

GRAD: Generative Retrieval-Aligned Demonstration Sampler for Efficient Few-Shot Reasoning

dlab

EPFL

Oussama Gabouj*, Kamel Charaf*, Ivan Zakazov*, Nicolas Baldwin, Robert West

oussama.gabouj@gmail.com, publication.charaf@gmail.com, ivan.zakazov@epfl.ch, nicolas.baldwin@epfl.ch, robert.west@epfl.ch

Manual Few-shot Prompting

- Tedious and time-consuming.
- Hard to scale across tasks.
- Requires careful prompt design.



RAG-based Few-shot Prompting

- Requires building a database.
- Database maintenance is costly.
- Hard to find relevant documents.



GRAD(i) Does the Heavy Lifting

- Provides shorter demos&outputs.
- No database / RAG needed.
- Scales across tasks and domains.



Example

Query

How much does it cost you for lunch today at Subway if you pay \$40 for a foot-long fish sub and thrice as much for a six-inch cold-cut combo sub?

RAG

Example 1

Question: The Early Bird Dinner offered dinner meals at half off the menu price if you eat between 2-4 pm. Curtis ordered the Salisbury Steak that costs \$16.00 and Rob ordered the Chicken Fried Steak at \$18.00. If they ate at 3 pm, what was the cost of their total bill?
Answer: Since Curtis and Rob ate during the Early Bird Dinner time frame, their meals were half off, making the Salisbury Steak \$8 and the Chicken Fried Steak \$9. Thus, the total cost of their meal was \$8 + \$9 = \$17. The answer is \$17. [END_DEMO]

Example 2

Question: Mike and John dined at the Taco Palace restaurant. They each ordered the Taco Grande Plate as their main meal, but Mike also ordered a side salad for \$2, a plate of cheesy fries for \$4, and a diet cola for \$2. As a result, Mike's lunch bill was twice as large as John's bill. What was the combined total cost, in dollars, of Mike and John's lunch?
Answer: Let x represent the cost of John's Taco Grande Plate and y represent the total cost of Mike's meal. [...] The combined total cost of their lunch is 8 + 16 = 24. The answer is 24 dollars. [END_DEMO]

Answer

Let the cost of the six-inch cold-cut combo sub be (x).
According to the problem, the cost of the foot-long fish sub is \$40, and the cost of the six-inch cold-cut combo sub is three times less than the foot-long fish sub. [...] Therefore, the total cost for lunch today at Subway is approximately \$53.33. However, since the exact value is required:
53.33

