

対話介入実験による対話の引き継ぎに有用な情報の調査

山下紗苗 東中竜一郎

名古屋大学大学院情報学研究科

xz.51d.7408@s.thers.ac.jp higashinaka@i.nagoya-u.ac.jp

概要

現状のチャットボットやコールセンタの自動応答では、完全に自律したやり取りは難しく、必要に応じてオペレータが介入する必要がある。しかし、どのような情報をオペレータに提示すればスムーズに引き継げるのかは明らかでない。本研究では、2人のオペレータがメモを引き継ぎながら一定時間ごとに交代してユーザと対話する実験を行った。そして、交代時に用いたメモを分析することで、対話を引き継ぐ際に必要な情報を実験的に調査した。雑談、相談、接客の対話タスクを対象に、240回のオペレータ切り替えを含む60対話を収集した結果、対話を引き継ぐ際には、対話の情報を隣接ペアや属性値対の形式で提示するのがよいという知見が得られた。

1 はじめに

近年、ニューラルネットワークの適用に基づく高精度な自律対話システムが多く研究されている [1, 2, 3, 4, 5]。しかし、しばしば矛盾 [6] や誤情報 [7] を含む発話が生成され、対話破綻を引き起こす。そのような場合、人間が適宜対話に介入し、対話システムから対話を引き継いで応対を進めることが有用だと考えられる [8]。しかしながら、現状、対話を引き継ぐ際にどのような情報があれば対話をスムーズに引き継げるかは明らかでない。

本研究では、対話を引き継いだオペレータが対話状況を迅速に把握し、スムーズに対話を続けられるようにするために、どのような情報を提示すればよいかを明らかにする。具体的には、2人のオペレータがメモを引き継ぎながら一定時間ごとに交代してユーザと対話する実験 (図1参照) を行い、引き継ぎに使ったメモを分析することで、対話を引き継ぐ際に必要な情報を実験的に明らかにする。雑談、相談、接客の対話タスクで実験を行い、複数の対話種別の引き継ぎに共通して必要な情報を調査する。

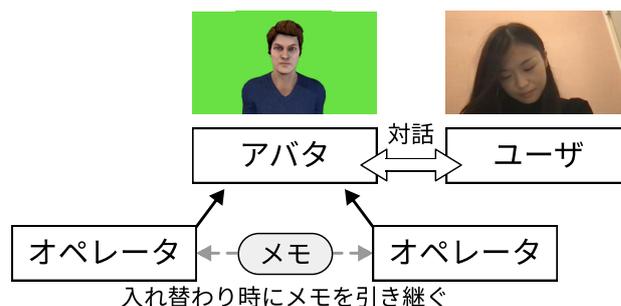


図1 対話介入実験の模式図。2人のオペレータが、メモをもとに、一定時間ごとに入れ替わりながらアバタを介してユーザと対話する。

2 関連研究

人間による対話の引き継ぎは、コールセンタでの入電を適切なオペレータに転送するコールルーティングの文脈で行われてきた [9, 10]。しかし、これらの研究では引き継ぎ時に提示すべき情報は扱われていない。

近年では、河原ら [8] が傾聴対話をオペレータに引き継ぐ研究を行っている。河原らは、オペレータに提示する情報として、音声認識結果とキーワードの強調表示を用いている。山下ら [11] は、雑談において、オペレータが対話を引き継ぐ際に、対話履歴の最後の5発話を提示することが有用であると報告している。しかし、これらの情報が対話の引き継ぎに最適かどうかは明らかではない。本研究では、引き継ぎに必要な情報に想定を置かず、オペレータが作成する引き継ぎ用のメモに着目し、必要な情報を実験的に調査する。

本研究は、対話内容を即座に把握するための対話の構造化に関するものと言える。対話の構造化には、トピックセグメント [12, 13]、談話セグメント [14]、情報状態 [15]、隣接ペア [16, 17, 18] などが利用されてきた。本研究では、これらの構造化の方法を念頭に置きつつ、オペレータが作成したメモをもとに、対話の引き継ぎにおける、適切な対話の構造化の方法を調査する。

