

対話破綻修復コーパスの収集と分析

—ユーザの個人特性とシステムとの関係性を考慮した修復文生成に向けて—

坪倉和哉¹ 岡田真依² 入部百合絵² 北岡教英³

¹ 愛知県立大学大学院 情報科学研究科 ² 愛知県立大学 情報科学部

³ 豊橋技術科学大学 情報・知能工学系

{id231001,is201022}@cis.aichi-pu.ac.jp

iribe@ist.aichi-pu.ac.jp kitaoka@tut.jp

概要

現状の対話システムとの対話では、未だ対話破綻が生じているため、破綻発生時には適切に対話を修復する必要がある。しかし、これまで提案されている修復方法はどのユーザに対しても同様の発話を行う。ユーザ適応の観点から、ユーザの個人特性やシステムとの関係性によって適切な修復方法が異なるかを明らかにするため、個人特性や関係性を含む対話破綻修復コーパスを作成した。破綻後のユーザ発話を分析した結果、自己主張スキルが高い人は破綻したシステムに同意しない、初対面や上司に対しては破綻後に依頼をしない、などの傾向が確認された。このことから、これらの傾向に合わせた破綻修復方法を検討する必要があることが示された。

1 はじめに

近年、大規模言語モデルの登場により対話システムの応答精度が高まっている [1, 2]。しかし、システムが事実に基づかない発話やペルソナと矛盾した発話を生成してしまう問題は未だ起こり得るため、関連研究が進められている [3, 4]。このような対話破綻を生じる発話は、ユーザの対話意欲を損ねる恐れがあるため、対話破綻が発生した後は適切に対話の流れを修復する必要がある。

しかし、対話破綻後にシステムが対話の流れを修復する研究は少ない [5]。対話破綻を修復する戦略としては、ユーザに謝罪する、ユーザに文を言い換えるように求める、などがある。現状のシステムは自己の破綻を認識したとしても、全てのユーザに対して同様の修復戦略を選択してしまうが、ユーザの個人特性やシステムとの関係性によって適切な修復文が異なる可能性がある。適切な修復文を生成するには、対話破綻後のユーザ発話の傾向を分析し、分

析結果に基づきユーザの個人特性やシステムとの関係性を考慮した修復戦略を選択する必要がある。

そこで本研究では、その第一歩として、ユーザの個人特性やシステムとの関係性が対話破綻後のユーザの発話の対話行為に影響を及ぼすかを調査する。そのために、まず個人特性と関係性のラベルが付与された対話破綻修復コーパスを作成した (2章)。次に、収集したコーパスを用いて破綻後のユーザ発話の傾向を分析し (3章)、破綻後のユーザ発話とシステムの修復発話の関連を確認した (4章)。

2 対話破綻修復コーパスの収集

本章では、本研究で収集した対話破綻修復コーパスについて述べる。ユーザの個人特性やシステムとの関係性が対話破綻後のユーザの発話の対話行為に影響を及ぼすかを調査するために、個人特性と関係性のラベルが付与された対話破綻修復コーパスを作成した。

2.1 対話破綻例の作成

本研究では、破綻修復コーパスを作成するため、破綻が発生した対話例をクラウドワーカーに提示し、その後続く発話を回答してもらった。そのため、まず対話破綻の例の作成を行った。

先行研究において対話破綻を生じさせる発話は17種類に分類されているため [6]、この分類を参考に対話破綻例を作成した。なお、大規模言語モデルの進展により、17種類の破綻のうちいくつかの破綻は発生頻度が極端に低いと考えられるため、本研究の対象からそれらを除き、計10種類とした¹⁾。また関係性は、コンボイモデル [7] を参考に7パターン

1) 誤情報、質問無視、期待無視、発話意図不明確、話題遷移エラー、情報不足、自己矛盾、相手の発話との矛盾、繰り返し、常識欠如の10種類。

質問6 (初対面)

