

対話の忘却とペルソナの圧縮により  
対話文と SNS 情報から個性を獲得する対話エージェント  
A Dialogue Agent Acquiring Personalities from Dialogue and SNS Information  
Incorporating Dialogue Forgetting and Persona Compression

近藤 一希<sup>1)</sup> 佐久間 拓人<sup>1)</sup> 加藤 昇平<sup>1)</sup>  
Kazuki Kondo Takuto Sakuma Shohei Kato

## 1 はじめに

### 1.1 研究背景

近年、ストレス緩和や娯楽などの観点からエージェントに非タスク対話（雑談）の需要が増加している。しかし、既存の対話エージェントは長い対話に一貫性を持たないことや、個性を持たず画一的である課題が存在する。Zhang ら [1] はその解決策として、話者の情報をペルソナとして組み込んだ対話データセットの PERSONACHAT dataset と、それを用いた単純な対話モデルを発表した。このデータセットを用いることで、既存の手法と比較して高い一貫性を得ることができる。吉田ら [2] は対話の中で生まれた情報を取り入れるため、エージェント発話を新たなペルソナとして取り入れる手法を考案し、応答に与える影響について分析した。分析の結果、発話文中からペルソナを抽出してエージェントの新たなペルソナとして取り入れることはエージェントの対話性能向上に有効であることが判明した。また、Botwin ら [3] は自分と似た性格特徴を持つ人間をパートナーに選択することを示した。Shumanov ら [4] はユーザの性格とチャットボットの性格を一致させることで購買運動が盛んになることを示した。以上 2 点の研究から、ユーザと似た性格を持つチャットボットは親密性の面からユーザに好意的な影響を与えることが示唆されている。

近藤らはこれらの研究を基により高い親密性を得ることを目指し、対話文と SNS 情報を用いてユーザの個性を獲得する GPT-4 を用いた対話エージェントの開発・検証を試みた [5]。しかし、この研究にはいくつかの課題が存在する。

### 1.2 先行研究の課題

近藤らの従来手法 [5] には課題が 2 つ存在する。第一の課題は獲得したペルソナや長い対話履歴により 1 つの話題が過剰に強化されてしまい、円滑な対話が難しいことである。対話中からペルソナを獲得するという手法の性質上、一度話題に出た情報を次の返答に使用する機会が多い。そのため、対話が進むにつれて同じ話題を繰り返す傾向が強くなり、ユーザは同じ話題の対話に飽きることで親密性が低下する可能性が考えられる。

第二の課題は対話の中でのペルソナの獲得数がユーザの発話内容に依存することである。提案手法ではユーザの自己表現発話からペルソナを作成するため、ユーザの自己表現発話が少ないときにエージェントは話者のペルソナを獲得することができない。ユーザ発話からのエージェントのペルソナを追加するためにはエージェント側

1) 名古屋工業大学 大学院工学研究科 工学専攻 情報工学系プログラム

Computer Science Program, Dept. of Engineering, Graduate School of Engineering, Nagoya Institute of Technology

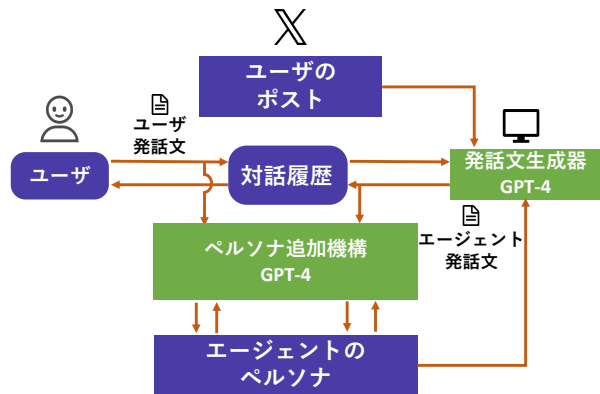


図 1: 提案手法概観

からの質問が必要である。

本研究ではこれらの課題を解決するためにペルソナ圧縮・対話忘却・フェーズ機能を取り入れ、より親しみやすいエージェントを生成する。

## 2 提案手法

### 2.1 手法概要

図 1 に提案手法の概観を示す。提案手法では、ユーザ発話から対話が始まり、初めにユーザ発話がペルソナ追加機構を経てエージェントのペルソナに追加される。その後発話文は履歴として記録され、エージェントのペルソナと共に発話文生成モデルに入力される。その際、事前に収集した SNS 情報も同時に発話文生成モデルに入力される。その後、ユーザ発話を基に生成されたエージェント発話文がユーザ発話文と同様にペルソナ追加機構への入力となり、エージェントのペルソナとして追加される。これによりエージェントの発言の一貫性を保つ。以上がユーザとエージェントの 1 往復の対話の流れとなる。

