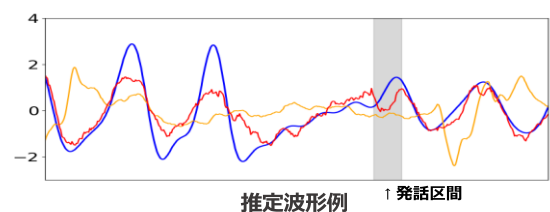
■ 対面動画からの非接触呼吸推定と呼吸情報の対話への応用(A01班)

人の音声対話は呼吸を二次的に利用した活動で、呼吸と対話は切っても切れない関係にある。A01 班の船越らの研究グループでは、これまで対話研究でほとんど着目されてこなかった呼吸情報の対話システムへの活用に取り組んでいる。発話の前兆が呼吸波形に現れることは人・人会話の研究から知られており、これを検出することで、対話システムはユーザの発話開始を予測しながら円滑に対話することができる。また、ユーザの呼吸に何らかの形でシステムが同調したり、システムから擬似的な呼吸を提示したりすることで、A01 班の目的である関係構築に寄与できる可能性がある。

これまでの研究成果として、センサーベルトを用いて呼吸計測を行いながらの対話動画を 80 名の参加者を対象に収集した。この内の一部は推定タスク用のデータセットとして右図のように個人情報保護した上で一般公開する。また対話動画自体も



呼吸情報付き研究用対話コーパスとして限定公開する予定である。データセットを用いて深層学習による呼吸波形推定に取り組み、一般的な web カメラを用いるだけで、特別なセンサーを装着しなくても被写体の呼吸波形を推定できること、さらに推定波形を元に発話開始を予測できることも示した。



青: 実呼吸波形, 黄: Optical Flow-based, 赤: DNN-based

今秋はさらに、上記の発話予測で実際に人・ロボット対話を円滑化できることを実証するとともに、ロボットによる疑似呼吸提示が関係性構築にどのように寄与できるのかを探るための新たな参加者実験を進めている。

船越 孝太郎(東京工業大学/A01 班 対話継続関係維持研究グループ公募班代表者)

■ 対話システムの評価と応用に向けたユーザのマルチモーダル行動モデリング(A02班)

対話システムの評価は、重要課題のひとつである。本研究は、対話を通じて醸成されるユーザのシステムに対する主観評価(例:対話満足度)を、ユーザの対話中のふるまいや発話から得られるマルチモーダル(MM)情報から自動推定する方法を確立することを目的とする(図1)。この研究では、システムに対してユーザの対話満足度が高い/低い、楽しんでいる/困惑している場合に、特有のふるまい・発話が観測されるという仮説に基づいている。この仮説に基づき、ユーザの音声・表情・ジェスチャ情報を含む MM 情報のセンシング技術に着目し、研究を進めてきた [1-3]。

本稿では現在までの研究で明らかになったことを報告する。発話内容を示す言語情報は主観評価を予測するために最も有効な特徴量であり、表情特徴量(アクションユニットと呼ばれる顔の部位の行動パターン)を加えることで精度が向上した。また対話中に変化するセンチメントの時系列を機械学習に援用することで、「対象の対話システムは対話を上手

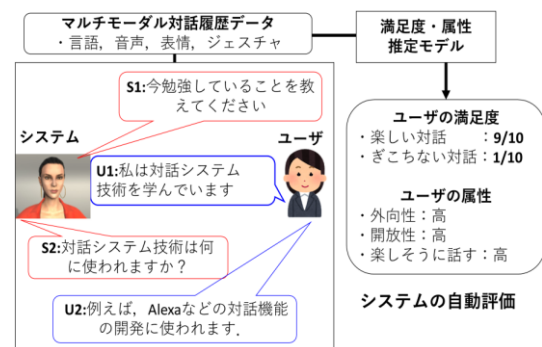


図 1:マルチモーダル情報に基づく主観的評価値の推定

く調整できていた」というユーザの評価スコアの高低を最高 0.791 (F 値)で推定できることがわかった [1]. 対話相手の Woz 操作者が同様の評価を行った結果、その精度は 0.71 であったことを考慮すると、機械学習と MM センシングによる対話評価は一定の有用性を持つことがわかる。この研究成果に基づいて明らかになった2つの課題に対して研究を行った。



