

# 発散・深掘り対話戦略に基づく LLM 対話システムによる 悩みの内省支援

鈴木香保 上乃聖 李晃伸  
名古屋工業大学大学院 工学研究科

k.suzuki.451@stn.nitech.ac.jp {sei.ueno,ri}@nitech.ac.jp

## 概要

内省は感情整理に寄与するが、適切に行われな  
場合、否定的な反芻思考に陥る危険性がある。本研  
究では、内省が苦手なユーザを支援するため、発  
散・深掘り対話戦略に基づく対話システムを提案す  
る。ユーザの悩みに対し情報収集と深掘りを繰り返  
すことで対話を通して段階的に内省を行うことを  
目指す。性格特性に基づく 120 ケースのシミュレ  
ータを用いた評価実験を実施し、その有効性を検証  
する。検証の結果、提案手法はベースラインよりも内  
省プロセスの質が高いことが示された。

## 1 はじめに

内省は、自らの経験や感情、行動を振り返り、そ  
の意味を再解釈する心理的プロセスである [1]。内  
省は感情の整理やストレス軽減に寄与し、精神的健  
康の維持において重要な役割を果たすことが知られ  
ている。効果的な内省を実現するためには、複数の  
視点から問題を捉え直す多層的な視点の取り入れ  
[2] や、具体的な経験と抽象的な考え方を繰り返し  
ながら内省を深めるといった構造的なプロセス [3]  
が必要とされる。

しかし、その複雑さから悩みを抱える当事者がこ  
れらのプロセスを独自で実践することは容易ではな  
く、戦略的な介入や構造的な支援が求められる [1]。  
適切な内省が行われない場合、ネガティブな考えを  
繰り返し考え続ける状態に陥り、かえってストレス  
を増大させる危険性がある [4] ため、対話を通じた  
内省が有効とされている。

また近年、大規模言語モデル (LLM) の発展によ  
り、LLM を活用した内省支援システムの研究が進  
んでいる [5][6]。Arakawa ら [7] は、LLM チャット  
ボットによる内省支援の特性を分析し、傾聴や情報  
整理については効果的である一方、ユーザの価値観

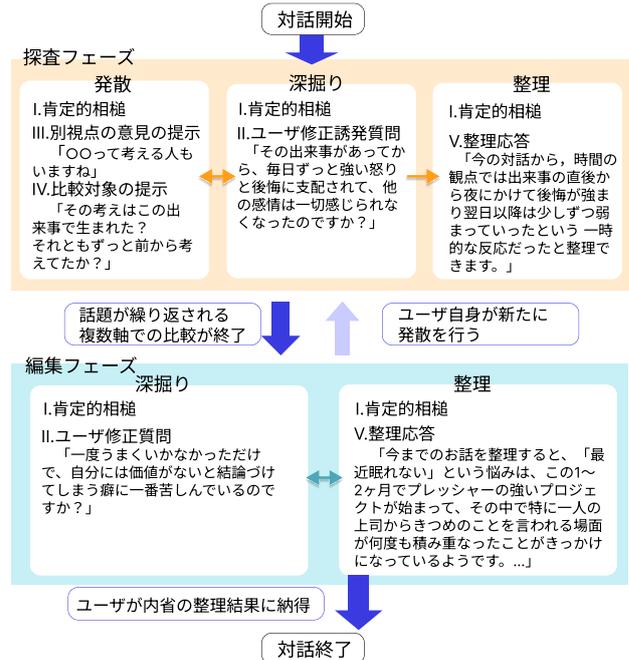


図 1 発散・深掘り対話戦略に基づく内省支援システムの  
全体構成

を揺さぶる批判的な問いかけ (例: "その目標は甘く  
ないか?") を行うことは困難であることを示した。  
これは、LLM がユーザの意図に従うよう訓練され  
ており、不快感を与える可能性のある質問を生成し  
にくいためである。この課題に対し、LLM が提示  
した複数の内省の切り口からユーザが選択するアプ  
ローチ [6] や、人間コーチとチャットボットを協働  
させる形態 [7] が提案されている。

## 2 発散・深掘り対話戦略に基づく 悩みの内省支援

本研究では、内省が苦手なユーザを支援するた  
め、システム主導で段階的に内省を促す対話システ  
ムを提案する。本システムは、2 段階のプロセスと  
誘発的な質問により、ユーザの認知的負担を軽減し  
つつ構造的な内省を実現することを狙う。

