Explanation of Revisions for "Large Language Models Exhibit Limited Reasoning Ability on Coding Problems"

- Main suggestions with our initial submission was the small scope of the problems we have evaluated. In order to address the review, we have added more questions to be 750 questions in total, which is most of the LeetCode questions to date that the models got correct
- We have also improved our synonym selection mechanism to consult an LLM instead of relying on a more static method as suggested by some reviewers.