(i) ふるまいの個人差: ユーザから観測される MM 情報は対話評価に関連するだけでなく、ジェンダー、年齢、性格、等の個人差によっても大きく変動する（例：対話評価に関わらずよく笑う/ポーカフェイスのユーザが存在する.）。この問題に対して、本研究では話者、ジェンダー、異なる言語で発話されたデータ間で共通の特徴量表現を学習することで、個人差に頑健な推定を可能とする Adversarial Domain Generalized Transformer (ADoGT)を提案した[2]。このモデルでは Adversarial Training と呼ばれるニューラルネットの学習方法を採用し、主目的タスク（対話評価の推定）における Loss を最小化すると同時に、その他の（年齢、ジェンダ等）識別タスクの Loss を減らさないよう（個人差要因を識別しないよう）学習する。4種類の音声感情認識タスク用のデータセットを用いて、評価を行った。訓練データとは異なる言語で発話されたデータでテストした場合でも、精度低下を軽減できることを示した。現在、この手法を対話評価推定タスクに応用する方法を研究している。

(ii) 対話評価推定結果の利用: システムがユーザからの対話評価を上手く推定できたとして、システム自身が獲得した対話評価結果をどのように用いて、対話を改善すればよいかは自明ではなく、対話状況、タスクによって適した対話生成系を研究する必要がある。システムにより対話評価結果を対話戦略に活かす事例として、インタビュー対話を対象に、インタビュー中に変化するユーザの発話意欲を自動推定するインタビューロボットを開発した [3] (図 2)。このシステムでは発話意欲推定結果に基づき、意欲が高ければ質問を掘り下げ、意欲が低ければ異なるトピックの質問をするよう質問戦略を切り替える。開発したロボットと 27 名の参加者との対話実験を通じて、インタビューロボットの発話意欲推定と適応的な質問戦略の効果を検証した。発話意欲の推定に基づき質問戦略を切り替える条件と、ランダムに質問戦略を切り替える条件とで比較し、参加者のロボットへの印象、参加者の発話意欲の変化、発話意欲の推定精度を検証した結果、

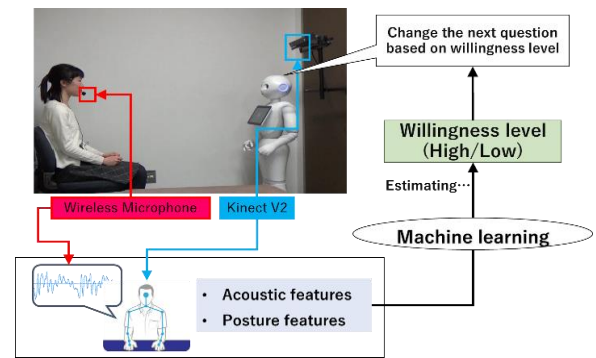


図 2: 発話意欲推定に基づき適応的質問選択を行うインタビューロボット

次のことが明らかになった。(1) 発話意欲の高低は平均 72.8% の精度で自動推定できる。(2) 提案する発話意欲に基づく適応戦略は、適切な聞き手としての態度を印象付けられる。(3) 意欲の高い発言を計数した結果、提案する発話意欲に基づく適応戦略は、有意に意欲の高い発言を増やせる。今後、対話評価の精緻な推定方法を研究すると共に、対話評価推定をいつ行い、推定結果をいかに対話戦略に取り入れるのかを探索し、研究を進めたい。ここまでの成果は([1]-[3]を含め)、当該分野のトップランク会議、ジャーナルである、ACM ICMI2022, IEEE Trans. Affective Computing に各 2 件採択された。

[1] Wenqing Wei, Sixia Li, Shogo Okada: Investigating the relationship between dialogue and exchange-level impression. ICMI 2022: 359-367

