

Few-shot Dialogue Strategy Learning for Motivational Interviewing via Inductive Reasoning (/pdf?id=PTLoZZ_rRFm)

Anonymous

16 Feb 2024 ACL ARR 2024 February Blind Submission Readers: February, Paper2665 Senior Area Chairs, Paper2665 Area Chairs, Paper2665 Reviewers, Paper2665 Authors Show Revisions (/revisions?id=PTLoZZ_rRFm)

Abstract: We consider the task of building a dialogue system that can motivate users to adopt positive lifestyle changes, Motivational Interviewing (MI). Addressing such a task requires a system that could infer \textit{how} to motivate the user effectively. We propose DIIR, a framework that is capable of learning and applying conversation strategies in the form of natural language inductive rules from expert demonstrations. Automatic and human evaluation on instruction-following large language models show natural language strategies descriptions discovered by DIIR can improve active listening skills, reduce unsolicited advice, and promote more collaborative and less authoritative conversations, outperforming in-context demonstrations that are over 50 times longer.

Paper Type: short

Research Area: Dialogue and Interactive Systems

Contribution Types: NLP engineering experiment, Approaches to low-resource settings


Languages Studied: English


Revealed to Zhouhang Xie, Bodhisattwa Prasad Majumder, Mengjie Zhao, Yoshinori Maeda, Keiichi Yamada, Hiromi Wakaki, Julian McAuley

15 Feb 2024 (modified: 15 Feb 2024) ACL ARR 2024 February Submission

Authors: Zhouhang Xie (/profile?id=~Zhouhang_Xie1), Bodhisattwa Prasad Majumder (/profile?id=~Bodhisattwa_Prasad_Majumder1), Mengjie Zhao (/profile?id=~Mengjie_Zhao1), Yoshinori Maeda (/profile?id=~Yoshinori_Maeda1), Keiichi Yamada (/profile?id=~Keiichi_Yamada1), Hiromi Wakaki (/profile?id=~Hiromi_Wakaki1), Julian McAuley (/profile?id=~Julian_McAuley1)


Previous URL: /forum?id=MLLN4ww46X (/forum?id=MLLN4ww46X)

Previous PDF:  pdf (/attachment?id=8G0QCxswEB&name=previous_PDF)

Response PDF:  pdf (/attachment?id=8G0QCxswEB&name=response_PDF)

Reassignment Request Action Editor: No, I want the same action editor from our previous submission and understand that a new action editor may be assigned if the previous one is unavailable

Reassignment Request Reviewers: No, I want the same set of reviewers from our previous submission and understand that new reviewers may be assigned if any of the previous ones are unavailable

Software:  zip (/attachment?id=8G0QCxswEB&name=software)

Preprint: no

Preprint Status: We are considering releasing a non-anonymous preprint in the next two months (i.e., during the reviewing process).

Preferred Venue: ACL

Consent To Share Data: yes

Consent To Review: yes

Consent To Share Submission Details: On behalf of all authors, we agree to the terms above to share our submission details.

A1: yes

A1 Elaboration For Yes Or No: sec. 8

A2: yes

A2 Elaboration For Yes Or No: sec. 9

A3: yes

A3 Elaboration For Yes Or No: end of sec. 1

B: yes

B1: yes

B1 Elaboration For Yes Or No: sec. 3, sec. A,B,C in the Appendix

B2: yes

B2 Elaboration For Yes Or No: sec. G in the Appendix

B3: yes

B3 Elaboration For Yes Or No: sec. 8

B4: n/a

B5: n/a

B6: yes

B6 Elaboration For Yes Or No: sec. 3 - Experiments

C: yes

C1: yes

C1 Elaboration For Yes Or No: In Appendix A, B, C

C2: yes

C2 Elaboration For Yes Or No: In Appendix A, B, C

C3: yes

C3 Elaboration For Yes Or No: sec. 3 - Experiments

C4: yes

C4 Elaboration For Yes Or No: In Appendix A, B, C

D: yes

D1: yes

D1 Elaboration For Yes Or No: In Appendix A, B, C

D2: yes

D2 Elaboration For Yes Or No: In Appendix C

D3: n/a

D4: n/a

D5: n/a

E: no

E1: n/a

Add

Author-Editors Confidential Comment

Withdraw

Reply Type: Author:

16 Replies

Visible To: Hidden From:

[-] Meta Review of Paper2665 by Area Chair fPMU

ACL ARR 2024 February Paper2665 Area Chair fPMU

09 Apr 2024, 22:15 ACL ARR 2024 February Paper2665 Meta Review Readers:
Paper2665 Senior Area Chairs, Paper2665 Area Chairs, Paper2665 Authors, Paper2665
Reviewers Submitted, Program Chairs Show Revisions (/revisions?id=cLmyWW0UQEK)

Paper Summary:

The work considers the problem of designing a Dialogue system for motivational interviewing. The paper proposes to learn dialogue strategies by inferring inductive rules ("best practices") in natural language on how to respond to specific situations using demonstrations. The proposed system is evaluated using a set of metrics, motivated by the Motivational Interviewing Integrity Treatment and Motivational Interviewing skill code. The evaluation shows that inducing such rules leads to better performance in the automatic and human evaluation.

Summary Of Strengths:

This is a resubmission. The authors greatly improved the readability of the paper.

The authors propose a method to learn strategies from a small number of dialogues using LLM, which ensure the rule captures the context and gist of the example.

The authors compare their results with GPT with open-source LLMs in Appendix E. Please move this reference from the table caption in to the main text of the paper, where it will be easier for readers to find.