ユーザA (あなた) : この前風邪を引いたんです。

ユーザB : それはよかったですね。

ここにユーザA (あなた) の応答文を入力してください

ここにユーザBの応答文を入力してください

図1 クラウドワーカーに提示した破綻発話の例。ユーザB (システム役) の最終発話は破綻を引き起こす発話である。

用意した (表1)。

2.2 破綻修復発話の収集手順

前節で作成した破綻例 140 個に対して、それぞれクラウドワーカーから回答を得た。クラウドワーカーに提示した破綻例を図1に示す。図1のように、クラウドワーカーには、ユーザAとユーザB間の関係性とユーザBの最終発話が破綻を引き起こしている対話履歴を提示する。クラウドワーカーは、破綻したユーザBに対する発話と、その発話に対するユーザBの破綻修復発話を作成する。なお、破綻例は[6]を参考に著者が作成した。

本研究では、作成した 140 個の破綻例を 28 個ずつの 5 つの組に分割し、各組 20 名のクラウドワーカーから回答を得た。そのため、合計 100 名のクラウドワーカーから 2800 件の対話破綻修復事例を収集した。

2.3 個人特性アンケート

クラウドワーカーの個人特性を調査するため、性格特性とコミュニケーションスキルに関するアンケートも実施した。

2.3.1 性格特性

クラウドワーカーの性格特性を測るため、BigFive[8]の5つの特性を10項目で測定する日本語版 Ten Item Personality Inventory (TIPI-J) [9]の質問文を利用した。BigFiveは、パーソナリティの全体的構造を外向性 (Extroversion)、協調性 (Agreeableness)、勤勉性 (Conscientiousness)、神経症傾向 (Neuroticism)、開放性 (Openness) の5つの次元で捉えるモデルである。このアンケートにより、BigFiveの各特性に対して2から14点の得点が得られる。

2.3.2 コミュニケーションスキル

クラウドワーカーのコミュニケーションスキルを測るため、ENDCORE (簡易版) [10]を採用した。ENDCOREは、コミュニケーションスキルの因子を階層構造として統合したモデルであり、自己統制、表現力、解読力、自己主張、他者受容、関係調整の6つの種類に分類されている²⁾。このアンケートでは、ENDCOREの各カテゴリーに対して7件法で回答を得る。

3 対話破綻後のユーザ対話行為の分析

3.1 対話行為の付与

本研究では、システムの対話破綻発話の後のユーザ発話とシステムの修復発話について分析を行う。分析方法として、本研究では対話行為レベルに着目し、ユーザの個人特性やシステムとの関係性によってどのように対話行為が異なるかを分析する。

そのために、まず日本語対話行為タガー [11]³⁾を使用し、クラウドワーカーが回答した破綻後のユーザ発話とシステムの修復発話に対話行為タグを付与する。日本語対話行為タガー [11]では、ISO 24617-2で定義された44種類の対話行為タグを付与することができる。対話行為タグを付与した結果の内、頻出度合が上位10個のタグを表2に示す。

3.2 性格特性と対話行為の関連

クラウドワーカーの性格特性と破綻後のユーザ発話に関係があるかを確認するため、クラウドワーカーの性格特性の得点 (各性格特性2から14点) と28個の破綻後のユーザ発話の対話行為の割合の相関係数を求めた。なお、対話行為は表2に示す頻出度合が上位10個のタグのみ使用した。また、100名のクラウドワーカーのうち、1名は性格特性のアンケートに不備があったため、99名のデータを使用した。Pearsonの積率相関係数を求めた結果、有意に相関のあった性格特性と対話行為の組を表3に示す。

2) ENDCORE (簡易版) の項目文を下記に記す。

自己統制: 自分の感情や行動をうまくコントロールする

表現力: 自分の考えや気持ちをうまく表現する

解読力: 相手の伝えたい考えや気持ちを正しく読み取る

自己主張: 自分の意見や立場を相手に受け入れてもらえるように主張する

他者受容: 相手を尊重して相手の意見や立場を理解する

関係調整: 周囲の人間関係にはたらきかけ良好な状態に調整する

3) https://github.com/Inever/commu_ai_tagger

表1 関係性の種類と定義

関係性	定義
初対面	今まで1度も会ったことがない人
顔見知り	社会的役割関係に直接結びつき役割変化の影響を最も受ける（顔と名前は知っているが話したことはない人）
知り合い	やや役割に関連し、時間の経過によって変化しやすい（数回話したことがある人）
友人	やや役割に関連し、時間の経過によって変化しやすく自己開示ができる 例：休日に出かける子
親友	長期的に安定的で役割に依存しない
会社の上司	同じ会社の上司
会社の部下	同じ会社の部下