GRAD

Example 1

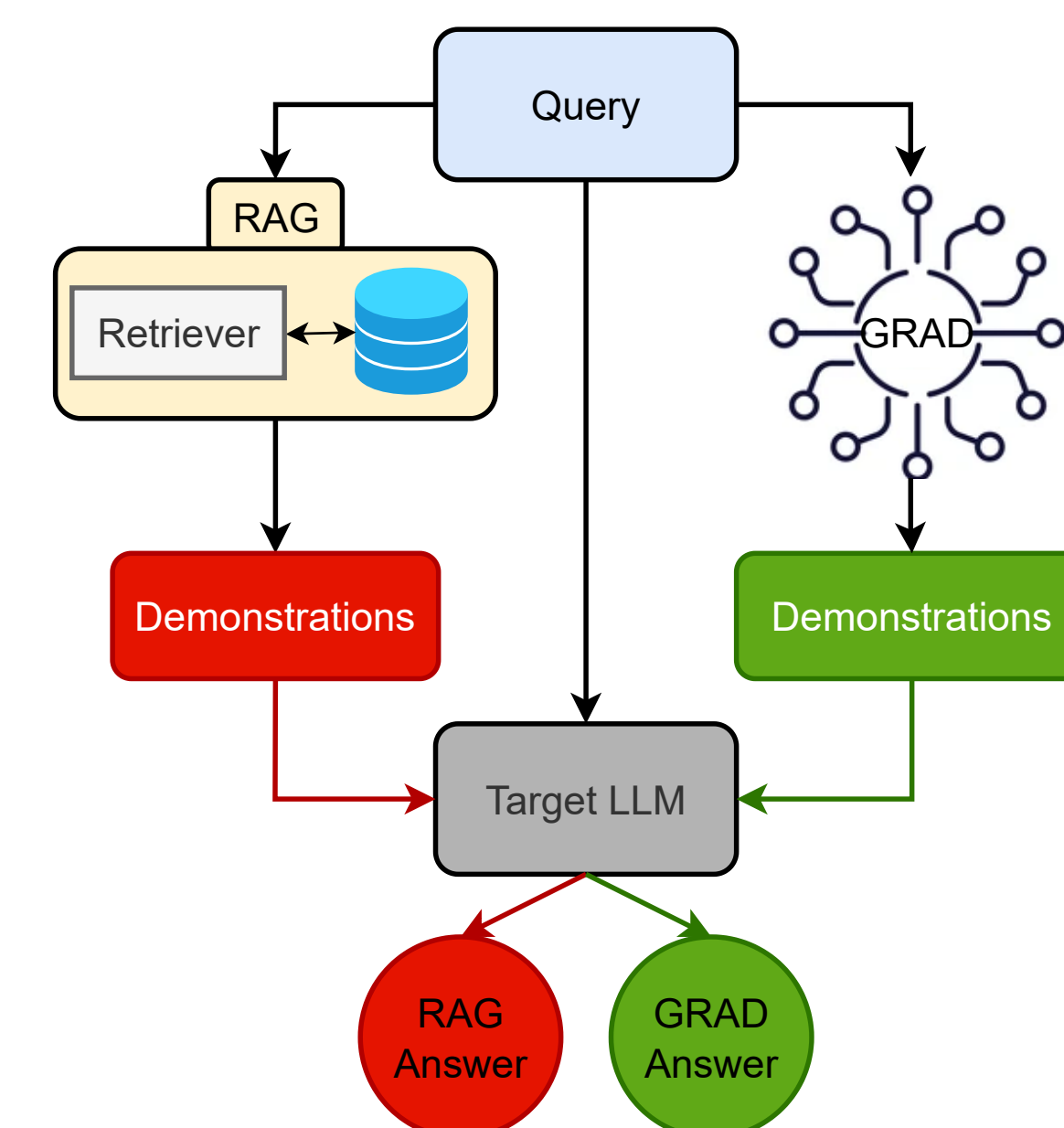
Example 1:
Scenario: At a local café, a customer buys a large coffee for \$3 and three times that amount for a slice of cake. How much did they spend in total?
Reasoning:
- Cost of coffee = \$3
- Cost of cake = 3 * \$3 = \$9
- Total cost = \$3 + \$9 = \$12
[12]
[END_DEMO]

Example 2

Example 2:
Scenario: A student buys a notebook for \$5 and spends three times that amount on a set of colored pencils. What is the total expenditure?
Reasoning:
- Notebook cost = \$5
- Colored pencils cost = 3 * \$5 = \$15
- Total cost = \$5 + \$15 = \$20
[20]
[END_DEMO]