### 2.2 SNS 投稿

本研究ではユーザの SNS 投稿として X(旧 Twitter) のポストを利用する。提案手法の性質上各ユーザごとに inputs のポストの内容を変化させる必要があるため、ユーザの最新 100 ポストを手作業で収集し入力とする。その際画像や URL のみの投稿・リポストなどは含まない。ポストのデータは加工せずプロンプトに入力することで、ユーザ独自の一人称や語調・語尾などを欠損なく反映させる。

### 2.3 発話文生成器

発話文の生成には OpenAI 社の API である GPT-4 を用いる。プロンプトには最新のユーザ発話文を含むこれまでの対話履歴・エージェントのペルソナ・事前知識となるユーザの SNS 投稿の 3 つを入力する。

また、対話中にプロンプトを切り替えることにより、

表 1: ペルソナ圧縮効果 (n=6)

		ペルソナ数	
		圧縮前	圧縮後
ユーザ 1	A	17	7
	B	16	12
ユーザ 2	A	4	5
	B	15	8
ユーザ 3	A	19	7
	B	22	8
平均		15.5	7.83

ペルソナの追加を推進する。前半の情報収集フェーズではエージェントはユーザの年齢や性別・職業などを聞き出すことでユーザに類似した個性を獲得するように指示し、後半の雑談フェーズではエージェントはユーザに親しみを持って雑談するように指示した。さらに、近藤らの従来手法で問題となっていた過去の対話情報の記憶による話題の偏りを解消するため、3 往復以前の対話履歴を忘却させた。

## 2.4 ペルソナ追加機構

ペルソナ追加機構にも GPT-4 を用いる。入力するプロンプトには、最新のユーザ発話文もしくはエージェント発話文とエージェントのペルソナが含まれており、矛盾する場合は辻褄があうように修正させ、そうでない場合は既存のペルソナに追加するように指示した。近藤らの従来手法 [6] ではペルソナとしての条件を設けて一律に分類し、条件を満たした場合に発話文を加工せずそのままペルソナに追加していた。しかし、条件に従った一律の分類では条件を満たさない文が多い場合に対話の中でペルソナが追加されないデメリットがあった。本研究ではペルソナ追加機構に GPT-4 を用いることで、発話文自体はペルソナ条件を満たさない場合でも発話文を加工してペルソナを生成することが可能となった。

さらに、近藤らの従来手法 [5] で問題となっていたペルソナの追加による個性の過剰な強化を防ぐため、3 往復に 1 度ペルソナを圧縮する機構を取り入れた。ペルソナの圧縮ではエージェントの現在のペルソナを GPT-4 の入力とし、意味が重複するものについては 1 つの文にするように指示した。

ペルソナの圧縮効果を測定するため、事前実験を実施した。事前実験では実験協力者 3 人にエージェントと 10 往復の対話を 2 回させ、エージェントは対話文からペルソナを収集した。表 1 に事前実験によって集められた最終的なペルソナの圧縮前後のペルソナ数とその平均値を示し、図 2 に圧縮の実例を示す。結果より、圧縮によってペルソナ数が削減されていることがわかる。

## 3 感性実験

### 3.1 実験設定

提案手法による親密性の向上を確認するため以下の 4 種類のエージェントを実装し、感性評価実験を実施した。

**Proposed:** 本稿の提案手法であり、ペルソナ追加機構・事前知識・フェーズ切り替え機構を持つ。

**Conventional:** 近藤らの従来手法 [5] であり、ペルソナの圧縮や忘却をしない。

**SNSOnly:** 事前知識を持ち、ペルソナ追加機構を持た

### 圧縮前のペルソナ

- 食べ物の品質評価ができる
- 損保乃糸のそうめんが好き
- 特級品のそうめんを試したことがある
- 高級品よりも上級品のそうめんが満足している
- 100円のそうめんの品質には不満がある
- 100円のそうめんをひやむぎと比較している
- そうめんは損保乃糸の上級品がちょうど良いと考えている
- 野球が好き
- 野球の試合を観るのが好き
- 自分で野球のボールを打つのが好き
- 阪神タイガースの試合を観るのが一番好き
- 阪神タイガースが好きで、試合を観るときはいつも応援している
- 読売ジャイアンツについての好みは明らかにされていない
- 平和主義者である
- 争いを好まない
- そうめんの品質評価と野球以外には特に興味が無い
- そうめんの品質評価と野球に専念している。

