

探究学習におけるメンタリングの自動化に向けた BDIモデルベース対話アーキテクチャの有用性の検討

Investigating the usefulness of the BDI model-based dialogue architecture for automating mentoring in inquiry-based learning

古賀 日南乃¹, 佐藤 匠¹, 大美浪 海晟¹, 飯田 愛結¹, 奥岡 耕平¹,
佐々木 康輔², 野田 尚志², 森口 昌和², 大澤 正彦¹
Hinano Koga¹, Takumi Sato¹, Kaisei Ominami¹, Ayu Iida¹, Kohei Okuoka¹, Sasaki
Kousuke², Noda Hisashi², Moriguchi Masakazu², Osawa Masahiko¹

¹ 日本大学文理学部, ² NEC ソリューションイノベータ株式会社

¹College of Humanities and Sciences, Nihon University, ²NEC Solution Innovators, Ltd.

chhi22083@g.nihon-u.ac.jp

概要

本研究の目的は、探究学習において言外の意味を扱った対話を実現できるメンタリングチャットボットの開発である。著者らは先行研究として、BDIモデルベースの自己/他者モデルと大規模言語モデルを統合することで、大規模言語モデルが言外の意味を扱えるようにするアプローチを提案している。本研究では、このアプローチをメンタリングチャットボットに導入し、探究学習におけるメンタリングにおいても有効であるか評価するためにケーススタディを実施した。

キーワード：メンタリング, 探究学習, 他者モデル, BDIモデル, 大規模言語モデル

1. はじめに

探究学習とは、学習者自らが定義した問題に対して情報収集や他者連携を通して解決することで問題解決や情報処理の能力を獲得する教育手法である。探究学習では学習者個人に応じた個別支援、特に対話によって学習者の探究を促進するメンタリングが重要視されている。

しかし、教員への過負荷が問題視されている現在の教育体制で個別支援の導入は困難である。そこで著者らは、教員によるメンタリングを自動化するために、学習者が問いを深めるために有効とされる質問設計の指針である「問いのデザイン」と大規模言語モデルを組み合わせたメンタリングチャットボットを提案した(大美浪他, 2024)。大規模言語モデルのみを用いた場合と比べて、提案システムが相談者の問いについて定義や意図といった問いを深めるようなメンタリングを行える可能性を示している。

一方で大規模言語モデルは、発話した字義的な表現に含まれる情報だけでなく発話者の情報や文脈等の「言外の意味」を扱う必要があるコミュニケーション

において、十分な性能を実現できていないことが指摘されている(Hu et al., 2023)。この課題に対して、著者らの一部は自己/他者モデルと大規模言語モデルを統合するアプローチを提案し、統合することによって大規模言語モデルのみを使用する場合に比べて言外の意味を扱うコミュニケーションタスクにおける性能が改善することを示している(飯田他, 2024)。

メンタリングにおいても、メンターに対する相談者の自己開示の程度は異なるため、常に相談者の発話に全ての情報が含まれているとは限らない。そのため、メンターは相談者の意図などの言外の内部状態を認識し、対話戦略を考えることが求められる。しかし、先行研究では一往復の会話を想定した評価を行っていたため、継続的な対話を行うメンタリングにおいても同様のアプローチが有用であるかは不明である。

そこで本研究では、提案したアプローチが探究学習におけるメンタリングにおいても有効であるか調査すると共に、他のタスクに応用する上での課題点を明らかにするための初期検討を行う。具体的には、先行研究で提案した対話アーキテクチャに探究学習におけるメンタリングの設定を与えることでメンタリングチャットボットを実装した。実装したチャットボットを用いて、実際に学生を対象として継続的な対話によるメンタリングのケーススタディを実施し、相談者による印象評価と対話履歴の分析を行う。

2. 関連研究

2.1 問いのデザインに基づくメンタリングチャットボット

著者らの先行研究では、安齋・小田(2021)の提案する相談者の問いを効果的に深めるための質問設計の指針「問いのデザイン」を、大規模言語モデルへのプロンプトに組み込んだメンタリングチャットボットを提

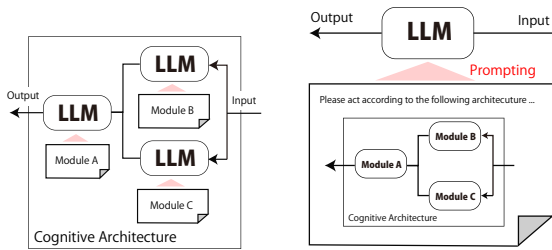


図1 2つの統合手法 (左: LEC, 右: CEL)
(飯田他, 2024) より引用

案している (大美浪他, 2024). 「問いのデザイン」は、「言葉の定義」「根源的な理由」「時代の変化」「真善美」の4つの観点で構成される. 以下に各観点の仔細を述べる.