Answer

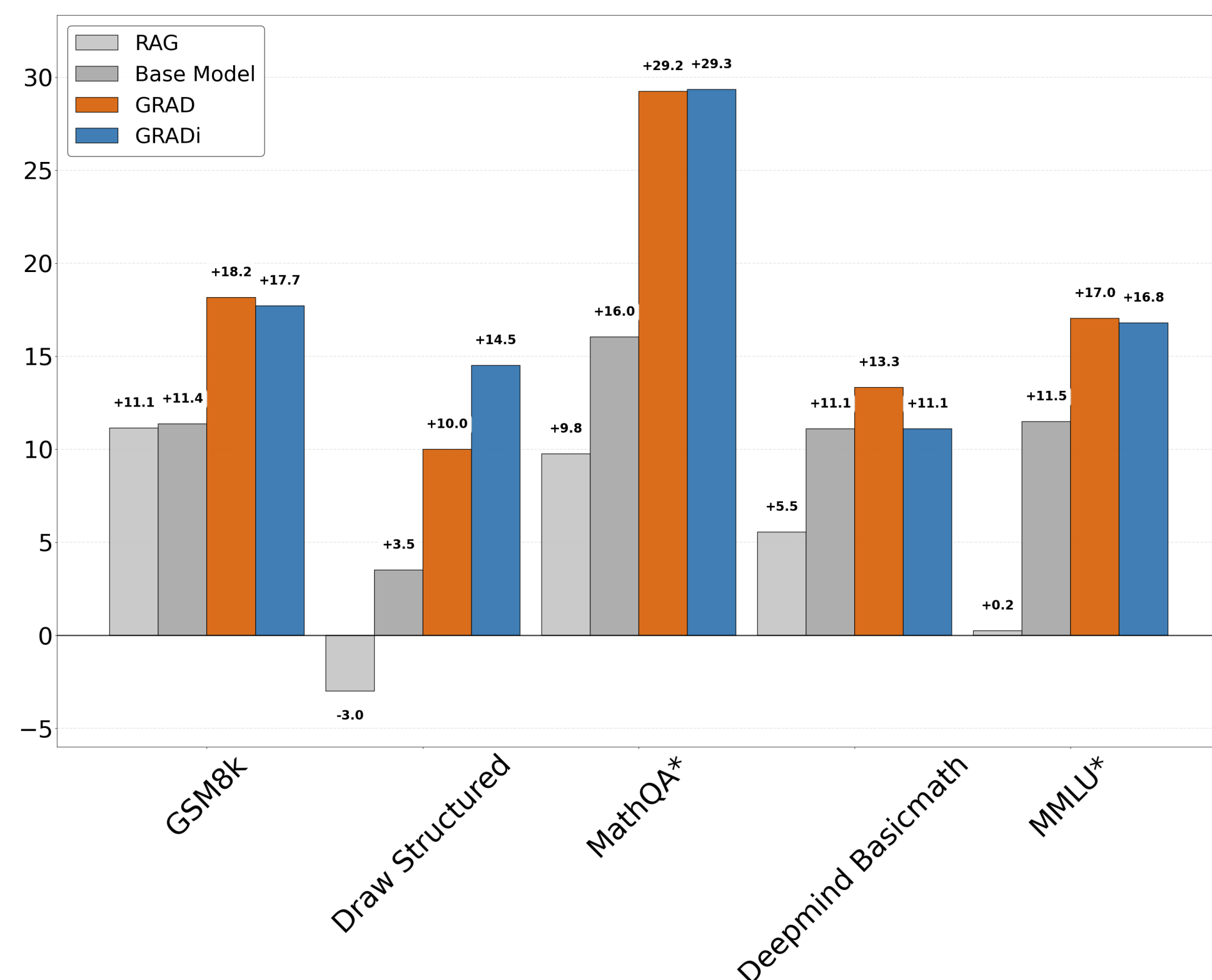
Reasoning:
- Cost of foot-long fish sub = \$40
- Cost of six-inch cold-cut combo sub = 3 * \$40 = \$120
- Total cost = \$40 + \$120 = \$160
[160]



- With GRAD, the problem is solved correctly!
- GRAD is trained under length constraints \Rightarrow both demos and final outputs are shorter!