表2 破綻後のユーザ発話の対話行為の割合（頻出上位10個のみ表示。割合は頻度を全発話数（2,800）で割った結果）

対話行為タグ	定義	割合
Answer	相手の質問や要求する情報に対する回答	25.6%
SetQuestion	5W1H 質問	13.7%
Inform	情報提供	13.1%
CheckQuestion	確認のための質問	7.8%
PropositionalQuestion	はい・いいえで答えられる質問	7.5%
Agreement	同意	5.9%
AddressOffer	相手の提案を考慮していることを示す発話。 提案に対する応答であるが、受諾でも拒否でも無い発話。	4.1%
Request	依頼。聞き手の同意を条件に指定した行動を行うことを約束させる発話	3.8%
Disconfirm	相手の質問や要求する情報に対する否定的な回答。No-Answer	3.4%
AcceptOffer	提案の受諾	3.1%

表3 性格特性と破綻後のユーザ発話の対話行為タグ割合との相関（Pearsonの積率相関係数を求めた結果、有意に相関があった性格特性と対話行為の組のみ表示）。 r は相関係数、sig.は有意水準のマークを表している。

性格特性	対話行為タグ	r	p 値	sig.
外向性	Answer	0.210	0.037	*
勤勉性	Disconfirm	0.258	0.010	*
勤勉性	AcceptOffer	-0.213	0.034	*
開放性	Answer	0.249	0.013	*

*: $p < 0.05$

表3より、外向性や開放性が高い人は、システムの破綻した質問や要求に対してきちんと回答をする傾向にあることがわかる。積極的な性格である外向性の高い人や思慮深い開放性が高い人は、システムが破綻しても、それを受け入れて返答する可能性がある。また、勤勉性が高い人はシステムの破綻した質問や要求に対して否定的に回答し、破綻した提案には受諾しない傾向にあることがわかる。勤勉性が高い人は責任感があり精力的に取り組むタイプであるため、システムが破綻した際に適当に受け入れることはせず、否定的に回答をする可能性がある。

3.3 コミュニケーションスキルと対話行為の関連

クラウドワーカーのコミュニケーションスキルと破綻後のユーザ発話に関係があるかを確認するため、クラウドワーカーのENDCOREの得点（各項目1から7）と28個の破綻後のユーザ発話の対話行為

表4 コミュニケーションスキルと破綻後のユーザ発話の対話行為タグ割合との相関（Pearsonの積率相関係数を求めた結果、有意に相関があったENDCOREと対話行為の組のみ表示）。 r は相関係数、sig.は有意水準のマークを表している。

ENDCORE	対話行為タグ	r	p 値	sig.
関係調整	Answer	0.198	0.049	*
自己主張	Agreement	-0.239	0.017	*

*: $p < 0.05$

の割合の相関係数を求めた。なお、対話行為は表2に示す頻出度合が上位10個のタグのみ使用した。Pearsonの積率相関係数を求めた結果、有意に相関のあったコミュニケーションスキルと対話行為の組を表4に示す。

表4より、関係調整スキルがある人は、システムの破綻した質問や要求に対してきちんと回答する傾向にあることがわかる。関係調整スキルがある人は、周囲と良好な関係に調整する能力があるため、システムが破綻発話をしてでも否定したり指摘したりすることはせず、システムの破綻した発言内容に合わせて対話継続を試みる可能性がある。また、自己主張の項目が高い人は自分の意見を受け入れてもらえるように主張する傾向にあるため、システムが破綻し想定外の発話をした場合にはシステムには同意しない傾向にある。

