Teacher LM: A Generalized Reasoning Model

Chenyang Zhao
Junior in Computer Science and Technology (CST)
Tsinghua University, Beijing, China
https://chenyangzhao.vercel.app/about

Intution

- Following the release of the GPT-3 model, researchers have demonstrated that In-Context Learning with fixed parameters can unlock the reasoning capability of large-scale language models. However, small-scale language models have difficulty using prompts to perform complex reasoning tasks with fixed parameters.
- CoT Instruction Finetuning with reasonings has also been shown to improve the reasoning ability of large-scale language models by relaxing the condition to allow parameter updates but requires a large amount of data with finetune processes to improve the reasoning ability of small-scale models.
- Even so, the models still perform poorly after using CoT Instruction Finetuning to improve their inference; even CoT Instruction Finetuning is often less effective than simply using Instruction Finetune without adding CoT.
- Therefore, we would like to propose a new language model, the Teacher LM, dedicated to generating reasoning processes that provide model-generated high-quality reasoning for existing QA tasks, replacing manually annotated CoT. Ultimately, our reasoning are used to perform CoT Instruction Finetune, improving small-scale language models' reasoning aility.

Contribution

- We proposed and open-sourced the first set of general-purpose inference language models dedicated to reasoning generation, i.e. the Teacher LM.
- We demonstrate that Finetune can unlock complex reasoning capabilities for small-scale language models on a small amount of data with model-generated CoT.
- Finally, we propose a new paradigm for training-friendly unlocking of inference power of language models based on Teacher-Student Learning and knowledge distillation.

Human Evaluation

On the left is the manually annotated reasoning and the ChatGPT-generated reasoning, while on the right is the Teacher LM-generated reasoning.

Problem

Question: What are is most likely to have a kosher deli?

(A) food court **(B)** jewish neighborhoods **(C)** jewish community **(D)** italian restaurant **(E)** jerusalem

Answer: The answer is (C).

Human

Analysis: Jewish Community means a community of Jewish people originating from the Israelites and Hebrews of historical Israel and Judah. Jewish community is most likely to have a kosher deli ... Jerusalem is the city and businesses are run by people.

流畅性：5　事实性：5　信息充分性：4

ChatGPT

Analysis: A kosher deli is a type of restaurant that serves kosher food, which is food that follows Jewish dietary laws ... However, the most likely place to find a kosher deli would be in a Jewish community or Jewish neighborhood.

流畅性：5　事实性：5　信息充分性：4

SenseModel

Analysis: Kosher food, which is food that has been prepared according to Jewish Law, is prepared and inspected for quality and safety by a certified kosher inspector.

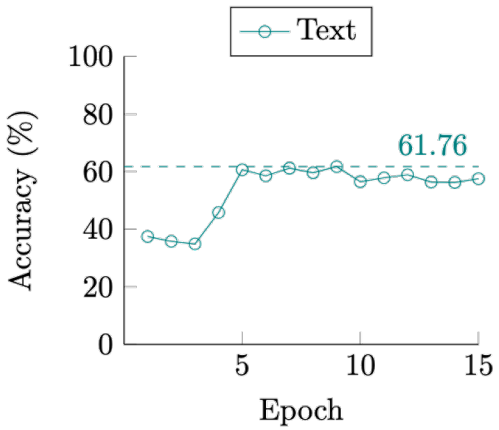
Foundamental: Various regions and states have their way of cooking according to their foods and the population of their areas ... Moreover, the differences in culture also define the differences between regions and areas.

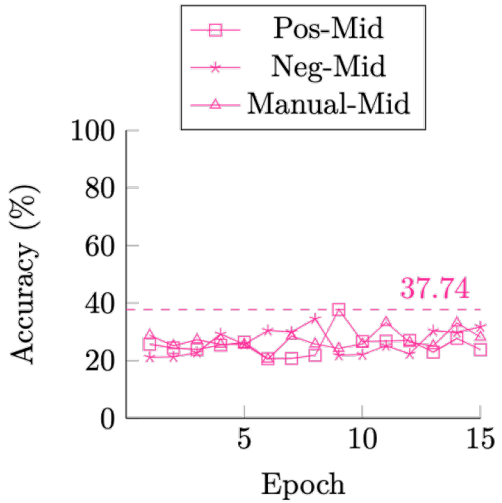
Error: Assuming that food can simply be prepared is incorrect. Because, food must be produced following certain religious rules and guidelines to be fit for consumption.

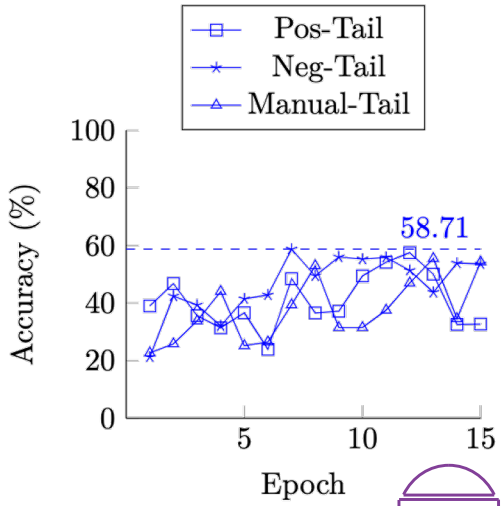
流畅性：5　事实性：5　信息充分性：5

Finetune Results

- Take the example of ECQA, a multiple-choice question dataset of moderate difficulty.
- Each sample has six fields: question, answer, five options, correct answer resolution (pos), incorrect answer resolution (neg), and overall resolution (manual), for which Teacher LM generates three fields: model CoT (cot), common errors (error), and related knowledge (fundamental).
- In our plot, the Text means only question + answer for Finetune; Manual-Mid means question + manual + answer; Manual-Tail means question + answer + manual, and so on.

