LLMPhy: Complex Physical Reasoning Using Large Language Models and World Models Supplementary Materials

Anonymous authors
Paper under double-blind review

1 Appendix

In this supplementary materials, we provide additional details on our dataset and qualitative results. Please have a look at our attached supplementary zip, which provide more insights into the working of our scheme, program synthesis, the prompts we used, as well as the full output of LLMs during its optimization process with the simulator. The following details will be a guide to understand the structure of the zip file.

1. We provide sample images from our TraySim Dataset in the folder named: 'TraySim_Samples' in the suppl.zip. This includes 6 sequences from the dataset (although we provide only the first and last frame of each sequence in the zip to save space, but we provide multiple views for each image).

2. We provide the full outputs from GPT-4o, the prompts used in both phases, the intermediate results in each step of our optimization, and the trajectory/layout visualizations. These are in the folder named: 'LLMPhy-Generated-Code-Including-Prompts'.

3. In folder: 'Sample-Prompts', we provide two prompts, from phase 1 and phase 2. Note that our promots are quite long and thus difficult to be included in this pdf.

4. In folder: 'Synthesized_programs', we provide a few examples of the programs synthesized by GPT-4o for solving our task. Note the the programs are common for both phase 1 and phase 2.

We also would like to note that our dataset and code will be made publicly available upon acceptance of the paper.