Summary Of Weaknesses:

The authors did not explain how the dialogues used in the evaluation were generated. Let's focus first on generating the clinician utterances. Reviewer p60D asked if the strategies are turn-wise [c2]. This question really helped me understand your work. It now seems to me that your strategies are rules to guide the generation of the next utterance based on the dialogue context. In your experiments, you learn rules/contexts from 5 high quality MI dialogues. Do you learn a separate context/rule for each clinician utterances in the 5 dialogues? Do you use 5 dialogues to give you a good coverage? When generating conversations, is each clinician utterance generated from a context/rule? Do you have any insight of how many of these context/rules are MI oriented and how many of them might just be about maintaining the dialogue?

Now that I think I understand how the clinician utterance are generated, how are the user utterances generated for your experiment?

Appendix A should be referred to from the main paper. You have your reference to Appendix E as part of a table caption. No evaluation of the rules learned.

As reviewer p60D suggested, the actual prompts used in learning the context-rules should be included in the appendix.

Overall Assessment: 3 = There are major points that may be revised

Best Paper Ae: No

Information Regarding The New ACL Policy On Deanonymized Preprints: I confirm I have read the information above about changes to the anonymity policy.

Add Author-Editors Confidential Comment

[-] A note re. review revision period and a gentle reminder

ACL ARR 2024 February Paper2665 Authors Zhouhang Xie (/profile?id=~Zhouhang_Xie1) (privately revealed to you)

29 Mar 2024, 11:58 ACL ARR 2024 February Paper2665 Official Comment Readers:
Program Chairs, Paper2665 Senior Area Chairs, Paper2665 Area Chairs, Paper2665
Reviewers, Paper2665 Reviewers Submitted, Paper2665 Authors Show Revisions
(/revisions?id=dL3uIA4c2Mm)

Comment:

Dear Reviewers,

As the discussion period comes close to its end, we wanted to thank you again for your time and effort in providing constructive feedbacks for our work. Meanwhile, we wanted to send a gentle reminder regarding our responses - please do not hesitate to contact us if you have any additional concerns and/or questions. We look forward to further

discussions with you.

Best Wishes, Authors

Add **Author-Editors Confidential Comment**

[-] General Response to Reviewers

ACL ARR 2024 February Paper2665 Authors Zhouhang Xie (/profile?id=~Zhouhang_Xie1) (privately revealed to you)

27 Mar 2024, 12:39 ACL ARR 2024 February Paper2665 Official Comment Readers:

Program Chairs, Paper2665 Senior Area Chairs, Paper2665 Area Chairs, Paper2665

Reviewers, Paper2665 Reviewers Submitted, Paper2665 Authors Show Revisions

(/revisions?id=taiqTzzxVd)

Comment:

We appreciate your time, effort, and valuable feedback. In this section we summarize main strengths and weaknesses mentioned by the reviewers, and our response.

Strength:

- This work studies a novel and non-trivial problem relevant to the *ACL community (LJA3). Reviewers agree that our proposed method is a strength (p6oD, LJA3, X9eg), mentioning our method has low inference-time overhead (x9eg), is training free and requires only a few examples (p6oD), explores the value of explanation rules for dialogue control (LJA3), and is an intuitive approach to gather meaningful dialogue strategies (x9eg). Reviewers also praised that our proposed automatic evaluation is derived from domain knowledge of MI (LJA3, X9eg).

Weakness and Comments/Suggestions/Typos:

- Two of the reviewers (p6oD, x9eg) asked about details of human evaluation - we will add the relevant information to the evaluation section/appendix (more details in the individual responses, as slightly different questions were asked). To sum up, we have 37 groups of human evaluations between each pair of methods compared, with 4 dialogue contexts in each group. The “additional wins” we report is the “number of wins by our method - number of wins by Opponent”.
- Two of the reviewers (p6oD, x93G) asked about open source models. However, the angles are different. We address the points along with other individual questions in individual responses to reviewers.
 - In summary, we currently have the results of Mistral-7b in appendix E (and points to these results/discussions under Table 1), we will add another pointer in the “Models and Baselines” paragraph to make this information more easily findable by readers.
 - Meanwhile, we note that DIIR is a LLM-reasoning-based framework, and it is common for LLM-reasoning-based frameworks to require strong underlying LLMs [1-8, see response to reviewer p6oD for details]. We note that even though our method is less impactful on the open source model, it is still comparable to several other ICL baselines, while having the benefit of producing interpretable dialogue strategies.
- One reviewer (LJA3) mentioned that it is unsure if our framework generalizes to other domains. We address this in individual responses to reviewers. To sum up, we agree that it would be interesting to generalize DIIR to other domains, but we wanted to clarify that our primary field of study in this work is motivational interviewing, and to this end, a general DIIR-like framework for various dialogue domains seems more aligned with potential future directions than weakness of the current work.
- We address other comments and additional details (such as the prompts we used for LLMs) in individual responses.