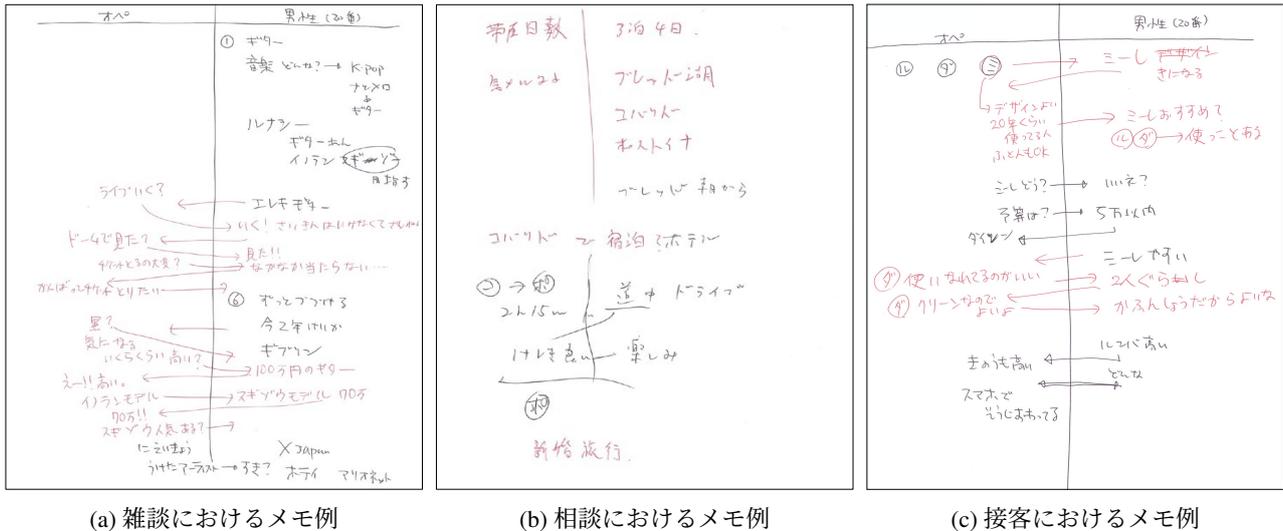


図3 実験終盤の、対話終了後のメモの例。(a)(c)オペレータとユーザの発話を左右に分け、隣接ペアのように発話同士を矢印で結んでいる。(b)隣接ペア形式と属性値対形式で情報をまとめている。

| | |
|----|-----------------------------------|
| U | あ、えと、Gibsonというメーカーで。 |
| O1 | あ、めっちゃめっちゃいいやつですね、はい。 |
| U | 奮発して買っちゃいました。 |
| O1 | あー、ですよねー、ちょっとね、お値段もしますよねー。 |
| U | そうです。 |
| O1 | はい。 |
| O1 | へえー、Gibsonの、何ですか？ |
| - | <入れ替わり> |
| U | えーと、ちょっと型番までは覚えてないですけども、色は黒いギターで。 |
| O2 | 黒！ |
| U | はい。 |
| O2 | Gibson。 |
| O2 | へえー。 |

図2 雑談の対話例。入れ替わり前のオペレータをO1、入れ替わり後のオペレータをO2、ユーザをUで示す。

3 対話介入実験

対話の引き継ぎに必要な情報を調査するために、対話に関するメモを引き継いで対話する実験（対話介入実験）を実施した。

3.1 実験の概要

対話介入実験の模式図を図1に示す。本実験では、2人のオペレータが一定時間ごとに入れ替わって1人のオペレータとして振る舞い、1人のユーザと対話する。なお、一方のオペレータがユーザと対話をしている間、もう一方のオペレータは別のユーザと対話をしている。切り替えでは、各オペレータは応対するユーザを交換する。すなわち、この実験では2つの対話が平行に実施されている。これ