図1に提案する対話戦略の全体像を示す。本研究の対話戦略では、ユーザの悩みについてまず探査フェーズで時間軸や対人軸など多角的に情報を収集する。悩みについて情報が十分に集まった段階で、編集フェーズに移行し、収集した情報を統合して内省結果として提示、結果をもとに内省の原因となったユーザの悩みや価値観について深掘りを行う。

## 2.1 全体構成

内省の苦手なユーザの負担軽減のため、対話戦略を探査と編集の2フェーズで構成する。探査フェーズで十分な情報を収集したのち、編集フェーズで統合することで段階的な内省を目指す。

**(1) 探査フェーズ：情報収集と発散** 探査フェーズでは、ユーザの悩みに関する情報を多面的に収集する。内省における具体と抽象を繰り返すプロセス[3]をなぞり、話題を詳細に聞きこむ深掘りと、別の話題に遷移する発散、および会話内容のまとめとフィードバックを行う整理、これらを交互に繰り返すことで段階的に情報を収集していく。

**(2) 編集フェーズ：統合と構造化** 編集フェーズでは、探査フェーズで収集した情報を統合し、時系列や因果関係を明示しながら前提・価値観を構造化して提示する。ユーザのリアクションに応じて深掘りを行い、新たな意見が出た場合は探査フェーズに戻る。

## 2.2 発散・深掘り・整理

探査フェーズでは多層的視点による情報収集（発散）、深掘り、整理を組み合わせる。編集フェーズでは特に悩みの本質部分についての深掘り、整理を行う。

**発散：多層的視点による情報収集** ユーザの悩みについて以下の視点から質問し、多面的に情報を収集する。

**時間軸：**過去の経験、現在の状況、未来への展望

**対人関係軸：**特定の人への反応の一般性やパターン

**自己比較軸：**過去の自分と現在の自分の比較

**理想比較軸：**理想や目標と現実とのギャップ

さらに、応答戦略（2.3に詳述）により客観的な第三者の視点や別の考え方を提示する。

**深掘り：段階的な内省の深化** ユーザ修正誘発質問（2.3に詳述）を用いて深掘りを行う。応答例は図1内に示す。初期段階では具体的な状況や出来事

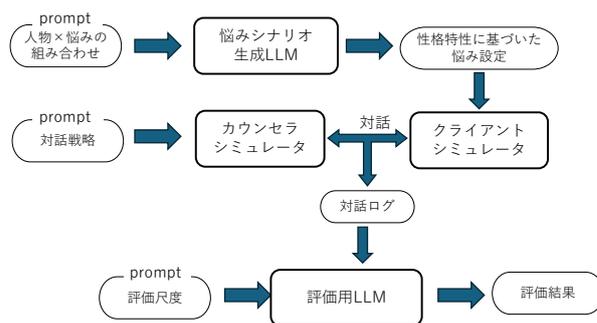


図2 シミュレータ全体図

について質問し、対話が進むにつれ、ユーザの発言から見える前提や価値観について質問する。

## 2.3 応答種別

対話戦略では、以下の応答戦略を使い分ける。発散では比較対象の提示と別視点の提示を、深掘りではユーザ修正誘発質問を、整理では整理応答を用いる。肯定的相槌は全体を通じて用いる。

**I 肯定的相槌：**ユーザの話を傾聴し共感を示す

**II ユーザ修正誘発質問：**誇張表現を用いた質問により、ユーザ自身が訂正・気づきを得るよう促す  
認知行動療法のソクラテス問答法[8]を参考に設計しており、感情・状況・頻度・空間の4つの誇張軸を用いる

**III 別視点の意見の提示：**客観的な第三者の視点を提示し、多様な考え方に触れる機会を提供する

**IV 比較対象の提示：**時間軸・対人関係軸での比較を示し、相対的な理解を促す

**V 整理応答：**探査フェーズでは部分的整理、編集フェーズでは全体的整理を行う

## 3 評価実験

本研究では、提案手法の有効性を検証するため、LLMシミュレータを用いた評価実験を実施する。

### 3.1 シミュレータシステム

シミュレータシステムは、悩みシナリオ生成LLM、カウンセラシミュレータ、クライアントシミュレータ、評価用LLMの4つのモジュールで構成される（図2）。

カウンセラシミュレータには、提案システムとベースラインシステムを用意する。提案システムは2節で述べた対話戦略と内省支援という対話目的をプロンプトに記述し、ベースラインは対話戦略を