References

[1] Prompt-Based Monte-Carlo Tree Search for Goal-oriented Dialogue Policy Planning

(<https://aclanthology.org/2023.emnlp-main.439.pdf>), EMNLP 2023

[2] PromptAgent: Strategic Planning with Language Models Enables Expert-level Prompt Optimization

(<https://arxiv.org/abs/2310.16427>), ICLR 2024

[3] Do Embodied Agents Dream of Pixelated Sheep?: Embodied Decision Making using Language Guided World Modelling (<https://arxiv.org/pdf/2301.12050.pdf>), ICML 2023

[4] Phenomenal Yet Puzzling: Testing Inductive Reasoning Capabilities of Language Models with Hypothesis Refinement (<https://arxiv.org/pdf/2310.08559.pdf>), ICLR 2024

[5] Plan, Verify and Switch: Integrated Reasoning with Diverse X-of-Thoughts (<https://aclanthology.org/2023.emnlp-main.169.pdf>), EMNLP 2023

[6] Hypothesis Search: Inductive Reasoning with Language Models (<https://arxiv.org/abs/2309.05660>), ICLR 2024

[7] Deductive Verification of Chain-of-Thought Reasoning (https://proceedings.neurips.cc/paper_files/paper/2023/file/f6b22ac37beb5da61efd4882082c9ecd-Paper-Conference.pdf), Neurips 2023

[8] Large Language Models Are Semi-Parametric Reinforcement Learning Agents (https://proceedings.neurips.cc/paper_files/paper/2023/file/f6b22ac37beb5da61efd4882082c9ecd-Paper-Conference.pdf), Neurips 2023

Add

Author-Editors Confidential Comment

[-] Official Review of Paper2665 by Reviewer p6oD

ACL ARR 2024 February Paper2665 Reviewer p6oD

21 Mar 2024, 04:15 ACL ARR 2024 February Paper2665 Official Review Readers:

Program Chairs, Paper2665 Senior Area Chairs, Paper2665 Area Chairs, Paper2665

Reviewers Submitted, Paper2665 Authors Show Revisions (/revisions?

id=407HGYE0huz)

Recommended Process Of Reviewing: I have read the instructions above

Paper Summary:

The paper proposes an approach to build Motivation Interviewing conversational agent. Using LLMs, they propose to assess the dialogue state, then generate (when learning) or retrieve (during inference) the best strategy to respond to users inputs in a given situation. It aims to increase the success rate of users' motivation by avoiding undesired or misplaced responses.

Summary Of Strengths:

1. No training is required and only few examples are needed
2. Strategies are generated by an LLM, which may be more suitable (detailed) to instruct LLMs for action.
3. An iterative process with discriminator, executor and generator to mimic/replace interactive environment that needs human intervention

Summary Of Weaknesses:

1. The method is shown to work well with undisclosed proprietary LLMs with more capabilities and to have less impact on open-source LLMs, which can be restrictive.

Comments, Suggestions And Typos:

1. Please add some details on the different LLMs instances (generators G, executor E and Discriminator). For instance their prompts. It is stated what their purpose is but not how they achieve it.
2. L171 "We repeat the process introduced above for all context-response pairs available" Does this mean that strategies are turnwise? Are previous turns strategies taken into account? As one may think that, in real-world conversations this could have an impact?
3. Could you provide precise figures from the dataset: how many high quality are retained? Consequently, how many dialogues "the rest" in L220 remain for leaning after removing 80 evaluation dialogues?
4. For Human evaluation how many tuckers? And how many total "battles" between each model (only the number of additional wins is provided)

Soundness: 3 = Acceptable: This study provides sufficient support for its major claims/arguments. Some minor points may need extra support or details.

Overall Assessment: 2 = Revisions Needed: This paper has some merit, but also significant flaws, and needs work before it would be of interest to the community.

Confidence: 4 = Quite sure. I tried to check the important points carefully. It's unlikely, though conceivable, that I missed something that should affect my ratings.

Best Paper: No

Limitations And Societal Impact:

Discussion on limitations is a bit shortsighted. Societal impact of such technology can be larger than the virtuous angle sold in the paper. And potential misuses are numerous and rather dangerous (changing users' convictions, misguidings, etc).

Ethical Concerns:

Impact of Motivational Interview, and its potential misuse, are briefly reviewed on the ethic side in the paper.

Needs Ethics Review: No

Reproducibility: 2 = They would be hard pressed to reproduce the results: The contribution depends on data that are simply not available outside the author's institution or consortium and/or not enough details are provided.

Datasets: 2 = Documentary: The new datasets will be useful to study or replicate the reported research, although for other purposes they may have limited interest or limited usability. (Still a positive rating)

Software: 2 = Documentary: The new software will be useful to study or replicate the reported research, although for other purposes it may have limited interest or limited usability. (Still a positive rating)

Knowledge Of Or Educated Guess At Author Identity: No

Knowledge Of Paper: N/A, I do not know anything about the paper from outside sources

Knowledge Of Paper Source: N/A, I do not know anything about the paper from outside sources

Impact Of Knowledge Of Paper: N/A, I do not know anything about the paper from outside sources

Reviewer Certification: p6oD

Add **Author-Editors Confidential Comment**

[-] Response to Reviewer p6oD (3/3) - Details of Prompts We used for LLMs

ACL ARR 2024 February Paper2665 Authors Zhouhang Xie (/profile?id=~Zhouhang_Xie1) (privately revealed to you)

27 Mar 2024, 12:59 (modified: 27 Mar 2024, 13:00) ACL ARR 2024 February
Paper2665 Official Comment Readers: Program Chairs, Paper2665 Senior Area
Chairs, Paper2665 Area Chairs, Paper2665 Reviewers Submitted, Paper2665
Authors Show Revisions (/revisions?id=OWz18IM5wqO)

Comment:

We document the prompts used for reasoning in this work as follows. All placeholders are wrapped in "<" and ">" marks. For example, dialogue context placeholder is denoted as "<Dialogue Context>"

Instruction for Generating Situation Description k

We ask the LLM to describe the situation with the following prompt. For a better inductive bias, we ground the generation in the scope of stages of change model (Prochaska 174 and Velicer, 1997) to help the model better infer user mental states by pasting a description of the model.