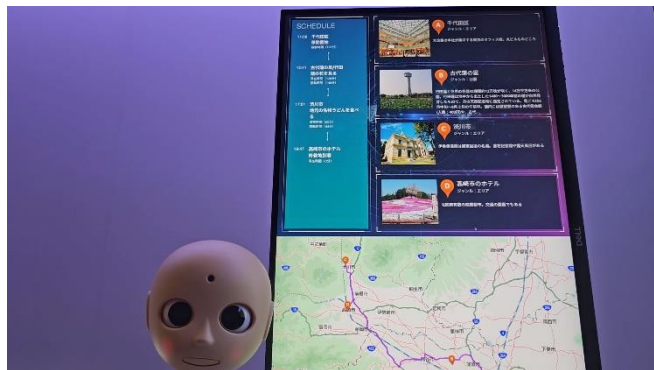
[2] Yuan Gao et.al. Adversarial Domain Generalized Transformer for Cross-Corpus Speech Emotion Recognition, IEEE Trans. on Affect. Comp., 2023

[3] Fuminori Nagasawa, et.al. Adaptive Interview Strategy Based on Interviewee's Speaking Willingness Recognition for Interview Robots, IEEE Trans. on Affect. Comp., 2023.

岡田 将吾 (北陸先端科学技術大学院大学 / A02 班 対話理解生成研究グループ公募班代表者)

■日々の困りごとの解決をサポートする対話システム(A03班)

対話システムが人と共生し日々の生活を支援するパートナーとなるためには、人との関係を構築する能力に加え、日々の困りごとの解決をサポートする能力も重要である。この課題に対し、我々はドライブ旅行中の困りごとの解決を念頭に、柔軟にユーザの状況を理解し旅行のスケジュールを自律的に計画するシステムを開発し NTT R&D フォーラムにて展示した。動作画面を下図に示す。



システムには、出発地や終着地、出発時刻の情報に加え、希望するアクティビティや食事などの要求や、チケットの時間など時間的な制約条件を自由に入力できる。システムはそ

れらを LLM を利用して理解し、移動時間を空欄とする仮のスケジュールへ変換する。次いで、ナビ API を呼び出し、移動時間を埋めることで、全体の移動スケジュールを仮組みする。この状態で、上記で入力した時間的な制約条件や食事の時間などの条件を考慮し、スケジュールの順序を入れ替えたり、食事のスケジュールを追加したり、一部のアクティビティの時間を調整するなどして、時間的に整合性の取れたスケジュールへ修正する。スケジュールの流れが定まった後、昼食を取る店や訪れる観光スポットについて、システムが観光 DBを検索して候補を例示し、対話を通して具体的な訪問場所を決定する。

上記は LLM 単独では 2023 年現在では実現不可能な処理であり、外部のツール・API や、それらを制御する別の LLM との連携が必須である。多数のシステムと人が協調し柔軟に課題を解決する、対話知能学がめざす未来の実現に向けて、引き続き研究に取り組んでいく所存である。

杉山 弘晃(NTT/A03 班 行動決定モデル推定研究グループ計画班代表者)

■ロボット倫理の論じ方を問い直す——「証拠に基づくロボット倫理」と「クリティカル・ロボティクス」(A04班)

筆者(呉羽)は、人間機械共生社会に向けた法社会規範の提案を目指す A04 班の中で、特にロボットに関する倫理について研究している。この過程で、哲学・倫理学を含む人文社会科学とロボティクスの協働のあり方について提言を行ってきた。

まず、ロボット倫理と文化の関係を主題とした研究の中で、「証拠に基づくロボット倫理」を提唱した。日本人はアニミズムを信奉しているからロボットのような人工物にも親和的だ、という「テクノアニミズム論」が世間に流布しているが、筆者は倫理学の知見と歴史学、宗教学、人類学、そして HRI の最新の証拠に基づき、この説が誤っていることを示した(『対話知能学ニュースレター vol. 8』の「3. 顕著な業績」の項を参照)。さらに現在は、ロボットユーザーに対するステレオタイプ・偏見(例えば、「ロボットユーザーは孤独な人々だ」、というもの)の問題を扱い、同様に証拠に基づく議論の必要性を訴えている。

また、ロボットを用いて社会に受け入れられた価値観に対して問題提起を行う「クリティカル・ロボティクス」を提唱し

た。このアプローチは、近年謝罪機能を実装したロボットが登場していることを踏まえて、自らの振舞いに責任を負えないロボットに謝罪をさせることの是非について考察した論文で唱えたものである。同論文では、結論として、人間社会の謝罪慣行に照らしてロボットに謝罪をさせることが許容される場合があると論じつつ、この慣行そのものを問い直す必要性を主張し、またロボットがその有効な手段になると論じた。(『対話知能学ニュースレター vol. 15』の「3. 顕著な業績」の項を参照)。そのためにも、今後は、クリティカル・ロボティクスの方法論を整備して、それをロボティクスの新しいアプローチとして確立しようと考えている。