は主に時間効率のためであるが、直前の対話から次の対話に即座に切り替えることは認知的負荷が高いと考えられ、そのような状況で作られるメモは対話を即座に把握できる有用なものになりうる。

オペレータ役として男女2名ずつ、合計4名を募集し、女性同士（オペレータペア1）、男性同士（オペレータペア2）の2つのペアに分けた。ユーザ役として20名を募った。

複数の対話タスクを実施することで多様な対話の引き継ぎに必要な要素を洗い出すことができると考え、雑談、相談、接客という3種類の異なる対話タスクを用意した。雑談では、ユーザが自身の最近ハマっていることについて話した。対話サービスを意識し、オペレータにはどちらかと言えば聞き手に回るよう指示した。相談では、オペレータが旅行代理店店員となってユーザの旅行計画の相談に乗った。相談にリアリティをもたせるため、旅行地には、対話の参加者があらかじめ知識を持たない地名としてスロベニアを選んだ。接客では、オペレータがユーザの要望を聞き、あらかじめ用意された3機種の掃除機から1機種を推薦した。

各オペレータペアは10名のユーザと3回ずつ対話（対話タスクをランダムな順番で1回ずつ実施）し、実験全体で60回の対話を行った。

3.2 オペレータの入れ替え

1つの対話の長さは10分とし、対話開始から2分、4分、6分、8分が経過した時点でオペレータを入れ替えた。オペレータが入れ替わることに

いて、ユーザには知らせていない。対話は、ウェブ会議システムのZoom¹⁾を用いビデオをオンにして行った。ただし、ユーザがオペレータの入れ替わりを判別できないように、オペレータはボイスチェンジャ²⁾を用い、アバタ³⁾を介して対話した。

オペレータには、対話を行っている最中に、対話を引き継ぐためのメモを取るように指示した。引き継ぎに際してどのような情報が必要かの予断を持たせないため、メモの書き方は指定しなかった。グラフィカルな情報が必要になる可能性があるため、メモは電子的ではなく、A4サイズの内紙1枚の片面に手書きで記入させた。メモは1つの対話につき1つとし、入れ替わりのタイミングで次のオペレータに手渡しで引き継がれるものとした（オペレータ同士は物理的に近いブースに配置した）。入れ替わったオペレータは、メモを参考にして対話を引き継ぎ、必要に応じて新たに情報を追記した。

図2に対話例、図3にメモの例を示す。2人のオペレータのうち、一方のオペレータは黒色でメモを取り、他方のオペレータは赤色でメモを取っていた。各対話の発話、手書きのメモ、ディスカッションで行われた発話の内容を分析するためこれらを書き起こした。

3.3 アンケート

オペレータとユーザは、各対話について、7段階リッカート尺度で対話に関する13個のアンケート項目（自然性、目的達成、満足度、文脈理解困難、次発話困難、他メモ有用、自メモ有用、連携、内容理解、個性、一貫性、信頼、注意）について同意の程度を回答した。

加えて、オペレータとユーザは、3種類の対話タスクを終えるごとに、対話タスクのうち最も話しやすかった／話しにくかったものを回答した。オペレータについてはさらに、もう一方のオペレータとの連携が最も円滑だと感じた／円滑でないと感じた対話タスクについても回答した。

オペレータのペアは、対話を終えるごとに、10分程度で、実施した対話の振り返りやメモの取り方の改善方法についてのディスカッションを行った。

4 対話およびメモの分析

1) <https://explore.zoom.us/products/meetings/>
 2) <https://clownfish-translator.com/voicechanger/>
 3) <https://store.steampowered.com/app/274920/FaceRig/>

