

# 経験に基づく知識の想起・深化を行う 対話システムの開発に向けたコーパスの構築

渡邊 寛大<sup>1,2</sup> 河野 誠也<sup>2</sup> 湯口 彰重<sup>2,1</sup> 吉野 幸一郎<sup>2,1</sup>

<sup>1</sup> 奈良先端科学技術大学院大学

<sup>2</sup> 理化学研究所ガーディアンロボットプロジェクト

watanabe.kanta.wm7@is.naist.jp,

{seiya.kawano, akishige.yuguchi, koichiro.yoshino}@riken.jp

## 概要

対話システムが人間社会で活用されるには、高い信頼性と長期的な関係性を築くための記憶・学習能力が必要である。具体的には、複数回行われる対話の中で以前の対話を記憶・利用し、またそこに出現するキーワードなどから自身の知識を深めて次回以降の対話に活用できる機能が必要である。本研究ではこうした対話モデルを構築することを指向して、過去の対話セッションとそれに紐づけられた外部データベースを考慮した雑談対話コーパスを収集し、分析した。

## 1 はじめに

自然言語を用いて人間と対話をする対話システムは、医療や教育、生活支援をはじめとした様々な分野で注目され [1, 2, 3], 研究開発が活発に行われている。一方で、実際に活用されるために改善すべき課題は依然として存在する。例えば、現在の対話システムの多くは、単一セッションで毎回異なる話者と初対面の対話を行うことを想定している。しかし、医療であれば患者、教育であれば生徒、生活支援であれば持ち主のように、長期間に渡って特定のユーザと対話し、過去の対話内容を想起・活用し、そこから学習する能力が必要となる。こうした場合、対話者が人間であれば、複数回行われる対話（過去の対話セッション）の中で行われた対話を記憶・利用し、またそこに出現するキーワードなどから、自身の知識を深めて次回以降の対話に活用することができる。対話システムもこのように、対話履歴（経験）から知識を想起・深化して次回以降の対話で活用することが望まれる。対話システムがこのような経験に基づく知識の想起・深化・活用を行うこと

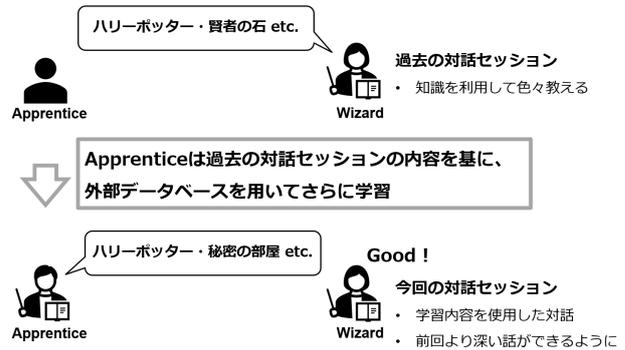


図1 本研究で想定する対話システムの概要

で、信頼関係の構築や、経験や知識の再検討による自己成長・自己理解・相互理解が可能となる。長期間に渡って特定のユーザと対話を行う対話システムを構築しようとする場合、こうした機能は必要不可欠である。

本研究では、過去の対話セッションに基づいて知識を想起し、外部データベースを用いて知識の深化を行い対話に活用する対話システムの開発を目指す (図1)。その第一段階として、過去の対話セッションとそれに紐づけられた外部データベースを考慮した雑談対話コーパスを構築し、分析した。結果として、考案した設定により目的に沿った対話を収集可能であることが確認できた。

## 2 関連研究

対話システムによる知識の想起やモデルの更新に関わる研究として、外部知識を対話システムの応答生成に組み込んだ研究 [4, 5, 6] や特定ユーザとの長期的な対話に焦点を当てた研究 [7, 8, 9] が存在する。本節では、これらの研究と本研究の違いについて説明する。