ロボット倫理は急速に注目度を高めているが、そのためには哲学・倫理学を含む人文社会科学とロボティクスの両者の協働が求められる。その協働のあり方自体がいまだ未整備であり、それについて積極的な提案を行っていくことが重要だと考えられる。

呉羽 真(山口大学/A04 班 人間機械社会規範研究グループ計画班分担者)

2. イベント報告・告知

■対話システムライブコンペティション 6

2023 年 12 月 13 日(水), 14 日(木)に,「第 14 回対話システムシンポジウム」(人工知能学会 言語・音声理解と対話処理研究会(SLUD))の一環として表記のイベントを開催した(会場は国立国語研究所 講堂, 同時にオンラインで配信)。「対話知能学」領域からは東中竜一郎(A02 計画班研究代表者)が登壇した。

対話システムライブコンペティションとは, オーディエンスの前で実際に対話システムを動作させ, 評価を行うイベントである。前回(「対話システムライブコンペティション 5」)同様, マルチモーダル対話システムを対象としている。システムを実装するためのツールとして, 対話理解生成研究グループ(A02班)が研究開発を行うマルチモーダル対話システムのミドルウェアが提供されるため, 開発者是对話のロジックに注力できる。また, 新学術領域研究で開発された CG エージェント(CGErica)をアバターとして活用している。



オンライン, オフライン合わせて 200 名程度の参加者に来ていただいた。コンペティションの結果は当イベントの公式ページに掲載されている。<https://sites.google.com/view/dslc6/%E7%B5%90%E6%9E%9C>

■ 対話ロボットコンペティション 2023

2023 年 12 月 23 日(土)に, 日本科学未来館 7 階 未来館ホールおよびイノベーションホールにて,「対話ロボットコンペティション 2023」本選会を開催した。「対話知能学」領域からは東中竜一郎(A02 計画班研究代表者), 港隆史(A02 計画班研究分担者)が登壇した。また, 本コンペティションには本領域から多数のチームが参加している。

対話ロボットコンペティション 2023(DRC2023)は対話ロボットの性能を競うコンペティションである。参加者は, 旅行代理店における対話タスクにチャレンジする。このタスクでは, カウンターセールス役となったロボットが, 対話を通してお客様の要望に応える。

本コンペティションは予選会と本選会からなり, 予選会は実際の旅行代理店の実店舗(JTB 名古屋ユニモール店, 及び, JTB トラベルゲート天神店)に対話ロボットを設置して実施した。予選会には全 12 チームが参加した。その様子は NHK(東海 NEWS WEB)で取り上げられた。(東海 NEWS WEB【特集】AIが「接客」に挑戦?“対話の力”はどこまで <http://www3.nhk.or.jp/tokai-news/20231121/3000032878.html>)

本選会への出場が決定したのは MIYAMA, ROS, JERRY_cis, SBIntuitions の 4 チームである。本選会では, 「対話知能学」領域関係の研究者と旅行代理店業務の専門家が審査員となり, 4 チームによる 1 次審査の結果, ROS と SBIntuitions が 2 次審査に進んだ。この 2 チームによる 2 次

審査の結果, 最終的に SBIntuitions が最優秀賞を獲得した。

本選会の様子は Zoom ウェビナー形式での配信も行い, 50 名ほどの会場参加者と 40 名ほどのオンライン参加者に見ていただくことができた。ポスターセッションでは, チーム間の交流も進んだ。コンペティションの結果は当イベントの公式ページに掲載されている。<https://sites.google.com/view/dialogrobotcompe3>

また, 当コンペティションは NHK の番組でも取り上げられた(12 月 22 日 午前 6:30~「おはよう日本」(全国放送) <https://www.nhk.jp/p/ohayou/ts/QLP4RZ8ZY3/episode/te/7YZ1QLRG1K/>)。