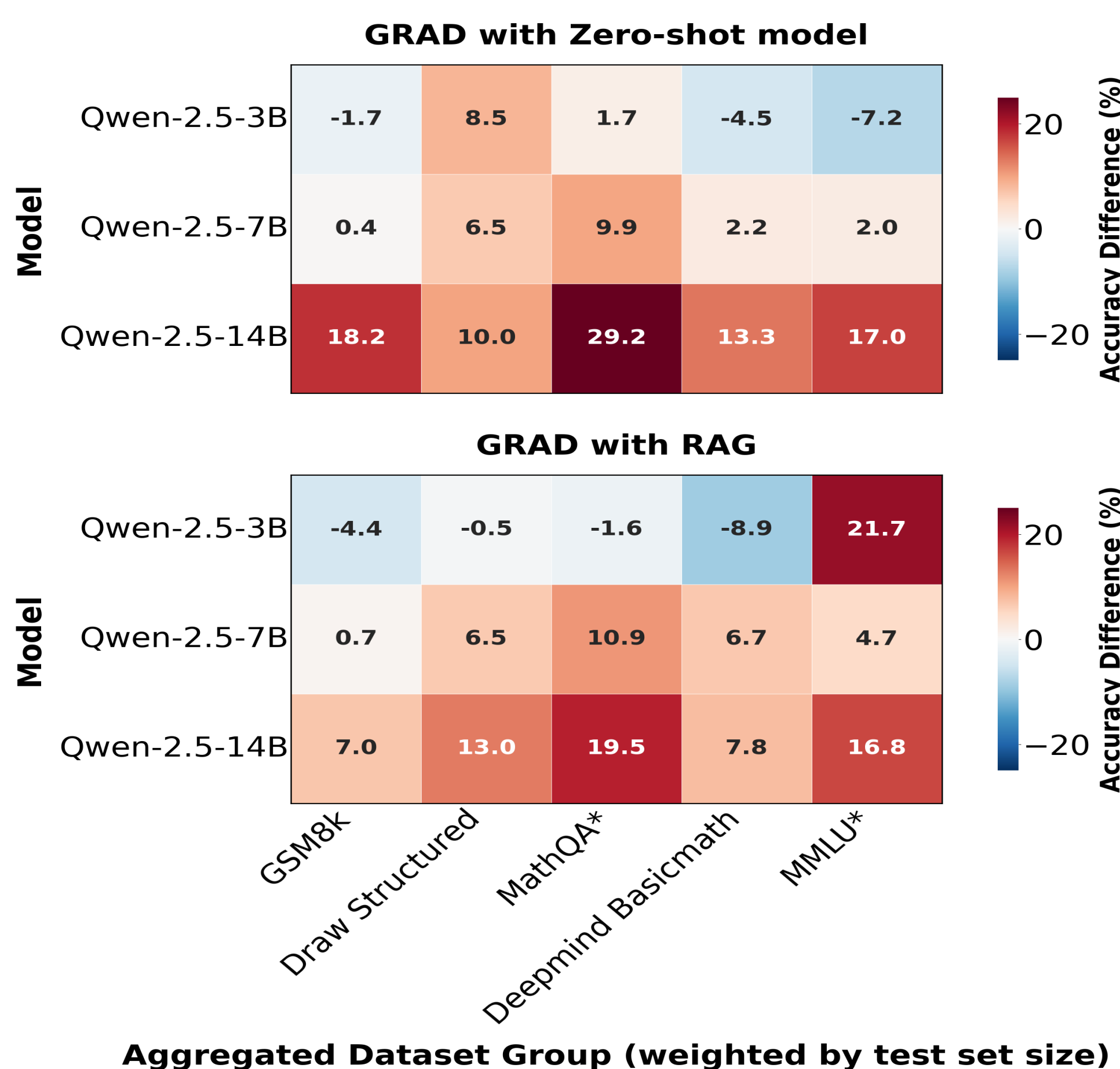
Results

14B GRAD(i) crushes all the baselines!



GRAD and GRADi outperform common baselines: zero-shot, RAG, and self-demo.

Performance Scales with Model Size



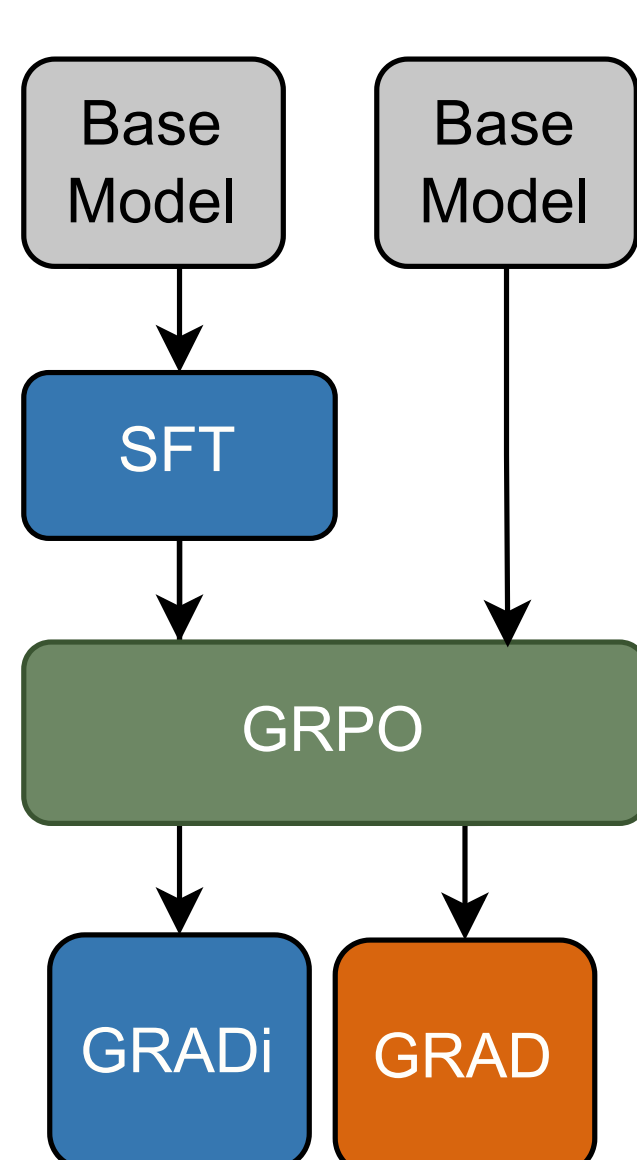
Can cheaper models provide demonstrations for larger ones?