## 2.1 外部知識に基づく応答生成

Dinan ら [4] は、応答に必要な外部知識を選択し、選択した外部知識に紐づいた応答を生成する対話応答生成モデルを提案した。また、このモデルを実際に学習・評価するために、Wikipedia の知識に基づくオープンドメインの対話の大規模なコーパスである Wizard of Wikipedia を収集してモデルを構築・評価した。Komeili ら [5] は、対話文脈から生成したクエリでインターネット検索を行い、その結果を知識として利用して応答を生成する対話応答生成モデルを提案した。Shuster ら [6] は、neural-retrieval-in-the-loop 構造を知識ベース対話に利用して、会話能力を維持したまま知識を最大化するモデルを構築・評価した。これらの研究では、過去数ターンの発話から外部知識を検索して応答生成に使用する。これに対して本研究では、過去の対話セッションから知識を想起し、外部のデータベースから知識の深化を行い、それによってモデルが動的に変化する対話モデルを構築することを指向する。

## 2.2 長期間の文脈を考慮した応答生成

Xu ら [7] は、過去の対話セッションの要約が注釈されたマルチセッション対話のデータセット Multi-Session Chat を収集し、過去の対話セッションを要約して応答生成に使用するモデルを構築した。Xu ら [8] は、ユーザとシステム双方のペルソナ情報（テキストで記述された個人の情報）を過去セッションの対話履歴から抽出・更新・参照し、応答を生成する対話応答生成モデルを提案した。Bae ら [9] は、ペルソナ同士が矛盾を生まないように情報を動的に更新する手法を提案した。これらの研究では応答生成に過去セッションの対話履歴の一部や要約を考慮するが、本研究では過去の対話セッションの考慮に加えて、過去の対話セッションに基づいて知識を想起し、外部のデータベースから知識を深化して対話システムの応答生成に活用することを指向する。

## 3 知識の想起・深化を伴う対話コーパス

ここまでの説明のように、本研究では過去の対話セッションが存在する状況で、その対話セッションに基づく知識の想起と外部データベースを用いた知識の深化から対話モデルを更新するような対話システムの構築を目的としている。このため、まずは知

識の想起・深化を行うシステム役を人間が行った場合の対話を収集して分析する。この節では収集における具体的な対話タスクの設定と、その収集方法の詳細について述べる。

### 3.1 タスク設定

本研究では、過去の経験（過去の対話セッション）に基づいて知識を想起して、外部データベースを用いて深化し、次の対話に活用した対話コーパスを構築した。過去の対話セッションとしては Wizard of Wikipedia データセット [4] を日本語訳したもの、外部データベースには Wikipedia を使用した。Wizard of Wikipedia では知識を伝える Wizard と知識について尋ねる Apprentice という役割分担で対話が行われ、Wizard 側のシステムを構築することが行われてきた。今回は我々の目的である経験に基づく知識の想起・深化を行う様子を収録するため、Apprentice 側が過去の対話セッションに基づいて自主学習を行うシステムという想定で対話収集を行った。具体的には、過去の対話セッションで Wizard から与えられた知識を Apprentice がより深く学習してきたという設定で行った。Wizard of Wikipedia の設定と対比させた本研究での設定は以下の通りである。

#### 3.1.1 Wizard of Wikipedia の役割設定

**Apprentice** 会話の各段階で、Apprentice は Wizard に自由に話しかけ、好奇心旺盛な学習者の役を演じ、おしゃべりに熱中する。目標は、自分自身や相手が興味を持っている選ばれたトピックについて、会話を魅力的で楽しいものにしなが、深く掘り下げていくことである。

**Wizard** Wizard は情報検索システムにアクセスし、Wikipedia から会話に関連しそうなパラグラフを見ることができる。会話の順番が回ってくる前に、これらのパラグラフを読み、観察された知識に基づいて次の返答をすることが可能である。

#### 3.1.2 本研究の役割設定

本研究では、Wizard of Wikipedia から以下のようにそれぞれの役割の設定を更新している。

**Apprentice** 前回教えてもらったことに対して、新たに学んできたことを Wizard に提示して、さらに議論を深めるのが目的である。Apprentice は過去の対話セッションとそれに紐づけされた Wikipedia の記事をもとに知識を深化した発話を行う。