You are a dialogue analyst and your job is to help us understanding motivational interviewing dialogues. You will be given a dialogue context, and you will help us determine which of the 4 stages of motivational interviewing the therapist is at: engaging, focusing, evoking, or planning.

1. **Engaging**: This is the first process, where the goal is to establish a trusting and respectful relationship between the therapist and the client. Engaging involves active listening and expressing empathy to create a comfortable environment that encourages the client to open up and talk about their experiences and issues. Effective engagement sets the stage for the work to come.

2. **Focusing**: During the focusing process, the therapist helps the client to determine the direction of the conversation and identify what changes are important and possible to work on. It involves clarifying the agenda and honing in on specific areas that may benefit from change. Setting goals and priorities is a key part of the focusing stage.

3. **Evoking**: In this stage, the therapist's role is to elicit the client's own motivations for change. The client is encouraged to talk about their desires, abilities, reasons, and need for change (known as DARN). The therapist uses reflective listening and open-ended questions to draw out the client's personal reasons for change, their understanding of the issue, and any ambivalence they may feel.

4. **Planning**: The final process involves developing a commitment to change and formulating a concrete plan of action. The therapist collaborates with the client to create strategies and steps to initiate and sustain change. This stage includes setting goals, considering options, discussing the pros and cons of different strategies, and planning for potential obstacles.

Look at the following dialogue snippet, which of the 4 stages is the dialogue in?
<Dialogue Context>

Format your answer in this format: {'prediction': "your answer"}, you do not have to explain anything.

Instruction for the Generator LLM

We use the following prompt for the Generator LLM.

You are trying to teach a student to follow true therapist's motivational interviewing behavior. Here is the current scenario: <Dialogue Context>
Client mental state seems to be: <Situation Description>
The Student Response is: <Executor Generated Response>
In comparison, the true therapist response is: <Gold Response>

From our annotation, it seems like the true therapist's actions in order, sentence by sentence, are: <Gold Response>, which is x sentences.
Analyze the current situation, and write a instruction for the student, in the format of "Based on the annotation, When the client ..., the therapist should ..., the therapist should not ..."
When mentioning what the therapist should do, be sure to include information on how many sentences are needed and what each sentence should do.
Important: this is not a general guideline, but should be specifically tailored to the flaw in the student response. Be general, make sure your rule is generalizable across topics. For example, simple use 'bad habit' instead of drug abuse/alcohol issues/smoking. such that we can reuse this rule for other topics in the future.

Instruction for the Discriminator LLM

For the ease of implementation, we simply ask the generator LM an additional question, checking whether the executor's response is good enough (however in principle this could be a separate LLM)

<...Omitted the portion that is a copy of the generator's prompt>
The student wrote a response based on your rule. <Executor Response>, did the student correctly follow your guideline and replicated the true therapist? Answer yes or no first.

Instruction for the Executor LLM

We ask the Executor LLM to conduct dialogue completion with the below instruction, following recent work's suggestion [2].

Look at the following therapist-client dialogue, predict what the therapist should say next.
<Dialogue State>
Follow these guidelines when producing response:
<Dialogue Strategy Description>
Start your response with [therapist]:

References:

- [1] The transtheoretical model of health behavior change, American Journal of Health Promotion, 1997
- [2] Is this the real life? Is this just fantasy? The Misleading Success of Simulating Social Interactions With LLMs, Arxiv Preprint, 2024

Add **Author-Editors Confidential Comment**

[-] **Response to Reviewer p6oD (2/3)**

ACL ARR 2024 February Paper2665 Authors Zhouhang Xie (/profile?id=~Zhouhang_Xie1) (privately revealed to you)

Comment:

We cover comments and suggestions in this response block.

- [C1] Please add some details on the different LLMs instances (generators G, executor E and Discriminator). For instance their prompts. It is stated what their purpose is but not how they achieve it.
 - Thanks for the suggestion, we will add the actual prompts we used and consolidate them with relevant discussions into a single appendix section.
 - We present the planned appendix section that will be added at the end of this response (it would take a bit too much space to paste it here).
- [C2] L171 “We repeat the process introduced above for all context-response pairs available” Does this mean that strategies are turnwise? Are previous turns strategies taken into account? As one may think that, in real-world conversations this could have an impact?
 - Yes, the strategies are turnwise. We agree that there could be room for improvement by integrating previous turn strategies.
- [C3] Could you provide precise figures from the dataset: how many high quality are retained? Consequently, how many dialogues “the rest” in L220 remain for learning after removing 80 evaluation dialogues?
 - Thanks, there are 110 high quality dialogues with 23 low quality dialogues in the original dataset. Among the high quality dialogues, we hold out 80 dialogues for evaluation; this leaves 30 dialogues. We do not use all these 30 dialogues for learning, but randomly sample 5 dialogues to generate the natural language rules. We will add this description to the dataset section.
- [C4] For Human evaluation how many tuckers? And how many total “battles” between each model (only the number of additional wins is provided)
 - Thanks for bringing this up, we will add the information below to section C.3 of the appendix (Human Evaluation Details Section).
 - (How many battles...) There are 37 “battles” between each pair of models. Each “battle” requires the Turkers to read both model’s responses given 4 dialogue contexts. (This adds up to 148 dialogue contexts - which is from the 10 dialogues we mentioned that are reserved for evaluation at L220+).
 - (How many turkers...) Each “battle” is rated by one individual Turker.
- [Limitations And Societal Impact] Discussion on limitations is a bit shortsighted. Societal impact of such technology can be larger than the virtuous angle sold in the paper. And potential misuses are numerous and rather dangerous (changing users' convictions, misguidings, etc).
 - Thank you for bringing up this very important problem. We agree (and do not want to downplay) that the persuasive nature of MI can impose potential harm. However, we wanted to clarify that MI techniques have been originally developed for and long been used for benefiting users in domains such as healthcare [9].
 - For this reason, we mentioned in the Ethical Concern section that “system deployers should carefully analyze the specific topic for Motivational Interviewing to eliminate risks of harm” - this covers cases such as misguiding and changing users’ convictions. We will incorporate concrete examples to make this clear: “system deployers should carefully analyze the specific topic for Motivational Interviewing to eliminate risks of harm, such as potentially changing user’s convictions or misguiding the user.” We are happy to hear your suggestions as to how the ethical concern section could be improved.