指定せず内省支援という対話目的のみを指定した LLM の標準的な対話能力を活用したシンプルなプロンプト設計とした。

全てのシステムにて LLM のモデルは gpt-5.1-2025-11-13 を使用する。

LLM ベースのシミュレータでは、過度に協力的で明晰な応答が返ってくるため、実際の人間らしさとは離れた対話になりやすい。本研究でも、実際の対話場面に近い、リアルなクライアント応答を実現する必要がある。そこで、MINDVOYAGER [9] の段階的情報開示アプローチを参考にする。MINDVOYAGER は、表層的な感情・行動と深層的な経験・スキーマを区別し、対話の進行に応じて段階的に開示することでリアリティのあるクライアント応答を実現する動的シミュレータである。

クライアントシミュレータには基本的なペルソナ（年齢、職業、性格特性）と表層的な悩みのみを提示し、深層的な価値観や詳細な背景情報は明示的に与えない。これにより、LLM が「最初から全てを知っている」状態を避け、対話を通じて徐々に思い出したり気づいたりするプロセスを再現する。悩みシナリオは、内閣府調査 [10] の 20 代の主要な悩み 6 種類と、表 1 に示す Big Five 性格特性 [11] の典型的 4 パターン [12] を組み合わせ、計 120 ケース（6 種類 × 4 パターン × 5 つの人物背景）を生成する。人物背景についても LLM を使用して生成した。評価用 LLM が人間の評価傾向を反映できるかを確認するため、KokoroChat データセット [13] を用いた検証を行った。KokoroChat データセットは、カウンセリング対話ログを対象に、専門家が 20 項目の評価基準に基づいて主観評価を付与した対話ログと、その評価結果を対にして構成されたデータセットである。実験 1 で使用する評価項目（3.2.1 に詳述）それぞれに対応する KokoroChat データセットの評価項目を選択し、データセット内の対話ログに対する人間評価スコアと評価用 LLM のスコアを比較した。その結果、全 4 項目で有意な正の相関 ( $r=0.63\sim 0.73$ ,  $p<0.01$ ) を確認した（表 2）。この結果から、評価用 LLM が人間の評価傾向と一定の対応関係を持つことが示唆されたため、以降の評価に用いる。

## 3.2 実験 1：カウンセリング対話品質による評価

提案手法の有効性を検証するため、一般的なカウンセリング対話の内省評価の枠組みでシステム全体を評価する。

表 1 Big Five 特性に基づいた 4 種類の性格特性

パターン	特徴
Role model	成熟・社交的・安定的
Reserved	落ち着いた・控えめ
Self-centred	衝動的で自己中心的
Average	一般的・多数派

表 2 KokoroChat データセット評価との相関係数

評価項目	KokoroChat 項目	相関係数 ( $r$ )	$p$ 値
意味の変化	問題の明確化	0.73	0.0001
洞察の深さと 具体性	新しい気づきや 体験があったか	0.65	0.0005
感情の変化	希望や期待を感じ られた	0.63	0.0007
クライアント の主体性	一緒に考えなが ら取り組めたか	0.63	0.0053

### 3.2.1 評価指標

効果的な内省には、経験の言語化、新たな気づき、といったプロセスと、それらにユーザが意欲的に取り組むことが重要であるため、カウンセリング対話の質として以下の項目を設定した。

対話の質を以下の 4 項目で評価した。

- 1.意味の変化：悩みが具体的に言語化されたか
- 2.洞察の深さと具体性：新しい気づきや発見があったか
- 3.感情の変化：前向きな感情や期待を感じたか
- 4.クライアントの主体性：クライアントが主体的に関わっているか

各項目は 10 段階評価とした。

### 3.2.2 実験 1 結果と考察

120 ケースの評価結果を表 3 に示す。総合スコアは、全体的にベースラインを提案手法が下回った。特に「感情の変化」項目で 1.73 ポイントの大きな差が見られた。

本評価は、カウンセリング対話の一般的な品質を測定するものである。一方、提案手法では、ユーザ修正誘発質問（2.3 の応答戦略 II）により、図 1 の応答例に示すような極端な誇張表現を含む問いかけを行う。これらは肯定・否定で答えられる形式を取っているが、「毎日ずっと強い怒りと後悔に支配されて」といった極端な表現により、自身の認識を修正しながら答えることを促される。

このような発話は、ユーザに自身の認識を見直させ、内省を促す効果がある一方で、答えづらさや心理的な抵抗を生じさせる可能性がある。結果として、ユーザに負担がかかる傾向があったと推察され、特に「感情の変化」や「希望や期待」の評価が

