

戦略的創造研究推進事業 CREST
研究領域「人間と情報環境の共生インタラクション
基盤技術の創出と展開」
研究課題「文脈と解釈の同時推定に基づく相互理
解コンピューテーションの実現」

研究終了報告書

研究期間 2019年10月～2025年03月

研究代表者: 今井 倫太
(慶應義塾大学 理工学部 教授)

§1 研究実施の概要

(1)実施概要

大規模言語モデルを利用した SCAIN アルゴリズム(文脈と解釈の同時推定)の構造を考案した。相互理解を扱うために、インタラクションの流れにおける解釈者個人が把握する情報抽出を再現する構造となっている。具体的には、解釈者が注目する(文脈把握の視点となる)発言コアテキストを、解釈者が認識する(誤認識も含む)認知文脈をベースに、大規模言語モデルなどの解釈駆動モデルで情報文脈化し、個人の把握情報を抽出する。同じ文脈であっても、コアテキストの選び方、認知文脈の違いに従って、個々人が把握している情報の違いが生じるモデルである。

言語インタラクション以外のインタラクション階層へも SCAIN アルゴリズムの構造は展開できる。具体的には身体・物理世界インタラクション、非言語インタラクション、言語インタラクション、個体間インタラクションであり、それぞれの階層で研究を行なった。

・高橋 G・杉浦 G は、身体・物理世界インタラクションを軸足に研究した。高橋 G では、物理的行動制約の下で行動するロボットの行動可能性を生成するアクショングラフ、人の命令の解釈を表現するシーングラフを用意し、人からの命令を人が納得する(相互理解の)形で実行することを実現した。また、同手法を長崎リハビリテーション病院に展開し、社会実装に取り組んだ。杉浦 G は、ロボットが、実空間の物体の場所を移動しながらスキャンし、人間からの自然言語文から物体を特定して、物体の把持・移動させることのできるマルチモーダルアーキテクチャを開発した。

・今井 G は、オンラインコミュニケーションの参加者の聞き逃し場面を題材に、言語インタラクションの研究を行った。聞き逃した人の認知会話履歴をシミュレーションし、聞き逃しが生じた認知会話履歴と生じなかった場合を比較することで、聞き逃すと意味が変わってしまう文章 SCAINsを特定するアルゴリズムを構築し、聞き逃しにおける相互理解支援システムを構築した。

・今井 G・植田 G は、個体間インタラクションにおける文脈と相互理解を扱った。植田 G が明らかにした、インタラクションを通して人が行う他者の選好推定の数理モデル(スパースモデリング)を基盤とし、人とのインタラクションで取得可能な少数のサンプルでユーザのファッションの選好推定・推薦をするシステム Interactive-SmartClerk を構築した。Interactive-SmartClerk は、杉浦 G による自然言語文によるファッションアイテム検索ならびに、高橋 G 構築のアクショングラフとも連携しており、本チーム全成果を統合したものとなっている。

以上の研究で見えてきた文脈と解釈の同時推定と相互理解を扱う上での重要な点は、解釈者毎に生じる認知文脈の違いを正しくモデル化することである。物理世界でのインタラクションでは人が環境情報のどこに注目するかで行動命令の解釈が異なる。言語インタラクションでは、聞き逃しなどの人が把握する会話の履歴の違いが相互理解に向けた発言解釈に大きく影響を与える。個体間インタラクションでは、価値判断サンプルの選択が重要であり、エキスパート店員のようなサンプルの選び方が良い選好モデル構築に重要である。インタラクションにおける認知文脈をシミュレーションし、各自が把握する文脈情報の個別性の違いを明示的に扱うことが、コンピュータで相互理解を扱う上で必要不可欠である。また解釈駆動モデル(行動規則・物理制約、大規模言語モデル、集団価値モデル)は、発言の解釈を文脈に位置付ける上で重要な役割を果たしている。インタラクションで交わされる情報を人がどう位置付けし、文脈情報を取り出すかは、認知文脈のシミュレーションと解釈駆動モデルの構成として実現可能であることが明らかになった。

(2)顕著な成果

< 優れた基礎研究としての成果 >

1.

概要: 人同士の会話における相互理解を支援するために、相互理解の妨げとなる対話者間の発言解釈の不一致の可能性がある箇所 SCAINs を推定する手法を開発し、難関国際会議 EMNLP'23 で成果発表した。相手の発言の聞き逃しや、発言を理解する際の知識不足が、その後の発言解釈に影響する箇所を SCAINs として抽出する手法である。情報文脈化手法と疑似的欠損文脈の二点により、会話参加者の文脈把握をシミュレーションすることを可能にした。

2.

概要: 人がインタラクショを通じて行う相手の選好推定メカニズムを、異性の顔画像を題材に明らかにした。人はごく少数の質問でインタラクショ相手の選好の推定が可能であり、大量データが必要なベイズ推定には劣るものの、その精度はかなり高かった。また、人の推定メカニズムがスパースモデリングで説明できることも明らかとなった。インタラクショの文脈を利用し個人適応するシステム構築に重要な知見が得られた。

3.

概要: 代理投資課題における他者のリスク推定と行動調整の実験を行った。インタラクショの中で、相手の価値(ここではリスク選好)を人が見積もる方法を調べる実験は世界初の試みである。実験の結果、投資者が、出資者の指示に従って投資回数を調整できるものの、出資者のリスク選好の推定はできなかった。一方、出資者の満足は十分に高かった。満足させるのが目的であれば、意図推定や行動調整までは必要ないことが明らかになった。

< 科学技術イノベーションに大きく寄与する成果 >

1.