Add **Author-Editors Confidential Comment**

[–] Response to Reviewer p6oD (1/3)

ACL ARR 2024 February Paper2665 Authors Zhouhang Xie (/profile?id=~Zhouhang_Xie1) (privately revealed to you)

27 Mar 2024, 12:51 (modified: 27 Mar 2024, 13:00) ACL ARR 2024 February
Paper2665 Official Comment Readers: Program Chairs, Paper2665 Senior Area
Chairs, Paper2665 Area Chairs, Paper2665 Reviewers Submitted, Paper2665
Authors Show Revisions (/revisions?id=bTWMG7EPiM)

Comment:

Thank you for providing valuable feedback for our work. We address the weakness and comments mentioned above as follows:

- [W1] The method is shown to work well with undisclosed proprietary LLMs with more capabilities and to have less impact on open-source LLMs, which can be restrictive.
 - Thanks for pointing this out - we agree that it would be better if the frameworks can work well for open-source LLMs. However, we wanted to note that DIIR is a LLM-reasoning-based framework, which commonly requires a strong underlying LLM. We provide an overview of various recent relevant works in the Table below (see end of this response block, "Overview for LLMs Studied in Recent LLM-reasoning-based Frameworks") - they either require a strong proprietary model, or find open-source LLMs less performant, similar to our findings. Meanwhile, we note that even though our method is less impactful on the open source model, it is still comparable to several other ICL baselines, while having the benefit of producing interpretable instructions like the ones presented in Table 9.

Table: Overview for LLMs Studied in Recent LLM-reasoning-based Frameworks

Title	Conference+Year	Models
Prompt-Based Monte-Carlo Tree Search for Goal-oriented Dialogue Policy Planning	EMNLP 2023	GPT3.5
PromptAgent: Strategic Planning with Language Models Enables Expert-level Prompt Optimization	ICLR 2024	GPT3.5, GPT4
Do Embodied Agents Dream of Pixelated Sheep?: Embodied Decision Making using Language Guided World Modelling	ICML 2023	OpenAI Codex
Phenomenal Yet Puzzling: Testing Inductive Reasoning Capabilities of Language Models with Hypothesis Refinement	ICLR 2024	primarily GPT-4, and find GPT-3.5, llama2-70b, claude less performant
Plan, Verify and Switch: Integrated Reasoning with Diverse X-of-Thoughts	EMNLP 2023	GPT3.5
Hypothesis Search: Inductive Reasoning with Language Models	ICLR 2024	GPT3.5, GPT4
Deductive Verification of Chain-of-Thought Reasoning	Neurips 2023	primarily GPT3.5, and find llama less performant
Large Language Models Are Semi-Parametric Reinforcement Learning Agents	Neurips 2023	GPT3.5, text-davinci-003

Add **Author-Editors Confidential Comment**

[-] Responding to rebuttal

ACL ARR 2024 February Paper2665 Reviewer p6oD

02 Apr 2024, 01:25 ACL ARR 2024 February Paper2665 Official
Comment Readers: Program Chairs, Paper2665 Senior Area Chairs,
Paper2665 Area Chairs, Paper2665 Reviewers Submitted, Paper2665
Authors Show Revisions (/revisions?id=7wAMie3pIjW)

Comment:

Thank you for the very (too?) thorough answer. All the information brought in during the discussion here could have a very beneficial effect on the paper's revision process. So I maintain my score for now.

Add

Author-Editors Confidential Comment**[-] Further author response, clarifications, and a thank-you note**

ACL ARR 2024 February Paper2665 Authors Zhouhang Xie (/profile?id=~Zhouhang_Xie1) (privately revealed to you)

02 Apr 2024, 15:50 (modified: 02 Apr 2024, 16:01) ACL ARR 2024

February Paper2665 Official Comment Readers: Program Chairs,

Paper2665 Senior Area Chairs, Paper2665 Area Chairs, Paper2665

Reviewers Submitted, Paper2665 Authors Show Revisions (/revisions?

id=M3dGRI7z7F9)

Comment:

Dear Reviewer,

We want to thank you again for your time and constructive feedback on our work.

In the meantime, we wanted to clarify that although we try to cover as much details as possible during the response period for clearer communication in fewer turns, many aspects discussed are clarifying discussions, while others can be easily incorporated via minor updates.