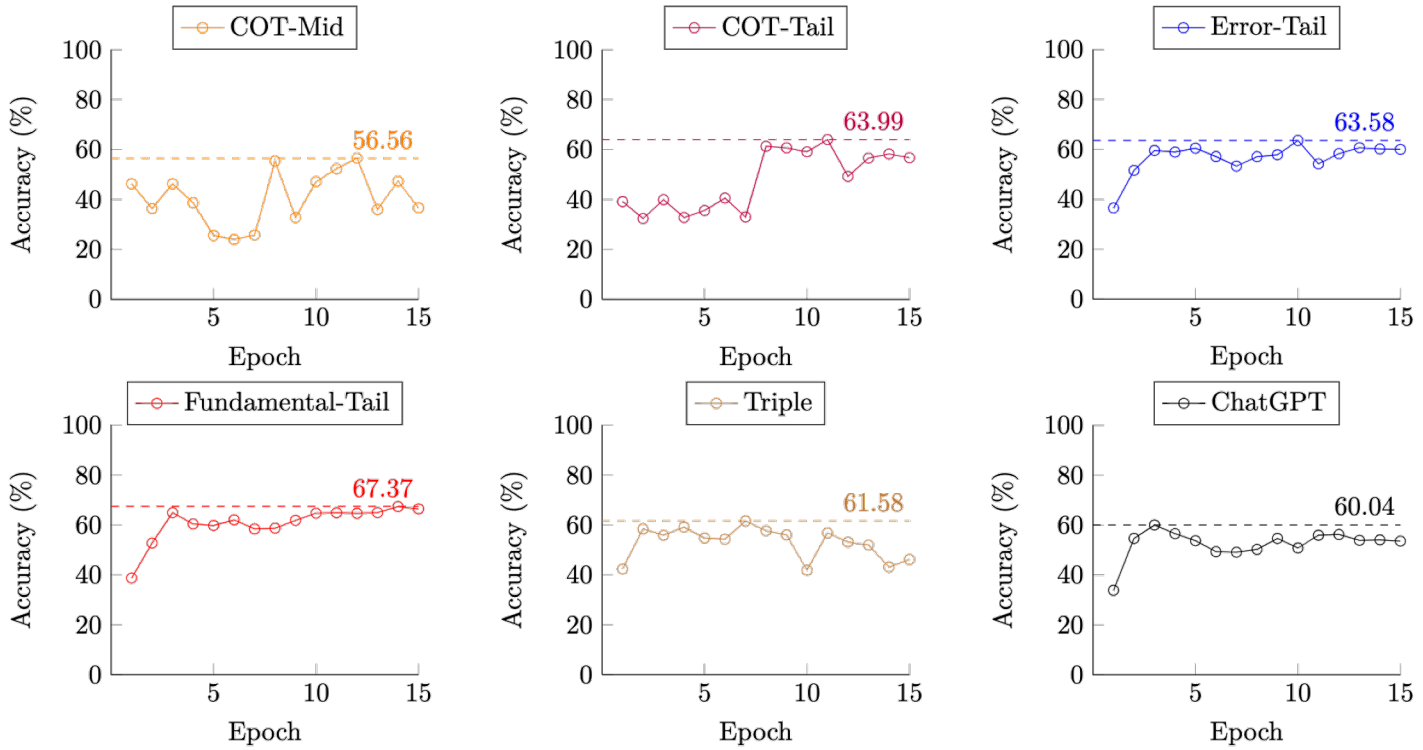






Finetune Results

- CoT-Tail means input question + answer + cot; Triple means that Error-Tail, CoT-Tail and Fundamental-Tail are directly overlapped and trained together; ChatGPT means the reasoning of CoT-Tail is replaced with that of ChatGPT.



Discussion

- Manual reasoning and 7B1 Teacher LM's reasoning are comparable in the human evaluation, but after Finetune of the small-scale model, Teacher LM's CoT is significantly better than manual reasoning.
- Tail-Training is less effective than Middle-Training combined with Self-Consistency. The model's performance could be further improved if the latter were used.
- Common Sense Task (ECQA, for example) does not reflect reasoning ability as properly as Math Word Problems (GSM8K) because the former relies on reasoning ability and the model's objective memory of relevant common sense.
- Lastly, CoT hurts are still common, probably due to the dual effect of the model's small size and the poor amount of data.

Thanks

Chenyang Zhao
Junior in Computer Science and Technology (CST)
Tsinghua University, Beijing, China
https://www.cs.tsinghua.edu.cn/csen/