表1 対話あたりの発話時間、発話回数、文字数、単語数、異なり単語数。括弧内は発話あたりの文字数。オペレータをO、ユーザをUで表す。

| | 雑談 | | 相談 | | 接客 | |
|--------|--------|--------|--------|--------|--------|--------|
| | O | U | O | U | O | U |
| 発話時間 | 3m57s | 6m26s | 6m19s | 4m21s | 6m13s | 4m39s |
| 発話回数 | 86.1 | 71.3 | 82.9 | 85.5 | 86.1 | 87.6 |
| 文字数 | 1312 | 2262 | 2175 | 1474 | 2057 | 1626 |
| | (15.2) | (31.7) | (25.0) | (17.2) | (23.9) | (18.6) |
| 単語数 | 778.0 | 1320.2 | 1217.3 | 883.3 | 1210.8 | 971.3 |
| 異なり単語数 | 177.3 | 278.2 | 280.2 | 192.7 | 279.2 | 212.2 |

表2 メモ1つあたりの統計量

| | 雑談 | 相談 | 接客 |
|--------|-------|-------|-------|
| 文字数 | 208.4 | 115.3 | 120.5 |
| 単語数 | 105.4 | 57.8 | 63.6 |
| 異なり単語数 | 40.7 | 18.8 | 21.1 |
| 矢印の数 | 151 | 116 | 115 |

4.1 対話の統計量

表1にオペレータ／ユーザごとの対話統計量を示す。10分程度の対話の中で、オペレータは80～90回、ユーザは70～90回程度の発話を行った。対話の傾向は対話タスクごとに異なっていた。具体的には、相談と接客では、オペレータの1発話あたりの文字数が多かった。これらの対話では、対話の主導権がオペレータにあったと考えられる。それに対して雑談では、オペレータが聞き手に回ったことから、対話の主導権はユーザにあったと考えられる。

4.2 メモの統計量

表2に、対話タスクごとのメモの文字数、単語数、異なり単語数、矢印の数を示す。メモの文字数と単語数は、ユーザの1発話あたりの平均文字数におおむね比例しており、相談、接客、雑談の順に多い。雑談は異なり単語数が多く、多様な話題が話されていることが分かる。

図4に、対話ごとに取られたメモの文字数の推移を示す。オペレータペア1のメモの文字数は2回目でピークを迎え、その後は文字数を減らして収束傾向にあった。最も序盤のメモを除き、オペレータとユーザの発話は左右に分けて記載されていた。オペレータペア2の相談と接客のメモの文字数は、実験の序盤から終盤まで大きく変化せず、メモのフォーマットも変化しなかった。全体としてメモの文字数とフォーマットは、オペレータペアごと、対話タスクごとにほぼ収束した。

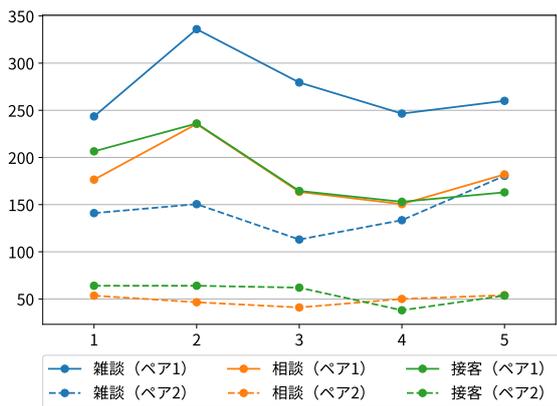


図4 メモの文字数の時系列遷移。横軸は2人のユーザへの応対を終えた回数を表す。3.2節の通り、オペレータは二人のユーザと入れ替わりながら対話をした。ここでは、並行して行われた2つの対話のメモの文字数の平均をプロットしている。

4.3 メモの分析

雑談や接客のメモでは、図 2a,2c のように矢印が多用された。矢印は「高い? → 100万円のギター → えー!! 高い。」のように、ユーザの発話に対するオペレータの反応という隣接ペアの関係を表していると言える。また、オペレータによると、雑談ではオペレータとユーザの発話を左右に分けてメモを取る方法が有用とのことであった。オペレータの発話をメモしておく理由は、矛盾をなくすためである。雑談の聞き手は、相槌を打つだけでなく、発話内容に同意したり、聞き手自身の意見や感想を述べたりする。したがって、聞き手であるオペレータが入れ替わった際に、入れ替わる前と異なる内容の発話をしないようにオペレータ自身についてのメモを取る必要がある。