**表 3 実験 1 評価結果 (10 点満点, 120 ケースの平均値)**

カテゴリ	提案手法	ベースライン
意味の変化	8.58	<b>8.80</b>
洞察の深さと具体性	8.22	<b>8.55</b>
感情の変化	5.96	<b>7.69</b>
クライアントの主体性	8.30	<b>8.78</b>

**表 4 実験 2 評価結果 (10 点満点, 120 ケースの平均値)**

評価項目	提案手法	ベースライン
自己理解の深化度	<b>8.73</b>	8.47
洞察の具体性	8.66	<b>8.68</b>
自己表現の主体性	<b>8.85</b>	8.46
内省志向性	<b>9.79</b>	8.94

下がったと考えられる。

しかし、本研究の目的は「内省が苦手なユーザの内省支援」であり、結果だけでなく内省プロセスそのものの質を評価する必要がある。そこで、内省により注目した詳細な評価を行うため、評価指標を再設計し実験 2 を実施した。

### 3.3 実験 2：内省プロセス品質による評価

#### 3.3.1 評価指標の再設計

内省は具体と抽象を繰り返す構造的なプロセスを必要とする [3] ため、カウンセリング対話の質として以下の項目を設定する。

- 1.自己理解の深化度 (深化度)：**対話を通じた構造的な深化ができているか (抽象→具体, 過去と現在の統合)
- 2.洞察の具体性 (具体性)：**「なんとなく」から具体的な状況・頻度の特定に変化しているか
- 3.自己表現の主体性 (主体性)：**受動的同意でなく、能動的な修正・精緻化をしているか
- 4.内省志向性：**行動計画でなく内省に重点を置いているか

各項目は 10 段階評価とした。

#### 3.3.2 実験 2 結果

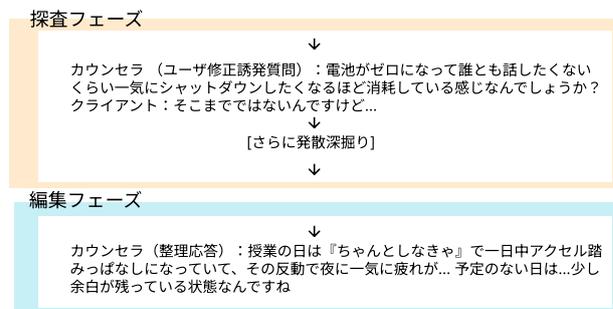
同一の 120 ケースに対して、シミュレータで評価した結果を表 4 に示す。提案手法はベースラインと同等以上の性能を示し、全体的に内省に関して良好な結果が得られており、提案する対話戦略によって内省を行うことができた。

応答種別と評価項目の相関分析結果を表 5 に示す。ユーザ修正誘発質問は主体性 ( $r=+0.223, p<0.05$ ) と深化度 ( $r=+0.236, p<0.01$ ) に正の相関を、編集フェーズの整理応答は深化度 ( $r=+0.319, p<0.01$ ) に強い正の相関を示した。

**表 5 応答種別と評価項目の相関 (提案手法,  $n = 120$ ) 主体性：自己表現の主体性, 深化度：自己理解の深化度, 具体性：洞察の具体性**

応答種別	1 対話の平均使用回数	深化度	主体性
ユーザ修正誘発質問	12.52	+0.236**	+0.223*
別視点の提示	2.28	+0.034	+0.036
比較対象の提示	10.91	+0.220*	+0.092
整理応答 (探査)	16.04	+0.145	+0.087
整理応答 (編集)	11.22	+0.319**	+0.139

\*  $p < 0.05$ , \*\*  $p < 0.01$



**図 3 提案手法の対話例 (探査フェーズから編集フェーズへの流れ)**

## 4 考察とまとめ

本研究では、内省が苦手なユーザを支援するため、発散・深掘り対話戦略に基づく対話システムを提案した。

実験 1 と 2 の結果から、提案手法の特性が示された。実験 1 では、カウンセリング対話としての全体的な品質評価においてベースラインを下回り、特に「感情の変化」「希望や期待」で顕著であった。これは、図 3 に示したようなユーザ修正誘発質問が極端な誇張表現を含み、ユーザに心理的負担をかけるためと考えられる。

一方、実験 2 では、内省プロセスの質という観点から評価すると、提案手法がベースラインを上回った。図 3 に示すように、ユーザ修正誘発質問により段階的に深掘りを行い、整理応答で情報を統合することで、具体と抽象を繰り返す構造的な内省が実現されている。実験 1 で負担となった発話が、内省という観点では効果的に機能したことが示唆される。この対比から、提案手法はユーザに心理的負荷をかける一方で、構造的な内省プロセスを実現していることが明らかになった。

