

Data and Code Appendix

The computing infrastructure for running the experiments should not affect the results of the experiment, however the program was run on a MacBook Pro M1 made in 2020 with 8 GB memory and an Apple M1 chip. The MacBook was running a macOS Monterey operating system version 12.3.

Code:

To run the code, download the files in the code and data folders. Pip install the requirements.txt file; do so in a Conda environment in order to save the packages in a directory that will not affect other programs. Then you can run the program by navigating to the code folder and executing:

```
python3 run.py --data_index 1
```

The data index to use to run the given data files are as follows:

- 0: 'SVAMP'
- 1: 'GSM8K'
- 2: 'Algebra'
- 3: 'SVAMPClean'
- 4: 'Trig300'

Before running, you will need to obtain a key from OpenAI in order to use the API, and then insert the key in the utils.py file where it says:

```
openai.api_key = "sk-***"
```

Data:

Within the data folder you can find the 5 datasets presented above and an additional .json file that includes the bad questions found within the original SVAMP dataset that were corrected when making the new SVAMPClean dataset.

Results:

In the results folder, you will find 5 .txt files with the final results of each dataset. To see all the final results, scroll to the bottom of each file. You will see a variety of information including but not limited to:

- The total questions solved (total_questions),
- The total correct questions (total_correct_questions),
- The amount of questions the symbolic solver got correct (correct_check_answers),
- The amount of times the answer passed verification the first time (estimate_passes),