相談のメモを図 2b に示す。メモの下半分には隣接ペアの形式でメモが取られているが、メモの上半分には「滞在日数 | 3泊4日」のように属性値対形式で情報が記入されている。相談や接客では、オペレータがユーザから聞き出すべき情報は決まることが多い。既に聞き出した情報とまだ聞き出していない情報がひと目で把握できるという点で、属性値対形式も有用と考えられる。

4.4 ディスカッションの分析

あるオペレータはディスカッションで、入れ替わり直後はメモをじっくり読む時間がないため、メモ

の一番下に記載された、最新の1行だけを参照して発話していたと述べた。そして、何かしら発話したうえで、メモの履歴をさかのぼって内容を確認していたとのことである。対話の引き継ぎでは、直前の対話履歴を参照でき、必要に応じて過去の対話履歴を効率的に参照できるような情報提示が有効であると考えられる。

4.5 アンケート結果の分析

アンケートの結果、オペレータは、3種類の対話のうち特に雑談について、自然性や満足度の高い対話ができ、メモも有用で、文脈理解や次発話作成に困難が少なかったと回答した。Bonferroni 補正を行った Wilcoxon の順位和検定を実施したところ、次発話困難（何を話せばよいか分からないことがあった）については雑談 - 相談間で、他メモ有用（もう一方のオペレータのメモは役に立った）については雑談 - 接客間で、1%の有意水準で差が見られた。メモにおいて隣接ペアが多用された雑談が引き継ぎやすいと評価されていることから、隣接ペア形式のメモの有効性が示唆された。他のタスクにおいても他メモ有用の評価は7段階中6前後と高かった。ユーザのアンケート評価は対話の種類に拠らずどれも7段階中4を超えており、オペレータの入れ替わりがあっても問題なく対話できていることが確認できた。

なお、3種類の対話タスクを比較するアンケートでは、オペレータは、3種類の対話タスクのうち雑談が最も話しやすく、かつオペレータ間の連携が円滑だったと回答した。相談と接客のような、タスク遂行の正確性が重要となる課題については、比較的引き継ぎが難しいことが示された。

5 おわりに

本稿では、対話の引き継ぎに必要な情報を実験的に調査した。具体的には、雑談、相談、接客に対して介入実験を行い、オペレータが作成したメモを分析した。その結果、隣接ペア形式および属性値対形式での情報が引き継ぎに有用であることが分かった。今後は、対話の引き継ぎに適したインターフェースを実装していきたい。例えば、対話履歴から隣接ペアを抽出して対話を要約する、対話履歴から属性値対形式で所定の情報を抽出する、などの要素技術に取り組む予定である。今後は、そうしたインターフェースの実装を行い、その有用性を検証したい。