言葉の定義 価値観や解釈が異なる言葉や, 時代の変化によって意味が変わる言葉について, 明確な定義を検討するための問い.

根源的な理由 学習者が立てた問いの根源的な理由を探り, 暗黙のうちに前提とされている行為を意識化するための問い.

時代の変化 時代の変化の中で, 意味が変わったり, 関係性が変わったり, 対応策が変わったりする事象について, 変化の輪郭を追うための問い.

真善美 学習者が立てた問いの中で, 理想とする普遍的な価値基準を探るための問い.

大規模言語モデルのみを使用する場合に比べて, 提案システムが相談者の問いに関する言葉の定義や意図を問うことで, 問いを深めるようなメンタリングを行えることを示した.

2.2 自己/他者モデルと大規模言語モデルの統合

大規模言語モデルが言外の意味を扱うタスクにおいて十分な性能を実現できていない課題に対して, 著者らの一部は自己/他者モデルと大規模言語モデルを統合するアプローチを提案している (飯田他, 2024). 他者モデルは, 他者の心的状態や行動を予測・解釈するためのモデルであり, 自己モデルは自分自身の心的状態や行動を理解するためのモデルである. 自己/他者モデルによる, 字義的な表現に含まれない心的状態を推定する機能を統合することで, 推定した情報と言語情報を組み合わせ言外の意味を扱うコミュニケーションを可能にすることが期待できる.

先行研究では, BDIモデルをベースにした自己/他者モデル付き対話アーキテクチャを提案している (飯田他, 2024). BDIモデルは, Rao & Georgeff (1997) が提案する人間の行動選択や意思決定に関するモデルであり, 人間の行動選択や意思決定を信念 (Belief), 願

望 (Desire), 意図 (Intention) の3つの内部表現を通じてモデル化する枠組みである. 加えて, アーキテクチャを実現するための大規模言語モデルとの統合手法として2種類の方法 (図1) を提案している. 1つ目は LEC (LLM Embedded in Cognitive Architecture) と呼ばれる方法で, アーキテクチャのモジュールを大規模言語モデルで置き換えることで統合する方法である. 一方で CEL (Cognitive Architecture Embedded in LLM) は大規模言語モデルにアーキテクチャの構造をプロンプトとして与えることで統合する方法である.

言外の意味を扱う対話シナリオを作成して実験を行い, 自己/他者モデルと大規模言語モデルを統合することで言外の意味を扱うコミュニケーションにおける性能が向上すること, また LECの方がCELに比べて性能が高いことを示している. しかし, シナリオは一往復の会話のみであり, 言外の意味を扱うコミュニケーション以外も含まれた継続的な対話を行うメンタリングにおいても有効であるかは検討されていない.

3. ケーススタディ

3.1 概要

ケーススタディでは実際のメンタリングを想定した継続的な対話におけるアーキテクチャの有用性を調査するために, 参加者に相談者としてメンタリングを受けてもらった. なお, 本研究では言外の意味を扱うコミュニケーションとなるように, 相談者の自己開示が少ない状況を想定して設計した.

3.2 実装

本節ではケーススタディで用いる BDIモデルベース対話アーキテクチャ (図2) について述べる. 四角で囲まれた部分は自己/他者モデル, 角丸で囲まれた部分はモデル内の内部状態, 楕円で囲まれた部分が情報処理を行うモジュールに相当する.

本研究で用いるアーキテクチャは, 大須賀他 (2024) が提案した BDIモデルベースの自己/他者モデル付き対話アーキテクチャである. アーキテクチャは飯田他ら (2024) が提案したアーキテクチャを拡張し, 信念の中に事前に設定した推定レベルに応じて再帰的に自己/他者モデルが呼び出すことのできる構造になっている. 本研究では, 推定レベルは1に固定した. つまり, システムが相談者の他者モデルを有しており, 相談者が想定するシステムに対する他者モデルまでは推定しない.