■ 対話ロボットとコミュニケーション

2023年12月23日(土)に、クロスパル高槻にて、標記のイベントを開催した。本企画の趣旨は、対話機能を持つ最先端のソーシャルロボットを一般来場者が体験できる場を提供し、対話を行うロボット技術や未来のソーシャルロボットに求められる機能を考える機会を設けることである。とりわけ、言葉だけでなく、対話の雰囲気を作り出す非言語情報に注目したコミュニケーション力、通称“コミュ力(りょく)”の体験(バトル)してもらい、最先端の対話ロボット技術を広く発信することがねらいである。

当日は「対話知能学」領域から酒井和紀(A01 計画班研究協力者・大阪大学)、吉川雄一郎(A01 計画班研究分担者・大阪大学)、瀬島吉裕(A01 公募班研究代表者・関西大学)らによって準備された対話ロボットの展示に加え、「子どもと考えるロボット共生社会」と題した特別講演も行われ、本領域から瀬島吉裕が登壇した。講演では、chatGPT に代表される生成 AI の可能性と限界について解説するとともに、

未来のソーシャルロボットに求められる機能を紹介した。講演の最後に、聴講していた子どもに未来のロボットについて意見や感想を求め、多くの子どもたちが「人間らしさ」をキーワードにしていることを共有した。また、ビンゴ大会やワークショップも併せて行われ、150 名を超える多くの来客で賑わった。

イベント公式ページ(関西大学):<http://www2.kansai-u.ac.jp/hri/GCS/>



■ 話者属性を考慮して言語表現の制御を行う対話システム ー人間と人工知能との親密なコミュニケーションの実現ー

東京工芸大学 創立 100 周年記念展「写真から 100 年」の一環として、標記の展示を行った。開催期間は 2023 年 11 月 11 日(土)～12 月 10 日(日)、場所は東京都写真美術館。

展覧会は 4 部構成となっており、第 4 部は「次の 100 年に向かって」として、未来を創造する東京工芸大学の工学部と芸術学部の共同研究によるプロジェクトやメディアアート作品などが紹介された。その中で片上大輔(A03 班 公募班研究代表者・東京工芸大学)、宮本友樹(A03 班 公募班研究協力者・電気通信大学)らによって準備されたのが標記の展示となる。

展示された研究の内容は以下である。この研究では、AI が対話する相手の推定された年齢・性別・性格・表情・文化

などの属性と会話の反応に応じて、話のレベルや丁寧さを調節し、敬語・友達口調などを駆使して話す、いわば言葉の配慮を備えた対話システムを構築する。これまでに言語学の分野で研究されてきた「人が会話の相手と良好な関係性を構築・維持するための言語的配慮」と定義される「ポライトネス理論」と、その会話における相互作用による変化を扱う「デイスコース・ポライトネス理論」で蓄積されてきた知見に基づいて、人間と AI の間での適切で親和性の高い対話のやり取りを実現する。

展示には上皇ご夫妻も私的に訪問されたとのことである(FNN プライムオンライン:<https://www.fnn.jp/articles/gallery/621332>)。

■ 対話知能学プロジェクト×日本科学未来館 vol.17 トークイベント(告知)

2024 年 1 月 27 日(土)に、日本科学未来館にて、表記のイベントを開催予定である。これまでの対話知能学における取り組みの振り返り、今後に向けた展望などについてのトークイベントを予定している。

「対話知能学」領域からは、石黒浩領域代表、河原達也(A01 計画班代表研究者)、東中竜一郎(A02 計画班代表研究者)、杉山弘晃(A03 計画班代表研究者)、新保史生(A04 計画班代表研究者)がゲストとして登壇予定である。

3. 日本科学未来館での活動

■ 常設展における傾聴ロボットの展示 (A01 公募班 瀬島吉裕)

日本科学未来館の常設展示のリニューアルが11月22日に行われ、4つの新しい展示が加わった。3階「未来をつくる」ゾーンの「ハロー！ ロボット」はその一つである。ロボットたちとのふれあいや、最新ロボティクス研究の紹介を通して、未来の多様なロボットとのくらしを想像する展示である。そこに、「対話知能学」において瀬島吉裕(A01 公募班研究代表者)が開発したロボット、Pupiloid(ピューピロイド)が展示されている。

