Discovering Hidden Patterns in the Data:
The Value of Input and Output Analysis in Optimizing LLM Prompt Chains

Author: Isabel Deibel, isabel.deibel@gmail.com

Dear organizing committee and reviewers:

In this cover letter, I will address the following comments left by reviewers:

Reviewer 1:

- The paper discusses an important and relevant problem, and the details of the experiments, the prompts, and the findings of the work are mentioned in the paper. The current analysis is complete and useful. **However, it would have been good to see the paper evaluating diverse prompting techniques (other than just CoT) like Program of Thoughts, and Tree of Thoughts for distractor generator could have been included/commented about.**

Reviewer 2:

- The article is well written and recommended for publication after the suggested changes have been applied. The objectives of the paper are well set out and relevant to the educational fraternity wherein the use of fill-in-the-gap assessments play a significant role in education. The research is therefore highly relevant when considering the application of Large Language Models in developing educational and instructional content, and in this case, specifically fill-in-the-gap type assessments. The works cited in the article are relevant and appropriate. **The article has a few language errors which need to be addressed before publication. While these errors are minor and do not materially affect the article, addressing the errors will make the article easier to read and comprehend.**

Thank you for the suggestions. I have addressed the errors pointed out by reviewer 2 and I agree with reviewer 1 that there could be additional techniques that could be suitable to generate distractors. The main goal of the current paper was to document that 1) prompts may not be reliably distinguished on pure success/failure counts and 2) input and output analyses are insightful for prompting strategy, rather than to compare Chain of Thought as a technique to other techniques. While comparing Chain of Thought prompting to other prompting techniques on this topic could yield very interesting insights as well, I believe that this comparison lies outside the scope of the current paper and deserves its own analysis and write-up in the future.

Thank you again for your thoughtful suggestions.

Sincerely,

Isabel Deibel