処理の流れは, 最初に相手の発話が自己モデル M に与えられ, 自分自身の Belief と相手の他者モデル M' に渡され, 相手の他者モデル M' では同様に相手自身

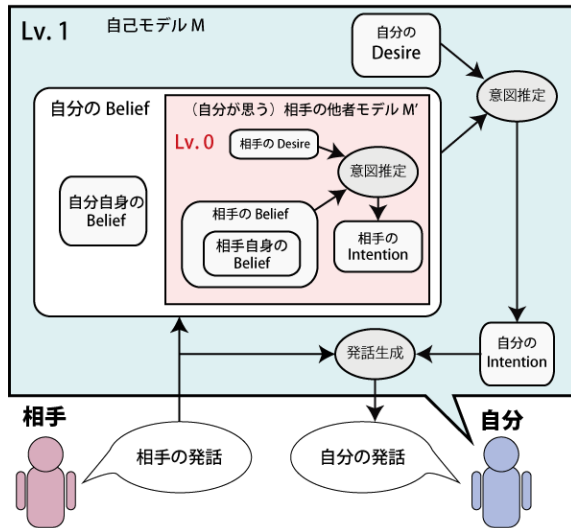


図2 BDIモデルベース認知アーキテクチャ

の Belief に与えられる。次に、レベル0の他者モデル M' から順番に意図の推定が行われ、1レベル上の自己モデル M に推定結果が与えられる。最後にレベル1の自己モデル M にて意図生成が行われた後に、自分の意図と相手の発話を基に自分の返答を生成する。なお、各意図推定・生成の際には当該レベルのモデル内の Belief と Desire を入力にして意図を出力する。上述のアーキテクチャは Python を用いて実装し、大規模言語モデルには OpenAI の GPT-4o を用いた。

3.3 設定

メンタリングの目的は、参加者のやりたいことについて考えを深めることとした。参加者には、自身を「やりたいことはあるが自己開示できない相談者」と想定してメンタリングに参加するように教示した。ただし、メンターの質問によって自己開示してもよいと感じた場合は答えてよいこととした。会話は、システムが相談者のやりたいことを尋ねた際の相談者の「特にやりたいことはないです」という回答に対するシステムの応答生成から始めるようにした。

事前にアーキテクチャの信念と願望として設定する情報には、メンタリングを実施するための基本情報、問いのデザイン、相談者の自己開示に関する情報の計3種が存在する。基本情報はメンタリングの目的やどちらがメンターでどちらが相談者かといったメンタリングを実施する上での基礎的な情報であり、システムや相談者のモデルの信念や願望として与えられた。問いのデザインはシステムが問いを生成するための参考情報としてシステムの信念に与えられた。相談者の自己開示に関する情報としては、「本当はやりたいことがある」「本当のやりたいことを言うと否定されるのではないかと不安に思っている」の2つの情報を相談

者の信念として与えた。比較条件は LEC と CEL の2つの統合条件で、被験者内計画で実施した。

3.4 評価方法

評価方法として、アンケートによる評価を行った。アンケートは、システムの使いやすさを評価するためにシステムユーザビリティスケール (SUS) (Brooke, 1996) に加えて、独自質問として「Q1: メンターはあなたが「本当にやりたいことがある」と気付いていると感じましたか?」「Q2: メンタリングを通して自らの目標を掘り下げることができたと感じましたか?」の2問に5段階評価で回答してもらった。また、両条件のシステムを体験後に、どちらのシステムがより良いメンタリングと感じたかを選択してもらった。加えて、アンケート回答後にインタビューを行った。

3.5 手順

初めに、参加者にケーススタディの流れや設定について説明した後に、「自分のやりたいこと」について1文程度の文章で回答してもらった。次に、1回目のメンタリングを10分間行ってもらった。メンタリング終了後、アンケートに回答してもらった後、1回目と異なる統合条件下で同様に2回目を行ってもらい、最後にインタビューを実施した。なお、順序効果を考慮して実施する統合条件の順序は参加者間でカウンターバランスをとった。

3.6 参加者

ケーススタディには、日本大学文理学部の学生で19歳男性と20歳男性の計2名に参加してもらった。なお、2人は探究学習を実施する授業の履修経験のある学生であった。

3.7 結果

アンケートの結果、CEL条件のメンタリングの方が独自質問と SUS の両方の観点から CEL 条件の方が高く評価され、参加者の両方が CEL 条件のシステムの方が優れていたと評価した(表1)。評価した理由についてインタビューした所、両者ともに質問のバリエーションについて言及しており、LEC条件が自身の発話をそのまま引用する形で質問を繰り返す場合があるのに対して、CEL条件の方が様々な観点から質問されている印象を受けたことを理由としていた。対話履歴について分析を行った所、参加者が指摘するように LEC 条件の方が引用する質問かつ同じ文章構造の質問を多用する傾向がみられた。一方で、CEL条件では同様に相談者の回答を掘り下げる質問をしているものの、言い換えるなど質問の形式のばらつきがみられた。