表 1 各コーパスの統計情報と語彙的特徴

コーパス	対話数	発話数	平均ターン数	トークン数 (N)	平均発話長	語彙数 (V)	Herdan's C
今回の対話セッション	10	200	20.0	5,771	28.9	1213	0.820
過去の対話セッション	10	93	9.3	2,130	22.9	481	0.806

表 2 役割ごとの統計情報と語彙的特徴

コーパス	役割	トークン数 (N)	平均発話長	語彙数 (V)	Herdan's C
今回の対話セッション	Apprentice	3,133	31.3	779	0.827
	Wizard	2,638	26.4	434	0.771
過去の対話セッション	Apprentice	1,029	21.9	186	0.753
	Wizard	1,101	23.9	295	0.812

**Wizard** Wizard は前回の対話で自身の知識を Apprentice に与えている。今回は Apprentice が学習してきたことを深掘りして、さらに議論を深めることが目的である。

## 3.2 収集方法

収集の方法には、テキストチャットを用いた。対話の収集は、日本語を母国語とする 3 名 (Wizard 役 1 名, Apprentice 役 2 名) の作業者によって行った。自然な対話を意識すること、過去の対話セッションとつじつまが合うこと、話し方を統一するため丁寧語で会話すること、想起対象である過去の対話セッションに紐づけられた Wikipedia の記事の情報をできるだけ発話に使用することを指示した。

### 3.2.1 実験の手順

1. 二人の作業者に役割 (Apprentice 役と Wizard 役) を与える。
2. 過去の対話セッションをお互い確認する。
3. Apprentice 役は記事を使用して知識を深めた対話の内容を考える。過去セッションの各単語に Wikipedia の見出し語へのリンクが付与されており、Apprentice 役も Wizard 役も自由に記事の内容を参照できる。
4. Apprentice 役から Wizard 役に「実験開始」の文字列を送信すると、実験が開始される。
5. Wizard 役が最初の発話を送信する。
6. 記事及び Wizard 役からの発話を考慮して、Apprentice 役が応答を送信する。
7. 応答の際に使用した知識が記載されているパラ

グラフのタイトルをスプレッドシートに記入する。

8. 応答の際に焦点を当てた過去の対話セッションの発話をスプレッドシートに記入する。
9. 5-8 の手順を合計 20 ターン行う。

## 3.3 対話行為ラベルの付与

収集した発話と過去の対話セッションを句点ごと分割し、対話行為ラベルを付与した。対話行為ラベルの定義には、ISO 24617-2 に基づいて Information-seeking, Information-providing, Commissives, Directives を使用した [10]。これらの対話行為のいずれにも属さない場合は Others とした。

### Information-seeking (IS)

話し手が聞き手から特定の情報を得ることを目的とする機能。(例：質問)

### Information-providing (IP)

話し手が聞き手に特定の情報を知らせることを目的とする機能。(例：応答, 賛成, 不賛成)

### Commissives (CM)

話し手が特定の行為を行うために自分自身の行為を聞き手に委ねることを目的とする機能。(例：依頼, 約束, 指示機能への承諾)

### Directives (DR)

話し手が聞き手に特定の行為を行わせることを目的とする機能。(例：指示, 要求, 提案)

## 4 分析

知識の想起・深化を伴う対話コーパス (今回の対話セッション) と過去の対話セッションの統計を表

1,2 に、実際の収集データの例を付録 A に示す。語彙の特徴については、各コーパスの基本的な統計情報に加えて、表層上の語彙の多様性を Herdan の  $C$  [11, 12] を用いて算出した。この指標は、総トークン数  $N$  と語彙数  $V$  を用いて次のように表される。

$$C = \frac{\log V}{\log N} \quad (1)$$

## 4.1 語彙多様性

語彙の多様性を表す  $C$  の値は、過去の対話セッションよりも知識の想起・深化を伴う対話コーパスの方がやや大きかった。このことから、知識の想起・深化を伴う対話コーパスには過去の対話セッションよりも多様な語彙が含まれていることが分かる。また、役割ごとの分析でも、Apprentice の  $C$  の値は過去の対話セッションと比べて大幅に向上している。このことから、過去の対話セッションから知識を想起し、外部データベースを用いて知識の深化を行い、次の対話に活用する対話設定は、語彙の多様性の向上に寄与すると考えられる。