Pupiloid は、Pupil(瞳)+ oid(ロボット)の名前のとおり、人の瞳のような見た目をしたロボットである。このロボットは人の会話の「熱量」(発話量)を計測していて、より熱心に話しかけるほど、瞳孔が大きくなる。瞳孔の変化があることによって、よりロボットに親しみを感じやすくなり、会話を促進させる効果が期待されている。このロボットを用いた研究については『対話知能学ニュースレター vol. 15』の「1. 研究紹介」をご覧ください。これまで長期間展示されていたアンドロイド「オルタ」の展示に引き続いて、コミュニケーションをテーマとしたロボットが展示されたことは、本領域研究が極めて重要だと認識されていることを示している。



常設展示リニューアルの PR: <https://www.miraikan.jst.go.jp/news/press/202311213252.html>

Pupiloid(ピューピロイド)含む「ハロー！ ロボット」のロボット群の PR: <https://www.miraikan.jst.go.jp/news/general/202310313217.html>

■ 実証実験「アンドロイド I(アイ)と対話してみよう！」(A03 公募班 片上大輔)

2023年12月15日(金)～17日(日)の3日間、日本科学未来館においてアンドロイド I を用いた標記の実証実験を行った。実施者は片上大輔(A03 公募班研究代表者)、宮本友樹(A03 公募班研究協力者)、太田泰介(A03 公募班研究協力者)である。

対話アンドロイドを用いて、ロボットと人間がコミュニケーションをとりながら協働する実験を行った。「旅行」という共通の趣味を持っているという設定で、人間にそっくりなアンドロイド I(アイ)」と対話していただく。主に、実験参加者が今までに行った旅行先でよかった場所について、アンドロイド I と雑談しながらコミュニケーションを行った。

被験者は3日間で81名となり、目標の72人を上回った。実証実験の告知ページはこちら: <https://www.commu-ai.org/activity/20231215>

4. 顕著な業績

■ 論文

岡田将吾(A02 公募班研究代表者)らによる, ユーザの主観対話評価推定に基づいた対話戦略の適応の効果検証についての論文が Affective Computing 分野のトップジャーナルである IEEE Transactions on Affective Computing に採択された。

この研究において著者らは, 対話中のふるまい(音声・動作)から人の発話意欲(質問に意欲的に回答したい/回答したくない)を推定して, 適応的に質問するインタビューロボットを開発した。また発話意欲を推定・適応する対話ロボットが, 人の発話意欲をさらに引き出せることを定量的に検証した。

書誌情報: F. Nagasawa, S. Okada, T. Ishihara and K. Nitta, "Adaptive Interview Strategy Based on Interviewees Speaking Willingness Recognition for Interview Robots," in IEEE Transactions on Affective Computing, doi: 10.1109/TAFFC.2023.3309640.

■ 論文

船越孝太郎(A01 公募班研究代表者)らによる発表が, マルチモーダルインタラクションに関するトップ国際会議である ACM ICMI '23 (The 25th International Conference on Multimodal Interaction) に採択された。

この研究において発表者らは, 一般的なRGB動画像を用いて, 被写体(対話システムのユーザ)の呼吸波形を従来よりも高精度に推定できることを示すと共に, 推定した波形からユーザの発話開始を予測できることを示した。

書誌情報: Takao Obi, Kotaro Funakoshi. Video-based Respiratory Waveform Estimation in Dialogue: A Novel Task and Dataset for Human-Machine Interaction. Proc. the 25th International Conference on Multimodal Interaction, pp.649-660, 2023.

■ 受賞

船越孝太郎(A01 公募班研究代表者)らによる発表が, 2023年12月13, 14日に開催された「第14回対話システムシンポジウム」にて, 若手優秀賞を受賞した。

本発表では, 同著者らによる ICMI'23 での発表(一つ前の項目「論文」参照)の内容に加えて, 進行中の呼吸情報を利用した人・ロボット対話実験とその途中結果が報告された。

書誌情報: 小尾 賢生, 船越 孝太郎. 「親和的な対話システムの実現にむけた非接触呼吸推定技術の開発」. 第99回人工知能学会 言語・音声理解と対話処理研究会(第14回対話システムシンポジウム), pp.7-12, 2023/12.