

# 各対話破綻類型の特徴を考慮した対話破綻検出システムの提案

初田 玲音

2021 年 7 月 16 日 中間発表

## 1 研究背景

近年、雑談対話システムの研究が盛んである。雑談対話システムは、単に娯楽としての目的だけではなく、コンピュータ・ロボットの社会性地位獲得の為に技術としても研究が進められている。上手く書けません！なんか違う気がします！ 雑談対話システムが人間のパートナーとして認められる為には、一般的に継続的に自然な発話をする必要がある。

しかし、雑談対話システムは一般的にオープンドメインであり、扱う対話内容が多様かつ複雑である。そのため、システムがユーザの発話を正しく理解出来ず、対話破綻を引き起こす発話を行ってしまうことがある。破綻した発話は、継続的な対話の妨げになるため、システムが事前に対話破綻を引き起こす発話を検出することが有効だと考えられる。

そこで、本研究では、マルチラベル分類問題へと発展させた対話破綻エラー類型検出システムについて、各破綻エラー類型の特性を考慮した特徴を用意することで、個々の破綻エラー類型の推定精度を改善し、それによって全体の破綻エラー類型の推定精度を向上させる手法を提案する。

## 2 関連研究

堀井ら [1] は、Project Next NLP の日本語対話タスクグループによる雑談対話の破綻原因類型化案に基づき、その類型毎に破綻識別器を作成し、それらを組み合わせる手法を提案した。破綻の原因を類型化した点は同様だが、出力については「破綻ではない (o)、破綻 (x)、違和感あり (t)」の 3 種のみに留まり、どの破綻類型で破綻するかは考慮していない。

また、Xu らは、談話特性における大局的な一貫性は、局所的な一貫性に分解出来るという仮定に基づき、談話特性における一貫性を識別する教師なし学習モデルとして、local coherence discriminator(LCD) を提案した。このモデルでは、与えられた連続する 2 つの文章対を正例、文章の順序を無作為に入れ替えた不自然な文章対を

負例とする。実際の学習は、学習対象の文書中の正例を、作成された負例から識別するように行った。

## 3 研究構想

対話破綻類型は、全部で 17 種類存在し、2 つの軸によって整理され、8 つのグループに分けられる。従来の対話破綻検出と異なり、各対話破綻類型についてマルチラベルに分類する必要があるため、それぞれの対話破綻類型に対して、個別の破綻検出アプローチを適用することを想定している。

破綻特定原因対象範囲が「発話」(「解釈不能」や「用法エラー」など) ならば該当のシステム発話のみを基本に特徴を抽出し、「応答」ならば該当のシステム発話に加え、直前のシステム発話を基本として特徴を抽出する。例えば、「質問無視」については、チュータリングシステムや対話行為推定の研究を、「話題遷移エラー」については、～を参考にすることを検討している。

また、「常識欠如」や「社会性欠如」などの、学習データ中に破綻が著しく少ない(もしくは存在しない)破綻類型については、そのままのデータによる学習が困難である。現時点では、破綻特定原因対象範囲が「応答」である「挨拶無視」ならば、新たな学習データは不必要、または少量で対応出来ると考えているが、「常識欠如」や「社会性欠如」といった、社会的な要素を必要とする破綻類型の学習をどのように行うかについては、現在調査、検討中である。

## 4 今後の課題

- 各破綻類型の特徴、特性を調べる
- 社会性、常識の取り扱いについて検討する
- どのような特徴が有用か決定する方法を検討する
- 実験方法、評価方法について検討する