謝辞

本研究は、JST ムーンショット型研究開発事業、JPMJMS2011 の支援を受けたものです。

参考文献

- [1] Tom B Brown, Benjamin Mann, Nick Ryder, Melanie Subbiah, Jared Kaplan, Prafulla Dhariwal, Arvind Neelakantan, Pranav Shyam, Girish Sastry, Amanda Askell, et al. Language models are few-shot learners. **arXiv preprint arXiv:2005.14165**, 2020.
- [2] Jan Pichl, Petr Marek, Jakub Konrád, Petr Lorenc, Van Duy Ta, and Jan Šedivý. Alquist 3.0: Alexa prize bot using conversational knowledge graph. **arXiv preprint arXiv:2011.03261**, 2020.
- [3] Wenjie Ou and Yue Lin. A clarifying question selection system from ntes_along in convai3 challenge. **arXiv preprint arXiv:2010.14202**, 2020.
- [4] Daniel Adiwardana, Minh-Thang Luong, David R So, Jamie Hall, Noah Fiedel, Romal Thoppilan, Zi Yang, Apoorv Kulshreshtha, Gaurav Nemade, Yifeng Lu, et al. Towards a human-like open-domain chatbot. **arXiv preprint arXiv:2001.09977**, 2020.
- [5] Stephen Roller, Emily Dinan, Naman Goyal, Da Ju, Mary Williamson, Yinhan Liu, Jing Xu, Myle Ott, Eric Michael Smith, Y-Lan Boureau, and Jason Weston. Recipes for building an open-domain chatbot. In **Proceedings of the 16th Conference of the European Chapter of the Association for Computational Linguistics: Main Volume**, pp. 300–325, 2021.
- [6] Ryuichiro Higashinaka, Masahiro Araki, Hiroshi Tsukahara, and Masahiro Mizukami. Integrated taxonomy of errors in chat-oriented dialogue systems. In **Proceedings of the 22nd Annual Meeting of the Special Interest Group on Discourse and Dialogue**, pp. 89–98, 2021.
- [7] K Shuster, S Poff, M Chen, D Kiela, and J Weston. Retrieval augmentation reduces hallucination in conversation 2021. **arXiv preprint arXiv:2104.07567**, 2021.
- [8] Tatsuya Kawahara, Naoyuki Muramatsu, Kenta Yamamoto, Divesh Lala, and Koji Inoue. Semi-autonomous avatar enabling unconstrained parallel conversations—seamless hybrid of woz and autonomous dialogue systems—. **Advanced Robotics**, pp. 1–7, 2021.
- [9] Allen L Gorin, Giuseppe Riccardi, and Jeremy H Wright. How may I help you? **Speech communication**, Vol. 23, No. 1-2, pp. 113–127, 1997.
- [10] Marilyn Walker, Irene Langkilde, Jerry Wright, Allen L Gorin, and Diane Litman. Learning to predict problematic situations in a spoken dialogue system: experiments with How may I help you? In **Proceedings of the 1st Meeting of the North American Chapter of the Association for Computational Linguistics**, pp. 210–217, 2000.
- [11] 山下紗苗, 東中竜一郎. 雑談に介入する際に必要な対話要約の調査. 人工知能学会研究会資料 言語・音声理解と対話処理研究会, pp. 86–91, 2021.
- [12] Marti A Hearst. Text tiling: Segmenting text into multi-paragraph subtopic passages. **Computational linguistics**, Vol. 23, No. 1, pp. 33–64, 1997.
- [13] Feng Jiang, Yaxin Fan, Xiaomin Chu, Peifeng Li, Qiaoming Zhu, and Fang Kong. Hierarchical macro discourse parsing based on topic segmentation. In **Proceedings of the AAAI Conference on Artificial Intelligence**, Vol. 35, pp. 13152–13160, 2021.
- [14] Barbara Grosz and Candace L Sidner. Attention, intentions, and the structure of discourse. **Computational linguistics**, Vol. 12, No. 3, pp. 175–204, 1986.
- [15] Staffan Larsson and David R Traum. Information state and dialogue management in the trindi dialogue move engine toolkit. **Natural language engineering**, Vol. 6, No. 3-4, pp. 323–340, 2000.
- [16] Emanuel A Schegloff. **Sequence organization in interaction: A primer in conversation analysis I**, Vol. 1. Cambridge university press, 2007.
- [17] Amy BM Tsui. Beyond the adjacency pair. **Language in Society**, Vol. 18, No. 4, pp. 545–564, 1989.
- [18] 榎本美香, 伝康晴, 松坂要佐. 3人会話における談話行為と受け手のラベリングとその基礎的分析. 人工知能学会研究会資料 言語・音声理解と対話処理研究会, Vol. 45, pp. 7–12, 2005.