表 5 関係性毎に算出したシステム破綻後のユーザ発話の対話行為の割合（職場の上司/部下はシステム側の属性）

破綻後のユーザ発話の対話行為タグ	初対面	顔見知り	知り合い	友人	親友	職場の上司	職場の部下
Answer	26.75%	27.05%	27.11%	32.50%	22.63%	24.50%	18.25%
SetQuestion	17.50%	10.00%	12.63%	12.75%	12.37%	13.25%	17.75%
Inform	11.25%	17.73%	9.21%	14.50%	11.32%	11.00%	16.25%
CheckQuestion	11.50%	6.82%	6.32%	8.50%	4.74%	9.00%	7.25%
PropositionalQuestion	9.75%	9.09%	8.68%	4.00%	6.84%	5.75%	8.50%
Agreement	4.25%	7.95%	6.58%	6.00%	6.05%	7.25%	3.00%
AddressOffer	1.75%	4.55%	6.05%	5.00%	6.05%	1.25%	4.25%
Request	0.50%	2.95%	7.37%	1.75%	8.42%	1.00%	4.75%
Disconfirm	2.50%	2.95%	4.21%	3.75%	5.00%	2.25%	3.00%
AcceptOffer	1.00%	2.50%	2.11%	3.75%	4.74%	3.50%	4.00%

3.4 関係性と対話行為の関連

本研究で用意した7つの関係性とシステム破綻後のユーザ発話に関係があるかを確認するため、関係性毎に対話行為の割合を求めた（表5）。表5より、関係性毎に各対話行為の割合が異なっていることが確認できる。例えば、否定的な回答であるDisconfirmは親友より初対面の方が少なく、部下よりも上司の方が少ないことがわかる。また、依頼を意味するRequestは初対面や上司の場合に少ない一方で、親友は7つの関係性の中で最も多いことが分かる。このように、ユーザとシステム間の社会的な距離が破綻後の対話行為に影響を及ぼすことが示されたため、これらの傾向が適切な破綻修復の選択・戦略に役立つ可能性がある。

4 破綻後のユーザ発話とシステムの修復発話の分析

前章では破綻後のユーザ発話のみを分析したが、本研究で収集したコーパスにはシステムの修復発話も含まれるため、破綻後のユーザ発話とシステムの修復発話の組合せに関しても分析を行った。なお、システムの修復発話に関しても3.1節と同様の方法で対話行為を付与した。破綻後のユーザ発話とシステムの修復発話の出現割合を求めた結果を表6に示す。

全体的な傾向としてAnswer-AnswerとSetQuestions-Answerの組が約18%を占めていることがわかる。上位3位以下は2%以下となっているため、割合が多い組合せは上位2つのみであると言える。このことから、対話破綻後のユーザの反応、およびその後のシステムに望む修復発話は多様であることが明らかになった。前章の議論からも、ユーザの個人特性、関係性に依拠してこのような多様性が生じている可能性が考えられる。今後はデータを拡充するとともに、破綻後のユーザ発話とシステ

表 6 破綻後のユーザ発話 × システムの修復発話の出現割合（上位10個のみ表示）

破綻後のユーザ発話の対話行為タグ	システムの修復発話の対話行為タグ	割合
Answer	Answer	10.39%
SetQuestion	Answer	7.39%
PropositionalQuestion	Answer	1.93%
Answer	Inform	1.89%
Answer	CheckQuestion	1.82%
Agreement	Agreement	1.71%
Inform	Inform	1.68%
PropositionalQuestion	Inform	1.32%
Answer	Agreement	1.25%
Answer	Disconfirm	1.25%

ムの修復発話の組合せと個人特性や関係性との関係をより明確にしていく必要がある。

5 おわりに

現状の対話システムでは未だ対話破綻が発生するため、破綻時には適切な修復を行う必要がある。しかし、対話修復に関する先行研究では誰に対しても同様の修復処理が行われてきた。ユーザの個人特性やシステムとの関係性によって適切な修復方法が異なる可能性があるため、本研究では、まず個人特性や関係性ラベルを含む対話破綻修復コーパスを構築した。次に、コーパスの分析の第一段階として、対話破綻後のユーザ発話の対話行為に焦点を当て分析を行った。個人特性と対話行為、関係性と対話行為との関係をそれぞれ分析した結果、一部の項目間で傾向が確認されたため、個人特性や関係性が破綻後のユーザ発話に影響を及ぼす可能性が高い。