Specifically, among the points discussed:

Re weakness:

- [W1] - We will add the additional references to the paper (adding these references to the sentence "We found open-source LLMs cannot consistently follow self-generated statements, coherent with recent findings (Qiu et al., 2024), and thus are less performant (See Appendix E)." under Table 1)

Re comments/questions/suggestions:

- [C2] and [Limitations And Societal Impact] are clarifying discussions, although we wanted to thank you again for bringing these points up to help us think about ways to further improve/update our work as we mentioned above.
- For [C1], the requested information is specific prompts we used, and we originally stated that the code will be open-sourced and thus readers would indeed be able to access these prompts even without the to-be-added prompts.
- The remaining details mentioned in [C3], [C4] can be added to the work with minor updates (i.e. two sentences, one at the dataset section, another at human evaluation section). We agree that it is a good idea to include the additional details discussed. However, we wanted to note that the dataset we used (AnnoMI) and its statistics is open-sourced by prior work (cited in the paper), and we already mentioned that we used 148 dialogues for evaluation (L222) in the meantime.

Please don't hesitate to contact us if you have any further questions/concerns.

Best wishes, Authors

Add

Author-Editors Confidential Comment**[-] Official Review of Paper2665 by Reviewer LJA3**

ACL ARR 2024 February Paper2665 Reviewer LJA3

Recommended Process Of Reviewing: I have read the instructions above

Paper Summary:

The work considers the problem of designing a Dialogue system for motivational interviewing. The paper proposes to learn dialogue strategies by inferring inductive rules ("best practices") in natural language on how to respond to specific situations using demonstrations. The proposed system is evaluated using a set of metrics, motivated by the Motivational Interviewing Integrity Treatment and Motivational Interviewing skill code. The evaluation shows that inducing such rules leads to better performance in the automatic and human evaluation.

Summary Of Strengths:

- the paper addresses a novel and non-trivial problem that is of interest to *ACL community
- the paper motivates a series of automatic evaluation metrics, based on evaluation protocols with experts trained to do this tasks
- the experiments performed include both automatic and human evaluation on the overall value of the system produced
- the paper explores an "inductive bias" on how the explanation rules are inherently valuable not only for the model prediction but as well as part of the dialogue control.

Summary Of Weaknesses:

- there is a question on the wider applicability of the method proposed (DIIR) to other domains.

Comments, Suggestions And Typos:

- from the abstract, I would suggest removing DIIR, and just using the word "framework" or "model" as you do not introduce what it means in the abstract. Not explicitly mentioning it is not necessary.
- line paragraph Line 127-138 you are using terms like "discriminator LLM D" and "generator LLM E". I would suggest dropping the usage of the term LLM, as 1) the framework is quite general and actually does not depend on the underlying discriminator/generator to be LLM: it could be VLM, or any other model supporting desired behaviour.
- For me, a sentence in lines 174-177 is a bit unnecessary, as it seems you did not have space to properly describe the relationship (which might be done in more extensive related work). Yet, at the same time, I find Adversarial learning and reference games quite distance topics to discuss to begin with.

Soundness: 4 = Strong: This study provides sufficient support for all of its claims/arguments. Some extra experiments could be nice, but not essential.

Overall Assessment: 3 = Good: This paper makes a reasonable contribution, and might be of interest for some (broad or narrow) sub-communities, possibly with minor revisions.

Confidence: 3 = Pretty sure, but there's a chance I missed something. Although I have a good feel for this area in general, I did not carefully check the paper's details, e.g., the math or experimental design.

Best Paper: No

Ethical Concerns:

No Ethical Considerations

Needs Ethics Review: No

Reproducibility: 4 = They could mostly reproduce the results, but there may be some variation because of sample variance or minor variations in their interpretation of the protocol or method.

Datasets: 1 = No usable datasets submitted.

Software: 4 = Useful: I would recommend the new software to other researchers or developers for their ongoing work.

Knowledge Of Or Educated Guess At Author Identity: No

Knowledge Of Paper: N/A, I do not know anything about the paper from outside sources

Knowledge Of Paper Source: N/A, I do not know anything about the paper from outside sources

Impact Of Knowledge Of Paper: N/A, I do not know anything about the paper from outside sources

Reviewer Certification: LJA3

Add

Author-Editors Confidential Comment

[-] **Response to Reviewer LJA3 (1/1)**

ACL ARR 2024 February Paper2665 Authors Zhouhang Xie (/profile?id=~Zhouhang_Xie1) (privately revealed to you)

27 Mar 2024, 13:04 (modified: 27 Mar 2024, 13:05) ACL ARR 2024 February Paper2665 Official Comment Readers: Program Chairs, Paper2665 Senior Area Chairs, Paper2665 Area Chairs, Paper2665 Reviewers Submitted, Paper2665 Authors Show Revisions (/revisions?id=C9NY9_DF2a)

Comment:

Thank you for your time valuable insights on our work. Regarding the weakness and comments mentioned, please find our response as follows:

Re. Weakness:

- [W1] There is a question on the wider applicability of the method proposed (DIIR) to other domains.
 - Thank you for pointing out this insightful direction. We agree that it would be interesting to generalize DIIR to other domains. In the meantime, however, we wanted to clarify that the contributions we claim in this work (and for DIIR) are within the domain of MI (for example, we are the first work to build and evaluate a MI dialogue system that automatically learns from data, while prior works focus on manually engineering response logic). To this end, a general DIIR-like framework that works for various domains sounds more aligned with future work than a desiderata for the scope of the current work.

