Explanation Of Revisions PDF

We sincerely thank the reviewers for their invaluable comments and suggestions, which we have carefully considered and incorporated into the resubmission. Here are the major changes we made:

1. **Including more new models in our experiment [Reviewer fFtV, 6ZLo, 3xPr].** We tested Gemini-2.0, Claude-Haiku-3.5, Claude-Sonnet-4.0, GPT-3.5-turbo, GPT-4-turbo, GPT-4o, DeepSeek-v3, Qwen-3, and Qwen-3-reasoning, to show that our method is effective on SoTA models.

2. **Proposing a new question selection method**. We proposed an iterative algorithm to select triplets from the knowledge graph to generate questions. This method aligns with our new title, focuses more on finding the weak points of LLMs and increases our novelty.

3. **Removal of ICL [Reviewer fFtV, 3xPr]**. We removed the ICL part from our previous work.

4. **Adding new visual elements and tables [Reviewer fFtV, 6ZLo].** We used more visual elements to better express our motivation, methodology and results.