今後の課題としては、データを拡充するとともに、システムの破綻修復文の対話行為についても詳細に分析を行い、個人特性や関係性を考慮した修復文生成手法を提案していく。

謝辞

本研究は JSPS 科研費 JP22K19793, JP23H00493 の助成を受けたものです。

本研究は愛知県が公益財団法人科学技術交流財団に委託し実施している「知の拠点あいち重点研究プロジェクト第 IV 期（第 4 次産業革命をもたらすデジタル・トランスメーション（DX）の加速）」の研究成果です。

参考文献

- [1] Kurt Shuster, Jing Xu, Mojtaba Komeili, Da Ju, Eric Michael Smith, Stephen Roller, Megan Ung, Moya Chen, Kushal Arora, Joshua Lane, Morteza Behrooz, William Ngan, Spencer Poff, Naman Goyal, Arthur Szlam, Y-Lan Boureau, Melanie Kambadur, and Jason Weston. Blenderbot 3: a deployed conversational agent that continually learns to responsibly engage, 2022.
- [2] Takato Yamazaki, Katsumasa Yoshikawa, Toshiki Kawamoto, Tomoya Mizumoto, Masaya Ohagi, and Toshihiko Sato. Building a hospitable and reliable dialogue system for android robots: a scenario-based approach with large language models. *Advanced Robotics*, Vol. 37, No. 21, pp. 1364–1381, 2023.
- [3] Souvik Das, Sougata Saha, and Rohini Srihari. Diving deep into modes of fact hallucinations in dialogue systems. In Yoav Goldberg, Zornitsa Kozareva, and Yue Zhang, editors, *Findings of the Association for Computational Linguistics: EMNLP 2022*, pp. 684–699, Abu Dhabi, United Arab Emirates, December 2022. Association for Computational Linguistics.
- [4] Zhenfeng Han, Sai Zhang, and Xiaowang Zhang. Persona consistent dialogue generation via contrastive learning. In *Companion Proceedings of the ACM Web Conference 2023, WWW '23 Companion*, p. 196–199, New York, NY, USA, 2023. Association for Computing Machinery.
- [5] Dennis Benner, Edona Elshan, Sofia Schöbel, and Andreas Janson. What do you mean? a review on recovery strategies to overcome conversational breakdowns of conversational agents. In *International Conference on Information Systems (ICIS)*, pp. 1–17, 2021.
- [6] 東中竜一郎, 荒木雅弘, 塚原裕史, 水上雅博. 雑談対話システムにおける対話破綻を生じさせる発話の類型化. *自然言語処理*, Vol. 29, No. 2, pp. 443–466, 2022.
- [7] Robert Louis Kahn and Toni C Antonucci. Convoys over the life course: attachment, roles, and social support. 1980.
- [8] Lewis R Goldberg. An alternative" description of personality": the big-five factor structure. *Journal of personality and social psychology*, Vol. 59, No. 6, p. 1216, 1990.
- [9] 小塩真司, 阿部晋吾, Pino Cutrone. 日本語版 ten item personality inventory (tipi-j) 作成の試み. *パーソナリティ研究*, Vol. 21, No. 1, pp. 40–52, 2012.
- [10] 藤本学, 大坊郁夫. コミュニケーション・スキルに関する諸因子の階層構造への統合の試み. *パーソナリティ研究*, Vol. 15, No. 3, pp. 347–361, 2007.
- [11] Michimasa Inaba, Yuya Chiba, Ryuichiro Higashinaka, Kazunori Komatani, Yusuke Miyao, and Takayuki Nagai. Collection and analysis of travel agency task dialogues with age-diverse speakers. In *Proceedings of the Thirteenth Language Resources and Evaluation Conference*, pp. 5759–5767, 2022.