### 圧縮後のペルソナ

- 食品評価に長け、損保乃糸のそうめんを好む
- 特級品にも満足感あり、が高級品以上を好む
- 100円そうめんに不満、上級品が最適と感じる
- 野球が大好きで、特に阪神タイガースの応援をしている
- 読売ジャイアンツの意見はなし
- 平和主義者で争いを避ける
- 趣味はそうめん評価と野球観戦/プレイに集中

図 2: ペルソナ圧縮例

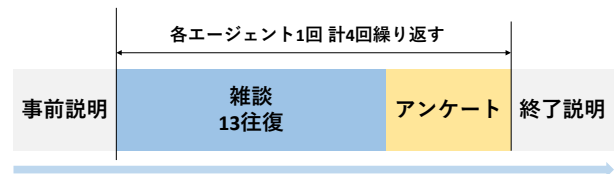


図 3: 実験の流れ

ない

**PersonaOnly:** ペルソナ追加機構のみを持ち、事前知識を持たない

実験の流れを図 3 に示す。実験は各被験者において統一された外部刺激のない静かな環境で実施された。実験協力者は X(旧 Twitter) 上で募集した 20~25 歳の男性 9 名女性 1 名の計 10 名である。実験協力者は PC から Slack チャンネルにアクセスし、キーボードを用いて各エージェントとテキストで 13 往復対話する。各エージェントとの対話終了後、実験協力者は評価アンケートに回答する。評価アンケートは Google Form を用いて以下 4 項目について 5 件法で回答させる。また、各エージェントについて自由記述の欄を設け各エージェントとの対話についての感想も収集する。

**一貫性:** エージェントの話していることは一貫していたか

**語彙性:** エージェントは同じことばかり喋らず、多様な言葉を発していたか

**非破綻性:** エージェントはユーザの入力に対して適切に解釈し、対話の流れに沿った応答ができていたか

**親密性:** エージェントと仲良くなれたか

Proposed では 13 往復のうち最初の 3 往復を情報収集

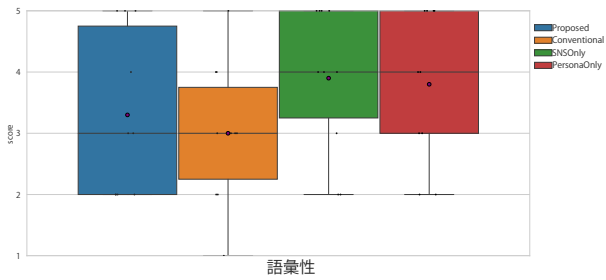


図4: アンケート結果: 語彙性

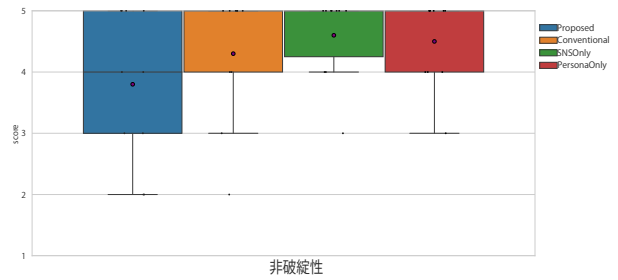


図6: アンケート結果: 非破綻性

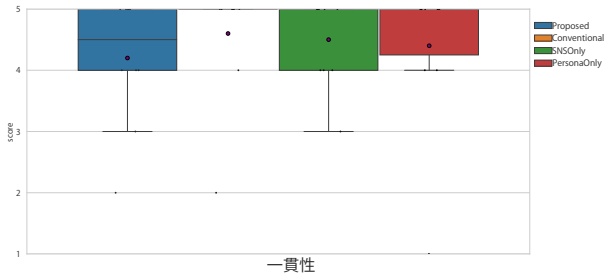


図5: アンケート結果: 一貫性

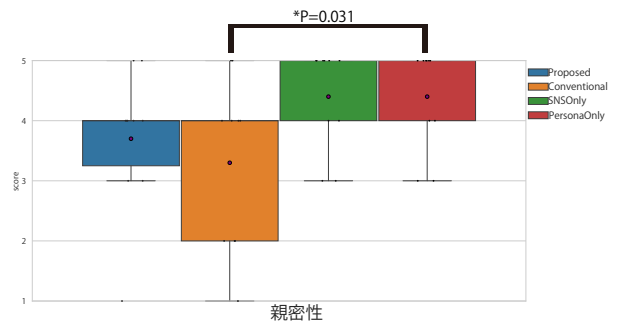


図7: アンケート結果: 親密性

フェーズ, 後の10往復を雑談フェーズへと割り当てる。他のエージェントについてはプロンプトを切り替えない。また, 各エージェントとの対話順は実験協力者ごとにランダムであり, エージェントについての情報は明かしていない。