Re. Comments, Suggestions And Typos:

- [C1] from the abstract, I would suggest removing DIIR, and just using the word "framework" or "model" as you do not introduce what it means in the abstract. Not explicitly mentioning it is not necessary.
 - Thanks, that is a good idea and we will update accordingly.
- [C2] line paragraph Line 127-138 you are using terms like "discriminator LLM D" and "generator LLM E". I would suggest dropping the usage of the term LLM, as 1) the framework is quite general and actually does not depend on the underlying discriminator/generator to be LLM: it could be VLM, or any other model supporting desired behavior.
 - Thanks for pointing this out - we updated accordingly, removing "LLM" to stay general.
- [C3] For me, a sentence in lines 174-177 is a bit unnecessary, as it seems you did not have space to properly describe the relationship (which might be done in more extensive related work). Yet, at the same time, I find Adversarial learning and reference games quite distance topics to discuss to begin with
 - Thanks for the suggestion - and indeed, the relationship between DIIR to adversarial learning and reference games more on inspiration level. We will remove this to smoothen the flow of the paper.

Add **Author-Editors Confidential Comment**

[-] **Official Review of Paper2665 by Reviewer x9eg**

ACL ARR 2024 February Paper2665 Reviewer x9eg

14 Mar 2024, 05:21 (modified: 31 Mar 2024, 11:03) ACL ARR 2024 February Paper2665 Official Review Readers: Program Chairs, Paper2665 Senior Area Chairs, Paper2665 Area Chairs, Paper2665 Reviewers Submitted, Paper2665 Authors Show Revisions (/revisions?id=RRW87xUTcn)

Recommended Process Of Reviewing: I have read the instructions above

Paper Summary:

The paper proposes a method to improve LLMs for motivational interviewing. The method generates a set of strategies for training examples. At inference time, a strategy for the most similar problem to the one at hand is used to condition the model for prediction. This results in improved outputs. The method does not require gradient-based updates and therefore can be used to improve closed models, such as, those only exposed by an API.

Summary Of Strengths:

- The method is intuitive and an interesting approach to gathering meaningful strategies
- The method is lightweight at inference time
- The method provides improvements
- Evaluation is done using MI-based criteria
- The current version of the paper is much easier to read and clearer than the previous version!

Summary Of Weaknesses:

- The human evaluation is not very clear to me. Could the authors maybe explain what is meant by "additional win" in the caption of Table 2?
- It would maybe be nice to see some models being evaluated that are not GPT-3.5 or GPT-4, for example, an open-source LLM like Mistral 7B to also see if such smaller models benefit similarly

Comments, Suggestions And Typos:

I have some smaller questions / suggestions:

- Will the data be released?
- Could you clarify whether any k's are grouped together or are they always distinct? Intuitively, there'll be many similar situations in such a set-up I think.

Typos:

- L47: hypothesis -> hypotheses
- L54: use -> uses
- L69 require -> requires
- L134: an generator -> a generator
- L206 vanillan -> vallia
- L299: focuses -> focus

Soundness: 4 = Strong: This study provides sufficient support for all of its claims/arguments. Some extra experiments could be nice, but not essential.

Overall Assessment: 4 = This paper represents solid work, and is of significant interest for the (broad or narrow) sub-communities that might build on it.

Confidence: 4 = Quite sure. I tried to check the important points carefully. It's unlikely, though conceivable, that I missed something that should affect my ratings.

Best Paper: No

Ethical Concerns:

None

Needs Ethics Review: No

Reproducibility: 4 = They could mostly reproduce the results, but there may be some variation because of sample variance or minor variations in their interpretation of the protocol or method.

Datasets: 1 = No usable datasets submitted.

Software: 1 = No usable software released.

Knowledge Of Or Educated Guess At Author Identity: No

Knowledge Of Paper: N/A, I do not know anything about the paper from outside sources

Knowledge Of Paper Source: N/A, I do not know anything about the paper from outside sources

Impact Of Knowledge Of Paper: N/A, I do not know anything about the paper from outside sources

Reviewer Certification: x9eg

Add

Author-Editors Confidential Comment

[-] **Response to Reviewer x9eg (1/1)**

ACL ARR 2024 February Paper2665 Authors Zhouhang Xie (/profile?id=~Zhouhang_Xie1) (privately revealed to you)

Comment:

Thank you for your effort in reviewing this work (for consecutive rounds) - we are glad we can improve the previous main concern re. writing. We provide responses to the weaknesses/comments as follows:

Re. Weakness:

- [W1] The human evaluation is not very clear to me. Could the authors maybe explain what is meant by "additional win" in the caption of Table 2?
 - For human evaluation, we break the 148 evaluation context-response pairs into groups of 4, resulting in $148/4=37$ groups. For each pair of methods, we let a Turker read responses generated from both methods and select the better model for each group; this results in 37 "battles" between models.
 - We report the additional wins as "DIIR's winning counts minus the opponent's winning counts". For example, if DIIR wins 29 times and the Opponent wins 8 times, the additional counts will be $29-8=21$. We will add a clarifying sentence to the human evaluation description paragraph.
- [W2] It would maybe be nice to see some models being evaluated that are not GPT-3.5 or GPT-4, for example, an open-source LLM like Mistral 7B to also see if such smaller models benefit similarly
 - Thank you for pointing this out. We originally included results and discussion using Mistral to the Appendix (section E), and included a general pointer under the main evaluation Table 1 ["...We found open-source LLMs cannot consistently follow self-generated statements, coherent with recent findings (Qiu et al., 2024), and thus are less performant (See Appendix E)."].
 - Our initial thought is that putting a pointer to these results under Table 1 might be better (more easily noticed by the readers) in case a future reader potentially wants to skim through the tables+figures. To make these results more easily noticed by the readers, we will additionally include discussions/pointers in "Models and Baselines" in the Experiment 3 section (currently line 198+).