Dataset	3B-Demo	7B-Demo	14B-BASE
GSM8K	66.67	72.22	84.12
draw_structured	42.5	47.00	34
MathQA*	51.23	54.13	43.78
deepmind basic_math	66.67	72.22	70
MMLU*	50.73	57.288	59.57

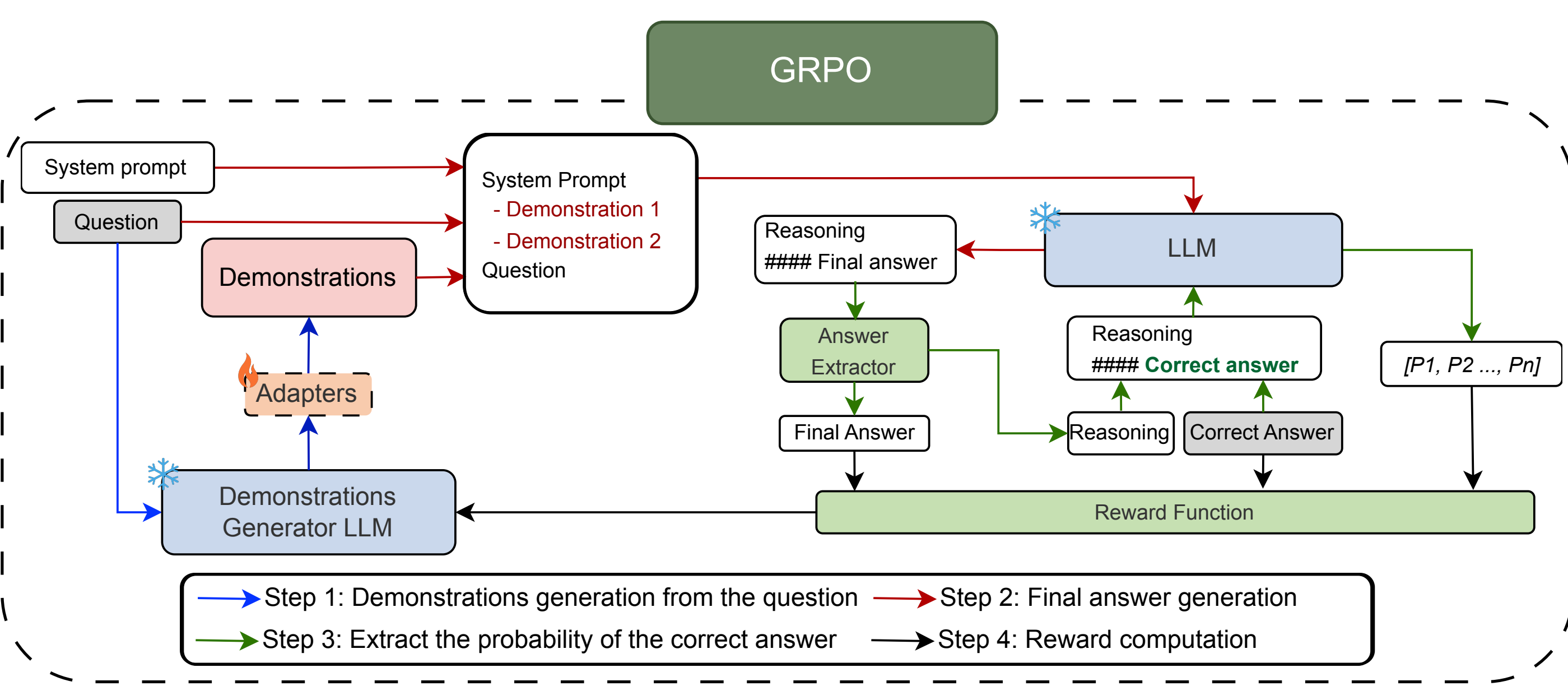
The target model (14B) still benefits from 3B and 7B demonstration samplers.

Pipeline

Training pipeline



GRPO pipeline and reward computation



$$\text{Reward} = R_{p[\text{correct_ans}]} + R_{\text{accuracy}} + R_{\# \text{demos}}.$$

$$R_{p[\text{correct_ans}]} = \frac{1}{1 + L_{\text{llm}}}$$

$$R_{\text{accuracy}}(a; a^*) = \mathbf{1}_{\{a=a^*\}}$$

$$R_{\# \text{demos}} = \frac{n}{D} \times \mathbf{1}_{\{n \leq 4\}}$$

Conclusion

- GRAD consistently outperforms Zero-shot and RAG across all the benchmark datasets
- GRAD generalizes beyond the dataset it was trained on: on MMLU subsets, GRAD improves Zero-shot and RAG performance by $>16\%$.
- Fine-tuned smaller models (3B, 7B) generate demonstrations that enable the bigger model (14B) to achieve competitive accuracy with minimal performance loss.