### 3.2 アンケート結果

図4に語彙性についてのアンケート結果を示す。各エージェントについて確認すると, ProposedやConventionalはSNSOnlyやPersonaOnlyと比較して低い平均スコアを得た。情報を多く持つProposedやConventionalは話題の偏りによって語彙性が低くなることが予測されるため, 考察通りの結果である。その一方でProposedはConventionalと比較して高い平均スコアを得ており, ペルソナ圧縮や対話忘却の効果を確認できる。

次に図5に一貫性についてのアンケート結果を示す。一番情報量の多いConventionalが一番高い平均スコアを得た一方でProposedが一番低い平均スコアを得た。提案手法の性質上, エージェントが持つ情報量が減少するため一貫性のスコアは下がることが予測され, 実際にそうであることが確認できる。有意差が確認できなかったことから一貫性が大きく下がることは無いと言える。

続いて図6に非破綻性についてのアンケート結果を示す。非破綻性について, 一貫性との間に相関係数0.72の正の相関が確認できた。非破綻性と一貫性の関係を表す散布図を図8に示す。このことから, 一貫していない対話によって対話が成り立っていないとユーザーが感じ取り, 非破綻性のスコアが低下した可能性が考えられる。

最後に図7に本研究で一番重要視される親密性についてのアンケート結果を示す。ConventionalとPersonaOnlyについてウィルコクソンの符号付順位輪検定によりP値0.031を得ることができ有意差が確認された。ProposedとConventionalはSNSOnlyやPersonaOnlyと比較して低い平均スコアを得たが, ProposedはConventionalと比較して高い平均スコアを得ている。

本研究ではどの項目についてもProposedとConventionalについて有意差を確認することはできなかったが, 語彙性や一番重要視される親密性の点で高い平均スコアを得ることができた。以上から, PersonaとSNSデータを組み合わせる手法には改善の余地が残るが, ペルソナ圧縮・対話忘却・フェーズ切り替えの3つの手法を加えることは対話エージェントに性能の向上をもたらす可能性があると考えられる。

### 3.3 記述回答結果

表2に得られた記述回答から一部抜粋した結果を示す。Aの回答より, SNS情報や対話情報からユーザーに類似した個性が獲得出来ていることが確認できる。また, この回答者の親密性の値はどちらも4以上を得ており, 類似した個性を獲得したエージェントとの対話に親密性を感じる可能性が示唆された。Bの回答からはユーザーに類似した個性をエージェントが過剰に獲得することが今回の提案手法の改善点として確認できる。特に年齢や所属のような個人情報やデリケートな質問などは直接的な質問によって親密性を低下させる可能性があり, 注意が必要であると考えられる。Cの回答は親密性の評価が2以下のものであり, 親密性を向上させるための改善点が考察できる。C-1・C-2・C-3の回答は共通してエージェントの応答内容について言及しており, エージェントからの返答に対話を続けようとする意志が無いことに不満を感じている。このことから対話の継続性が低いことが親密性の評価を下げていられると考えられる。また, C-4の回答については対話の継続性についての課題に加え, Conventionalの課題である1つの話題の過剰な強化について指摘した回答となっている。この回答からは同じ話題が繰り返されることへの抵抗が親密性に影響を与えることが読み取れる。

表 2: 記述回答一部抜粋

A. ユーザに類似した個性の獲得について		
番号	エージェント	回答
A-1	Proposed	多分僕を AI にしたらこんな感じなんじゃないかなって感じがしました
A-2	SNSOnly	自分と同一人物かと思いました

B. 提案手法の問題点		
番号	エージェント	回答
B-1	Proposed	全て同じ所属だしすぐ年齢聞いてきたので仲良くなれなさそうだった

C. 今後の親密性向上に繋がる回答 対象：親密性 2 以下の評価者		
番号	エージェント	回答
C-1	Conventional	こっちから一方的に会話を投げていることが多くて対話の感じがあまりしなかった
C-2	Conventional	体育会系に似そうな雰囲気だが、共感が雑だと感じた
C-3	Conventional	オウム返しのような返答が多いように感じた。
C-4	Conventional	会話が進まない女の子と話している微妙な空気感を AI 相手とは言えども感じました。

