

<p>どんな研究</p>	<p>人が他者と仲良くなっていくときの話し方をチャットボットに学ばせるため、長期間の対話データの需要が高まっています。本研究では、人同士の8週間にわたる長期のテキストチャットを収集し、長期雑談対話モデルの実現に向けた基盤データの構築と分析に取り組んでいます。</p>
<p>どこが凄い</p>	<p>従来は、人が長期のチャットを想像して作成した、作例対話データが用いられていました。本研究では、実際に初対面話者が会話を重ねて親しくなる過程を記録し、従来の作例対話データよりもスピーチレベルや対話行為の観点で自然な長期間チャットデータを得ることに成功しました。</p>
<p>めざす未来</p>	<p>本データを用いて、過去の対話履歴との整合性だけでなく、話者の関係構築が進む観点からも自然な長期雑談対話モデルの開発に取り組めます。今後はAIと直感的にコミュニケーションできる社会をめざし、よく見知った他者と話すような感覚で利用できる対話システムを実現します。</p>

会話と親しさ

■人間の会話は相手との親しさによって異なります。

- 会話の目的が同じ雑談でも、初対面の相手と話すときと、親しい相手と話すときでは、その話題や話しぶりは違います。



初対面では、敬語が多く、当たり障りない話題



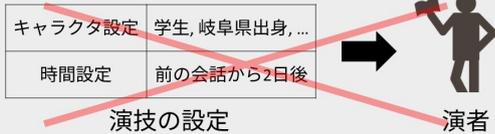
親しい相手とは、敬語が少なく、話者の内面に关わる話題

対話システムが人間と自然に長期間話すためには、会話を重ねて親しくなるときの話し方を**人間の会話から学ぶ必要**があります。

本研究の目的

- 人間と長期間自然に話せる対話システムの実現に向け、人間が親しくなるときの自然な長期間チャットデータを収集しました。

■従来の長期間チャットデータ：作例



対話相手と親しくなる様子が**演技では再現できないおそれ**

■本研究の長期間チャットデータ：実例



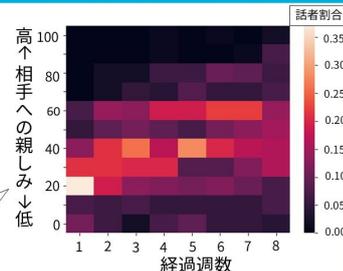
初対面話者が**実際に会話を重ねて親しくなる様子**を記録

初対面ペアによるチャットの長期間収集と分析

■データ収集

- 言語：日本語
- 参加ペア：20代の男女60ペア
- ペア相手：初対面の同性
- 対話期間：最長56日
- 対話量：1ペア1日10発話程度

対話相手への親しみは、実際に次第に高まりました



話題の変遷の例

日	話題	日	話題
1	気温が低いことと防寒対策	23	刑事法と民法の扱う分野の幅の広さ
2	緊急事態宣言中の人流	24	刑事法や学問の学習方法
3	お互いのプロフィールと話者Aの期末試験の感想	25	学習方法の続き
4	期末試験におけるレポート作成	26	将来子どもがほしいか
5	話者Aの残りの試験の予定	27	複数兄弟がいる人は将来子どもは少数でよい理由
6	話者Bが昨日アルバイトの残業だったこと、話者Aの今日行う勉強内容	28	子どもを持つ場合、希望性別
⋮	⋮	⋮	⋮

初日は相手と親しくないの**当たり障りない話題（天気）**

26日は相手に親しみが**わき深い話題（子どもがほしいか）**

■言語的特徴の評価

本研究の実例データと従来の作例データ（先行研究*の日本語版を作成）について、3週間の言語的特徴を比較しました。

→仮説検定の結果、話者の親しさと強く関連する敬体（です・ます文）と対話行為の出現割合が有意に異なることがわかりました。

長期間チャットデータの収集方法を作例から実例に変えることによって、人間が対話相手と親しくなるときの発話の傾向をより明瞭に反映できることを確認しました。

指標	週数	作例	実例
敬体	1	0.69	0.86
	2	0.59	0.86
	3	0.78	0.85
自己開示	1	0.43	0.46
	2	0.41	0.49
	3	0.40	0.47
質問	1	0.14	0.10
	2	0.13	0.08
	3	0.12	0.07
情報提供	1	0.09	0.06
	2	0.12	0.07
	3	0.12	0.08

※数値は1対話あたりの該当文の出現割合（中央値）

謝辞：本研究の一部はJSPS科研費JP19H05690, JP19H05693の助成を受けたものです

* Xu, Jing, et al. "Beyond goldfish memory: Long-term open-domain conversation." arXiv preprint arXiv:2107.07567(2021).

関連文献

[1] T. Arimoto, H. Sugiyama, H. Narimatsu, M. Mizukami, "Comparison of the Intimacy Process between Real and Acting-based Long-term Text Chats," in Proc. The 2024 Joint International Conference on Computational Linguistics, Language Resources and Evaluation (LREC-COLING 2024), 2024.

[2] 有本庸浩, 杉山弘晃, 成松宏美, 水上雅博, "長期雑談対話システムの実現に向けた長期テキストチャットの収集と分析," 第37回人工知能学会全国大会, 2023.

連絡先

有本 庸浩 (Tsunehiro Arimoto) 協創情報研究部 実世界インタラクション研究グループ