Re. Comments, Suggestions And Typos:

- [C1] Will the data be released?
 - Yes - the original dataset is publicly available (and comes with the dialogue quality labels already, so there is no variation in how the dataset is filtered), and we will release our code and splits as well.
- [C2] Could you clarify whether any k's are grouped together or are they always distinct? Intuitively, there'll be many similar situations in such a set-up I think.
 - Currently, the k's are distinct.
 - From the evaluations we can see that just doing rule retrieval using distinct k's already yields a good improvement. We agree that there could be potential improvements by grouping the k's and performing further reasoning. Still, currently the dense retrieval step will ensure that the retrieved dialogue strategy description belongs to a relevant situation (and also the most relevant one among many similar situations, up to the limit of the retriever capability).
- [C2] Typos.
 - Thanks for the catch! We updated the writing accordingly.

Add **Author-Editors Confidential Comment**

[-] Thank you for your response!

ACL ARR 2024 February Paper2665 Reviewer x9eg

Comment:

Thanks to the authors for the response. The human evaluation is now a lot clearer to me. Sorry also for having missed the Mistral results that were already in the paper. Given that this is a short paper I don't have any further requests / concerns and will update my scores accordingly.

Add **Author-Editors Confidential Comment**

[-] Thank you for reviewing!

ACL ARR 2024 February Paper2665 Authors Zhouhang Xie (/profile?id=~Zhouhang_Xie1) (privately revealed to you)

31 Mar 2024, 12:21 ACL ARR 2024 February Paper2665 Official

Comment Readers: Program Chairs, Paper2665 Senior Area Chairs, Paper2665 Area Chairs, Paper2665 Reviewers Submitted, Paper2665 Authors Show Revisions (/revisions?id=Kg-8BhZvT7)

Comment:

We are glad that we can address your concerns - and we wanted thank you again for your time and valuable feedbacks for our work.

Bests, Authors

Add **Author-Editors Confidential Comment**

[-] Supplementary Materials by Program Chairs

ACL ARR 2024 February Program Chairs

16 Feb 2024, 13:35 ACL ARR 2024 February Paper2665 Supplementary

Materials Readers: Program Chairs, Paper2665 Reviewers, Paper2665 Authors, Paper2665 Area Chairs, Paper2665 Senior Area Chairs Show Revisions (/revisions?id=QfPGvnpftK1)

Software: [↓](#) zip (/attachment?id=QfPGvnpftK1&name=software)

Previous URL: /forum?id=MLLN4wv46X (/forum?id=MLLN4wv46X)

Previous PDF: [↓](#) pdf (/attachment?id=QfPGvnpftK1&name=previous_PDF)

Response PDF: [↓](#) pdf (/attachment?id=QfPGvnpftK1&name=response_PDF)

Reassignment Request Action Editor: No, I want the same action editor from our previous submission and understand that a new action editor may be assigned if the previous one is unavailable

Reassignment Request Reviewers: No, I want the same set of reviewers from our previous submission and understand that new reviewers may be assigned if any of the previous ones are unavailable

A1: yes

A1 Elaboration For Yes Or No: sec. 8

A2: yes

A2 Elaboration For Yes Or No: sec. 9

A3: yes

A3 Elaboration For Yes Or No: end of sec. 1

B: yes

B1: yes

B1 Elaboration For Yes Or No: sec. 3, sec. A,B,C in the Appendix

B2: yes

B2 Elaboration For Yes Or No: sec. G in the Appendix

B3: yes

B3 Elaboration For Yes Or No: sec. 8

B4: n/a

B5: n/a

B6: yes

B6 Elaboration For Yes Or No: sec. 3 - Experiments

C: yes

C1: yes

C1 Elaboration For Yes Or No: In Appendix A, B, C

C2: yes

C2 Elaboration For Yes Or No: In Appendix A, B, C

C3: yes

C3 Elaboration For Yes Or No: sec. 3 - Experiments

C4: yes

C4 Elaboration For Yes Or No: In Appendix A, B, C

D: yes

D1: yes

D1 Elaboration For Yes Or No: In Appendix A, B, C

D2: yes

D2 Elaboration For Yes Or No: In Appendix C

D3: n/a

D4: n/a

D5: n/a

E: no

E1: n/a

Note From EiCs: These are the confidential supplementary materials of the submission. If you see no entries in this comment, this means there haven't been submitted any.

Add

Author-Editors Confidential Comment

[About OpenReview \(/about\)](/about)

[Hosting a Venue \(/group?id=OpenReview.net/Support\)](/group?id=OpenReview.net/Support)

[All Venues \(/venues\)](/venues)

[Sponsors \(/sponsors\)](/sponsors)

[Frequently Asked Questions](#)

[\(https://docs.openreview.net/getting-started/frequently-asked-questions\)](https://docs.openreview.net/getting-started/frequently-asked-questions)

[Contact \(/contact\)](/contact)

[Feedback](#)

[Terms of Use \(/legal/terms\)](/legal/terms)

[Privacy Policy \(/legal/privacy\)](/legal/privacy)

[OpenReview \(/about\)](/about) is a long-term project to advance science through improved peer review, with legal nonprofit status through [Code for Science & Society \(https://codeforscience.org/\)](https://codeforscience.org/). We gratefully acknowledge the support of the [OpenReview Sponsors \(/sponsors\)](#). © 2024 OpenReview