今後の展望として、内省支援の効果を維持しつつ感情的満足感を両立させる研究の実施、また実際の人間ユーザを対象とした主観評価の実施が挙げられる。

## 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] Simone Jennissen, Julia Huber, Johannes C Ehrental, Henning Schauenburg, and Ulrike Dinger. Association between insight and outcome of psychotherapy: Systematic review and meta-analysis. **American Journal of Psychiatry**, Vol. 175, No. 10, pp. 961–969, 2018.
- [2] Alain Morin. Levels of consciousness and self-awareness: A comparison and integration of various neurocognitive views. **Consciousness and cognition**, Vol. 15, No. 2, pp. 358–371, 2006.
- [3] Jack Mezirow. **Transformative dimensions of adult learning**. ERIC, 1991.
- [4] Anthony M Grant, John Franklin, and Peter Langford. The self-reflection and insight scale: A new measure of private self-consciousness. **Social Behavior and Personality: an international journal**, Vol. 30, No. 8, pp. 821–835, 2002.
- [5] Taewan Kim, Seolyeong Bae, Hyun Ah Kim, Su-woo Lee, Hwajung Hong, Chanmo Yang, and Young-Ho Kim. Mindfuldiary: Harnessing large language model to support psychiatric patients' journaling. In **Proceedings of the 2024 CHI Conference on Human Factors in Computing Systems**, pp. 1–20, 2024.
- [6] Inhwa Song, SoHyun Park, Sachin R Pendse, Jessica Lee Schleider, Munmun De Choudhury, and Young-Ho Kim. Exploreself: Fostering user-driven exploration and reflection on personal challenges with adaptive guidance by large language models. In **Proceedings of the 2025 CHI Conference on Human Factors in Computing Systems**, pp. 1–22, 2025.
- [7] Riku Arakawa and Hiromu Yakura. Coaching copilot: blended form of an llm-powered chatbot and a human coach to effectively support self-reflection for leadership growth. In **Proceedings of the 6th ACM Conference on Conversational User Interfaces**, pp. 1–14, 2024.
- [8] Aaron T Beck, A John Rush, Brian F Shaw, Gary Emery, Robert J DeRubeis, and Steven D Hollon. **Cognitive therapy of depression**. Guilford Publications, 2024.
- [9] Minju Kim, Dongje Yoo, Yeonjun Hwang, Minseok Kang, Namyoun Kim, Minju Gwak, Beong-woo Kwak, Hyungjoo Chae, Harim Kim, Yunjoong Lee, et al. Can you share your story? modeling clients' metacognition and openness for llm therapist evaluation. In **Findings of the Association for Computational Linguistics: ACL 2025**, pp. 25943–25962, 2025.
- [10] 内閣府男女共同参画局. 特 - 19 図 悩みやストレスの原因別割合 (男女、年齢階級別・悩みやストレスのある者・令和 4 (2022) 年) . 男女共同参画白書 令和 6 年版 (特集図表ページ) , 2024. 図表 00-19a, White Paper on Gender Equality 2024, Cabinet Office, Government of Japan, HTML 表形式. [https://www.gender.go.jp/about\\_danjo/whitepaper/r06/zentai/html/zuhyo/zuhyo00-19a.html](https://www.gender.go.jp/about_danjo/whitepaper/r06/zentai/html/zuhyo/zuhyo00-19a.html) (最終アクセス: 2025 年 1 月 9 日) .
- [11] Samuel D Gosling, Peter J Rentfrow, and William B Swann Jr. A very brief measure of the big-five personality domains. **Journal of Research in personality**, Vol. 37, No. 6, pp. 504–528, 2003.
- [12] André Kerber, Marcus Roth, and Philipp Yorck Herzberg. Personality types revisited—a literature-informed and data-driven approach to an integration of prototypical and dimensional constructs of personality description. **PLoS one**, Vol. 16, No. 1, 2021.
- [13] Zhiyang Qi, Takumasa Kaneko, Keiko Takamizo, Mariko Ukiyo, and Michimasa Inaba. KokoroChat: A Japanese psychological counseling dialogue dataset collected via role-playing by trained counselors. In Wanxiang Che, Joyce Nabende, Ekaterina Shutova, and Mohammad Taher Pilehvar, editors, **Proceedings of the 63rd Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)**, pp. 12424–12443, Vienna, Austria, 2025. Association for Computational Linguistics.