## 4.2 対話行為

役割設定を更新したことで、Apprentice と Wizard の対話行為ラベルの分布傾向が逆転した。知識の想起・深化を伴う対話コーパスの各発話の対話行為として Apprentice 役からは Information-seeking が高い割合 (90.4%) で使用されたことが分かった。また、Apprentice 役は過去の対話セッションに紐づけられた Wikipedia の記事の情報を利用した発話を、全体の 84.0% 行っていた。この結果から、過去の経験に基づく知識の想起・深化を活用した発話を含む対話収録として、おおむね適切な設定ができたと考えられる。しかし、今回用いた Wizard 役が知識の深掘りに特に協力的であったことがこれらのスコアに大きく影響を与えた可能性も考えられる。雑談の中で自然に知識の深掘りを行ってもらうために、今後は Wizard 役をもう少し自由な設定にして収集を行う必要がある。

## 4.3 使用記事

知識の想起を伴う対話コーパスの 1 対話あたりの平均使用記事数は 2.6、1 記事による最大持続ターン平均は 12.2 であった。つまり、20 ターンの対話に

表 3 各対話行為の割合 [%]<sup>1)</sup>

コーパス	役割	IS	IP	CM	DR	Others
本研究	A	4.8	90.4	1.2	0.6	3.0
	W	28.1	33.0	5.4	7.7	25.8
WoW	A	21.0	43.2	3.7	8.6	23.5
	W	2.9	82.4	1.5	5.9	7.4

おいては多くの場合、複数の話題が現れ、話題の切り替わりが生じるような対話データが収集された。

## 4.4 対話パターン

今回は話し始めを Wizard 役で固定していたため、同じような発話が多くなり、話の展開が Wizard 役主導になることが多いという問題があった。発話の多様性と Apprentice 役主導の展開の対話を収集するため、話し始めが Apprentice 役の対話パターンも今後検討する必要がある。

さらに、Wizard 役の個性に関わらず雑談の中で自然に知識の深掘りを行う対話を収録する必要もある。そのため、Wizard 役に深掘りを行う設定を与える場合と与えない場合についても比較の検討の余地がある。

## 5 おわりに

本研究では、経験に基づいて知識を想起し、外部データベースを用いて知識の深化を行い対話に活用する対話システムの開発を目的として、知識の想起・深化を伴う対話コーパスを構築した。構築したコーパスを分析した結果、今回の対話収集設定はおおむね目的に沿った対話が収集可能であることが確認できた。今後の方針としては、データ規模の拡大、構築した対話コーパスを活用した対話システムの構築に取り組む。具体的には、焦点を当てるべき過去の対話セッションの発話・知識を検索するモデル、Wikipedia の記事から知識を深化するモデル、構築したコーパスで Finetuning を行った言語モデル、以上を組み合わせた対話システムの構築に取り組む。