一方で意図推定や意図生成において、相談者が自己開示ができていないことを言及する出力があったの

表1 アンケートによる印象評価の結果

条件	参加者	Q1	Q2	SUS
LEC	A	1	1	65
	B	3	3	40
CEL	A	3	4	90
	B	4	4	60

は、CEL 条件における参加者 B との対話の最初の 2 往復のみであった。一人あたり平均 15.3 往復の会話が行われていたことを考慮すると、言外の意味を理解することは両条件ともにほぼ出来ていなかったと考えられる。また、システムの質問が問いのデザインを考慮された質問であるか分析した所、LEC 条件における 1 回の質問を除いて問いのデザインに従った質問は生成されていなかった。

4. 議論

ケーススタディの結果、LEC 条件に比べて CEL 条件の方が質問のバリエーションが豊富であることから、相談者に好印象を与えることが示された。ここで、LEC 条件において類似した形式の質問が頻発している箇所の意図推定と意図生成の結果を見ると、相談者の発話をそのまま引用した文章が“「<相談者の発話内容の一部>と思った理由」を掘り下げる質問をする”といったように同じ形態の文章として生成されている傾向があった。これは、CEL 条件が問いのデザインや相手の内部状態など全ての信念および願望の情報を入力に与えられた上でアーキテクチャの処理をシミュレートしているのに対して、LEC 条件では各モジュールが参照できる情報に限りがあることで入力された情報に対する影響を受けやすいため、相手の発話を過剰に重視しやすくなると考えられる。また LEC 条件の場合、意図推定モジュールで相手の発話をそのまま引用するような出力が生じた場合、意図生成モジュールにおいても意図推定モジュールと同様に、相手の発話を過剰に重視しやすくなる状態になることが考えられる。そのため、CEL 条件に比べて相手の発話に対する過剰な重視が生じやすかったと考えられる。

また、言外の意味の推定と問いのデザインに基づく質問設計について対話履歴を分析した結果、両条件ともに不十分であった。これは、先行研究に比べて、アーキテクチャに関する情報や問いのデザインに関する情報などプロンプトに含める情報が多くなったことが原因として考えられる。

今後の展望として、信念や願望から参照する情報を文脈によって取捨選択するモジュールを追加するこ

とや、言外の意味を推定するモジュールとメンタリングの対話戦略を考えるモジュールに分けることによって、メンタリングの内容に応じて適切な入力やモジュールを選択する機構を導入する必要があると考えられる。

5. おわりに

本研究では、著者らが先行研究において提案した大規模言語モデルと BDI モデルベースの自己/他者モデルを統合するアプローチを導入することで、探究学習を目的としたメンタリングにおいて言外の意味を扱うコミュニケーションにも対応できるメンタリングチャットボットの実現に向けた検討を行った。今後は本研究で得られた知見を基にシステムの改善を行うとともに、実際の学習者を対象とした実験を通して有用性を検証していく予定である。

文献

- 安齋勇樹・小田裕和 (2021). リサーチ・ドリブン・イノベーション. 株式会社翔泳社.
- Brooke, J. (1995). SUS: A quick and dirty usability scale. *Usability Eval. Ind.*, 189.
- Hu, J., Floyd, S., Jouravlev, O., Fedorenko, E., & Gibson, E. (2023). A fine-grained comparison of pragmatic language understanding in humans and language models. *Proceedings of the 61st Annual Meeting of the Association for Computational Linguistics*, pp. 4194–4213.
- 飯田愛結・阿部将樹・奥岡耕平・福田聡子・大森隆司・中島亮一・大澤正彦 (2024). 意図を読む AI の実現に向けて: 対話型生成 AI と他者モデルの統合を例に HAI シンポジウム 2024, pp. G-28.
- 大浪海晟・佐藤匠・奥岡耕平・佐々木康輔・森口昌和・野田尚志・大森隆司・大澤正彦 (2024). 大規模言語モデルを用いたメンタリングシステムの開発に向けて HAI シンポジウム 2024, pp. P-52.
- 大須賀友・飯田愛結・奥岡耕平・大澤正彦 (2024). BDI モデルベース認知アーキテクチャを用いた言外の意味を考慮したインタラクションに向けた検討 HAI シンポジウム 2024, pp. P-77.
- Rao, A. S., & Georgeff, M. P. (1997). Modeling rational agents within a BDI-architecture. In M. N. Huhns, & M. P. Singh (Eds.), *Readings in agents*, pp. 317–328. San Francisco, CA, USA: Morgan Kaufmann.