概要: オンライン会話における聞き逃しを支援し場面を支援する **SCAInS Presenter** を構築した。人が聞き逃すと発言の意味が変わってしまう箇所 **SCAInS** を発見し、会話参加者に提示するシステムである。聞き逃した人や、複数のオンライン会議に同時に参加する必要がある人が、会話文脈に迫りつくことを支援できる新産業分野の創出につながる成果である。知的インタフェースの難関国際会議 **IUI'24** で成果発表した。

2.

概要: 進行中の文脈に基づき任意のテキストを情報文脈化し、会話に関連する情報提供サービスの基本的枠組みを考案した。研究では、デジタルサイネージの前で会話している人の文脈に合わせて、任意の広告文を情報文脈化し、ユーザに訴求する広告を提示するシステムを構築した。知的インタフェースの難関国際会議 **IUI'24** ならびに知的インタフェースのに関する学術論文 **ACM TiiS** で発表した。

3.

概要: シーングラフに基づいて日本語の画像キャプションを自動評価する指標 **JaSPICE** を提案した。従属節と述語-目的語構造からシーングラフを生成し、同義語を用いてさらにグラフを拡張した。**STAIR Captions** と **PFN-PIC** で学習した 10 種類の画像キャプション生成モデルで実験を行った。その結果、提案手法は人間評価との相関係数において、**BLEU** や **METEOR** などのベースライン手法を上回る結果となった。パターン認識の難関国際会議 **CoNLL2023** で成果発表を行った。

< 代表的な論文 >

1.

概要: コンテキストライディングに関する研究として、ある個人が他人に物を推薦する際に、自分の推薦度合いに加えて、世の中の一般的な評判も加味させることで、より相手の好みに近い推薦が可能になることを明らかにした(一人集合知に基づく選好のキャリブレーション)。本研究は、インタラクショによる相手とのすり合わせで、個々人の推薦が形成される過程の解明に繋がる大変インパクトの大きい知見であり、**Scientific Reports(2022)**に掲載された。

2.

概要: 画像キャプションモデルの教師あり自動評価指標である **Polos** を提案した。**Polos** は、大規模な対照学習によってトレーニングされた埋め込みを活用する並列特徴抽出メカニズム

を使用して、マルチモーダル入力からスコアを計算する。Polos をトレーニングするために、人間のフィードバックに基づいて Multimodal Metric Learning from Human Feedback (M2LHF) で指標を開発した。結果、Composite、Flickr8K-Expert、Flickr8KCF、PASCAL-50S、FOIL、Polaris データセットで最先端のパフォーマンスを達成し、その有効性と堅牢性を実証した。

3.

概要: マルチターン参照対象同定問題に関する研究として、人の選択物体をマルチターンで特定する質問生成と物体特定をする深層学習アーキテクチャ UniQer を構築した。類似物体が複数ある画像においても対話の履歴と画像情報を関連付ける手法と、物体特定に繋がる質問生成の手法を実現しており、シングルターンが主流の視覚言語課題に大きなインパクトを与えられた。CLEVR ASK データセットにおいて SOTA を達成し、IEEE ICCV2021(the IEEE/CVF International Conference on Computer Vision 2021)に掲載された。

§2 研究実施体制

(1)研究チームの体制について

- ①「今井」グループ 研究代表者: 今井 倫太(慶應義塾大学理工学部 教授)

研究項目

- ・SCAIN の音声対話システム化
- ・SCAIN の特性解析・評価
- ・低ストレス音声対話システムの構築
- ・低ストレス・マルチモーダル対話システムの構築
- ・相互理解度推定の実現

- ②「植田」グループ 主たる共同研究者: 植田 一博(東京大学大学院総合文化研究科 教授)

研究項目

- ・人の関係性と言語表現の運用の研究
- ・文脈による価値推定に根ざした相互理解基盤モデルの実現

- ③「高橋」グループ 主たる共同研究者: 高橋 正樹(慶應義塾大学理工学部 教授)

研究項目

- ・環境情報の分散表現変換
- ・低ストレス・マルチモーダル対話システムの構築
- ・文脈依存コントロールの実現

- ④「杉浦」グループ 主たる共同研究者: 杉浦 孔明(慶應義塾大学理工学部 教授)

研究項目

- ・SCAIN の音声対話システム化
- ・環境情報の分散表現変換
- ・低ストレス・マルチモーダル対話システムの構築
- ・文脈依存コントロールの実現

(2)国内外の研究者や産業界等との連携によるネットワーク形成の状況について

産業界: 次の企業との共同研究を行なっている。

NTT コミュニケーション科学研究所・ATR インタラクション科学研究所、トヨタ自動車株式会社、アイシン株式会社、パラマウントベッド株式会社、村田機械株式会社、サントル株式会社、本田技術研究所、Apple、伊藤忠テクノソリューションズ、KDDI、ソニーリサーチ、ソフトバンク、東京海上ホールディングス、TOPPAN ホールディングス、トヨタシステムズ、NEC、三菱 UFJ フィナンシャル・

グループ、吉本興業株式会社、ソニーグループ株式会社、トヨタ自動車株式会社、株式会社 JTB
国内：Human-Agent Interaction 研究において共同研究、研究会開催を行った。

国外：慶應義塾大学とカーネギメロン大学の人工知能分野における共同研究が開始された。また、
国際会議 IEEE/ACM HRI, HAI, IEEE RO-MAN における会議運営やワークショップ開催の連携を
行った。さらに、南カリフォルニア大学にてシンポジウムを開催しチームの成果発表を行った。