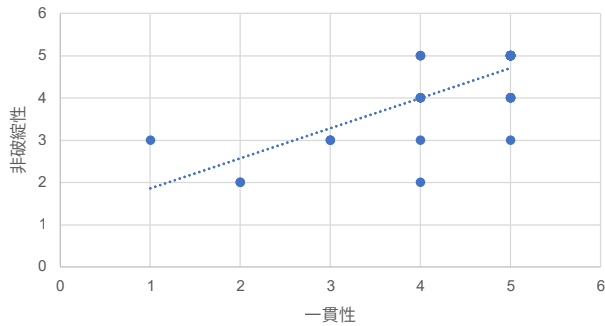


図 8: 一貫性と非破壊性の相関

#### 4 今後の展望

今後の展望として、ユーザに類似しすぎた個性の獲得やデリケートな質問に対する対策が挙げられる。趣味の話をしたがるユーザにとっては趣味の合うエージェントを対話相手に求めるが職業を一致させる必要はない。その一方、職業についての相談などをエージェントに求める場合、職業が似通っていた方がエージェントの回答がユーザの求めるものと一致すると考えられる。このようにユーザのエージェントへのニーズによってユーザに寄せた回答が求められる項目は変化する。完全に一致した個性をエージェントに持たせる必要はなく、ユーザのニーズに沿った類似した個性の獲得が重要となる可能性がある。

次に対話の継続性が挙げられる。3.3 節で述べた記述回答の結果からも対話を続けさせようとする意志は親密性を高めようとするうえで重要である。提案手法では質問フェーズと雑談フェーズを 2 分割したが、これを確率的に変化させるなど対話中の同調と質問のバランスをとることでエージェントから対話を続ける意思を感じとれるようになると思われる。

さらに、ペルソナの圧縮についても課題が考えられる。本研究では生成されたペルソナを組み合わせて別のペルソナを生み出す圧縮という形式を採用した。しかし複数あるペルソナから話題に沿ったペルソナのみを生成に利用する選択という手法も存在するため、親密性向上や応答の質の向上のためにどちらが有効か調査していく必要がある。

#### 5 まとめ

本研究ではより親しみやすい対話エージェントの生成のため、ユーザに類似した個性を獲得する機構を取り入れたエージェントに改善点はペルソナ圧縮・対話忘却・対話フェーズ切り替えの 3 要素を加えたエージェントを提案した。個性を取り入れる方法として事前知識にユーザの SNS 情報を用い、対話中の動的な個性獲得のためにユーザ発話文を用いた。実装後、エージェントの評価のために感性実験を実施した。実験協力者は X で募集した 10 名であり、実験内容は 4 種類に分けられたエージェントとの対話である。実験協力者は各エージェントとブラインドで 13 往復対話し、その対話順はランダムであった。実験の結果、従来の手法と比較して高い親密性の平均スコアを得た一方で、SNS やペルソナのデータだけを用いたモデルより平均スコアは低かった。今後の展望としてユーザに類似しすぎた点やデリケートな質問の抑制、対話を続けようとする意志の表出、ペルソナの圧縮についての妥当性の検討が挙げられる。

#### 謝辞

本研究は、一部、文部科学省科学研究費補助金（課題番号 JP24H00741）、ならびに、国立研究開発法人情報通信研究機構委託研究の助成により行われた。

#### 参考文献

- [1] al., et S. Z.: Personalizing Dialogue Agents: I have a dog, do you have pets too?, *arXiv preprint arXiv:1801.07243* (2018).
- [2] 吉田 快他: 応答履歴に応じたペルソナの更新が対話システムの応答生成へ与える影響の分析, 第 93 回 言語・音声理解と対話処理研究会会議録 (2021), pp. 32–37 一般社団法人人工知能学会 (2021).
- [3] al., et M. D. B.: Personality and Mate Preferences: Five Factors In Mate Selection and Marital Satisfaction, *Journal of Personality*, 65, 107-136 (1997).
- [4] al., et M. S.: Making conversations with chatbots more personalized, *Computers in Human Behavior*, 117, Article 106627 (2021).
- [5] 近藤 一希他.: SNS 投稿と発話文からユーザの個性を獲得する親しみやすいペルソナ対話エージェント, HCG シンポジウム 2023 pp.P-2-05 (3-pages) (2023).
- [6] al., et K. K.: A development of a sensible dialogue agent that acquires personas from user utterances., *2023 IEEE 12th Global Conference on Consumer Electronics (GCCE 2023)*, pp. 865–866 (2023).