## 謝辞

本研究は JSPS 科研費 JP22H04873 の助成を受けた。

1) A : Apprentice 役 W : Wizard 役

## 参考文献

- [1] Guangtao Zeng, Wenmian Yang, Zeqian Ju, Yue Yang, Sicheng Wang, Ruisi Zhang, Meng Zhou, Jiaqi Zeng, Xianguyu Dong, Ruoyu Zhang, Hongchao Fang, Penghui Zhu, Shu Chen, and Pengtao Xie. MedDialog: Large-scale medical dialogue datasets. In **Proceedings of the 2020 Conference on Empirical Methods in Natural Language Processing (EMNLP)**, pp. 9241–9250, Online, November 2020.
- [2] Diane Litman, Steve Young, Mark Gales, Kate Knill, Karen Ottewell, Rogier van Dalen, and David Vandyke. Towards using conversations with spoken dialogue systems in the automated assessment of non-native speakers of English. In **Proceedings of the 17th Annual Meeting of the Special Interest Group on Discourse and Dialogue**, pp. 270–275, Los Angeles, September 2016.
- [3] Eugene Agichtein, Yoelle Maarek, and Oleg Rokhlenko. Alexa prize taskbot challenge. In **Alexa Prize TaskBot Challenge Proceedings, 2022**.
- [4] Emily Dinan, Stephen Roller, Kurt Shuster, Angela Fan, Michael Auli, and Jason Weston. Wizard of wikipedia: Knowledge-powered conversational agents. In **International Conference on Learning Representations, 2019**.
- [5] Mojtaba Komeili, Kurt Shuster, and Jason Weston. Internet-augmented dialogue generation. In **Proceedings of the 60th Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)**, pp. 8460–8478, Dublin, Ireland, May 2022.
- [6] Kurt Shuster, Spencer Poff, Moya Chen, Douwe Kiela, and Jason Weston. Retrieval augmentation reduces hallucination in conversation. In **Findings of the Association for Computational Linguistics: EMNLP 2021**, pp. 3784–3803, Punta Cana, Dominican Republic, November 2021.
- [7] Jing Xu, Arthur Szlam, and Jason Weston. Beyond goldfish memory: Long-term open-domain conversation. In **Proceedings of the 60th Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)**, pp. 5180–5197, Dublin, Ireland, May 2022.
- [8] Xinchao Xu, Zhibin Gou, Wenquan Wu, Zheng-Yu Niu, Hua Wu, Haifeng Wang, and Shihang Wang. Long time no see! open-domain conversation with long-term persona memory. In **Findings of the Association for Computational Linguistics: ACL 2022**, pp. 2639–2650, Dublin, Ireland, May 2022.
- [9] Sanghwan Bae, Donghyun Kwak, Soyoung Kang, Min Young Lee, Sungdong Kim, Yuin Jeong, Hyeri Kim, Sang-Woo Lee, Woomyoung Park, and Nako Sung. Keep me updated! memory management in long-term conversations. In **Proceedings of the 2022 Conference on Empirical Methods in Natural Language Processing (EMNLP)**, 2022.
- [10] Harry Bunt, Jan Alexandersson, Jean Carletta, Jae-Woong Choe, Alex Chengyu Fang, Koiti Hasida, Kiyong Lee, Volha Petukhova, Andrei Popescu-Belis, Laurent Romary, Claudia Soria, and David Traum. Towards an ISO standard for dialogue act annotation. In **Proceedings of the Seventh International Conference on Language Resources and Evaluation (LREC’10)**, Valletta, Malta, May 2010.
- [11] G. Herdan. **Type-token Mathematics**. Janua linguarum, Studia memoriae Nicolai van Wijk dedicata. Series maior, 4. Mouton, 1960.
- [12] G. Herdan. **Quantitative Linguistics**. Butterworths, 1964.

# 付録

## A 収集データの例<sup>2)</sup>

### A.1 過去の対話セッション

表 4 過去の対話セッションの例

ターン	話者	発話
1	W	私はパラシュートとスカイダイビングが好きです。
2	A	どちらもやったことがないのですが、高所恐怖症の私には怖そうです。
3	W	でも、面白いゲームです。この歴史上初のパラシュートジャンプはアンドレ・ジャックによって行われました。
4	A	そうなんですか？フランス人みたいな名前だけど、何年に行ったのですか？
5	W	1797年12月22日に水素気球から飛び降りるという実験を行いました。
6	A	彼は成功したのでしょうか、それともそのスタントで挫折したのでしょうか？
7	W	軍がパラシュート技術を開発したのだから、成功でしょう。
8	A	そうですね。今はもっと安定していて、よくできていますね。
9	W	はい。飛行中の航空機に有用で、航空機乗務員を救うことができます。
10	A	そうですね。最近ほとんどもよく使われています。

### A.2 知識の想起・深化を伴う対話コーパス

表 5 収集した対話の例

ターン	話者	発話
1	W	お久しぶりです。前回、パラシュートについてお話ししましたが、あれから何か調べましたか？
2	A	もちろん調べました。パラシュートの形状はマッシュルーム型とラムエアー型の2種類あるんですね。
3	W	非常に興味深いです。もう少し詳しく教えていただけますか。
4	A	マッシュルーム型はキャノピーと呼ばれるパラシュートの傘が潰れにくいかわりに、コントロールが難しいみたいです。一方、ラムエアー型はその逆の特徴があります。
5	W	なるほど。パラシュートの歴史については何か分かりましたか。
6	A	はい。現代のパラシュート技術の基礎は1783年にフランスのルノルマンが開発したそうです。それ以前も史実として歴史が残っていますよ。
7	W	そうなんですね。非常に勉強になります。ちなみに日本人でのパラシュートでの落下第一号は空中分解事故で三菱IMF2試作機から脱出した中尾純利だそうですよ。
8	A	中尾さんは偶然にもパラソル翼が機体から外れて、パラシュートを使用したそうです。私は高所恐怖症なので、その場にも同じように使えないと思います。
9	W	私も同じ意見です。一応お伺いしますが、高所恐怖症についてはなにか調べてきましたか？
10	A	はい、調べました。単に高い所が苦手なだけだと、高所恐怖症と呼ぶそうで、いろいろ想像して恐怖を感じる症状がみられたら、高所恐怖症と言えるようですね。
11	W	そのような違いがあるのですか。高所恐怖症の治療法は確立されているのでしょうか？
12	A	治療法だと、曝露療法が短期的に有効みたいです。最近だとバーチャルリアリティ、つまりVRを利用した曝露療法が試されています。
13	W	VRを利用した曝露療法というのは初めて聞きました。少し話は変わりますが、興味深いエピソードとして赤ちゃんは本当は「高い高い」が怖いのではないかという説もあるようですよ。
14	A	へえ、TBSの「どうぶつ奇想天外！」で言われていそうですね。私からすれば「そんなこと当たり前だ」と思ったりもします。
15	W	高所恐怖症についてはよく分かりました。次はアンドレ・ジャックについてお話しませんか？
16	A	そうですね。アンドレ・ジャックの師匠は、水素気球の発明者ジャック・シャルルらしいですよ。知っていましたか？
17	W	それは初耳でした。アンドレ・ジャックの有名な逸話はご存じですか？
18	A	かの有名なナポレオンの戴冠記念で3000個の気球の飛ばした逸話ですね。やはり発明家のやることは常軌を逸しています。
19	W	同意見です。そのほかにアンドレ・ジャックについて何か調べたことはありますか？あれば是非教えてください。
20	A	彼の経歴によると、ナポレオン戦争の初期でイギリス軍に捕まって、オーストリア軍に引き渡され、その後ハンガリーのブダにて3年間の捕虜生活を送ったそうですよ。

表 6 発話に使用された知識と焦点を当てた過去の対話セッションの発話の例

ターン	発話に使用された Wikipedia の記事/パラグラフ	焦点を当てた過去の対話セッションの発話
2	パラシュート/解説	私はパラシュートとスカイダイビングが好きです。
4	パラシュート/解説	私はパラシュートとスカイダイビングが好きです。
6	パラシュート/歴史	軍がパラシュート技術を開発したのだから、成功でしょう。
8	パラシュート/歴史	はい、飛行中の航空機に有用で、航空機乗務員を救うことができます。
10	高所恐怖症/サマリー	どちらもやったことがないのですが、高所恐怖症の私には怖そうです。
12	高所恐怖症/治療	どちらもやったことがないのですが、高所恐怖症の私には怖そうです。
14	高所恐怖症/サマリー	どちらもやったことがないのですが、高所恐怖症の私には怖そうです。
16	アンドレ=ジャック・ガルヌラン/サマリー	この歴史上初のパラシュートジャンプはアンドレ・ジャックによって行われました。
18	アンドレ=ジャック・ガルヌラン/逸話	この歴史上初のパラシュートジャンプはアンドレ・ジャックによって行われました。
20	アンドレ=ジャック・ガルヌラン/経歴	この歴史上初のパラシュートジャンプはアンドレ・ジャックによって行われました。

2) A : Apprentice 役 W : Wizard 役