THE 2024 JOINT INTERNATIONAL CONFERENCE ON COMPUTATIONAL LINGUISTICS, LANGUAGE RESOURCES AND EVALUATION

QDMR-based Planning-and-Solving Prompting for Complex Reasoning Tasks

Jinfeng Huang 黄金凤

University of Science and Technology of China

Email: huangjf11@mail.ustc.edu.cn

Background

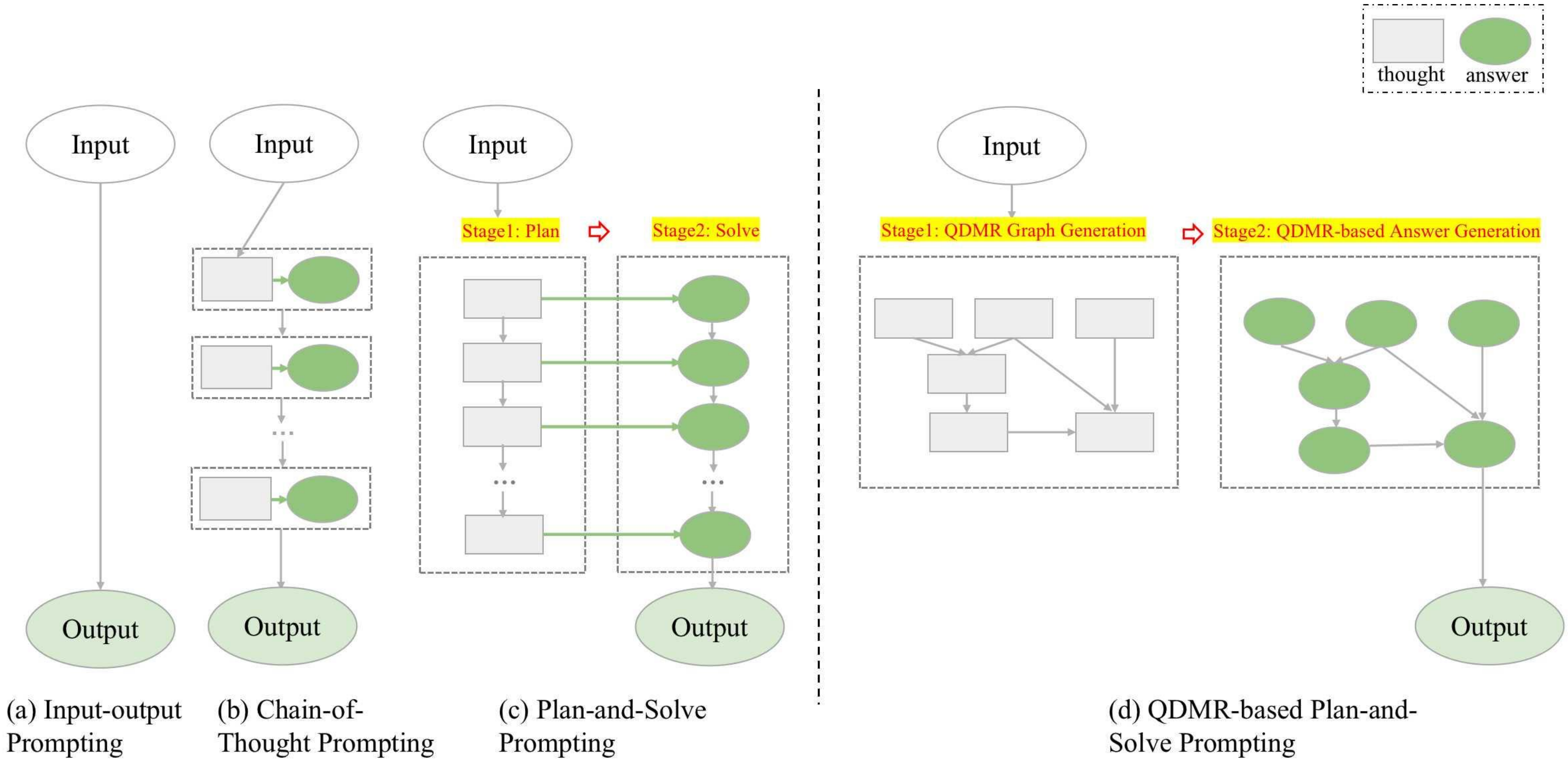


Fig 1: Comparison of different prompting techniques

Outline

1. Research Task
2. Research Goal
3. Research Method
4. Model Performance
5. Conclusions

1 Research Task

Question

Erwan went on shopping. He purchased a pair of shoes at \$200 but discounted 30%, and two shirts at \$80 each. Upon checkout, the cashier said that there is an additional 5% discount. How much did he spend after all the discounts?

Reasoning

Step 1: The total discount of a pair of shoes is \$200 x 30/100 = \$60.
Step 2: He spent \$200 - \$60 = \$140 for a pair of shoes.
Step 3: The total cost of two shirts is \$80 x 2 = \$160.
Step 4: The initial cost of his purchase is \$140 + \$160 = \$300.
Step 5: He got an additional \$300 x 5/100 = \$15 discount .
Step 6: Therefore, Erwan spent \$300 - 15 = \$285 after discounts.

Answer

The final answer is ##285##.

Fig 2: Example of complex reasoning problem

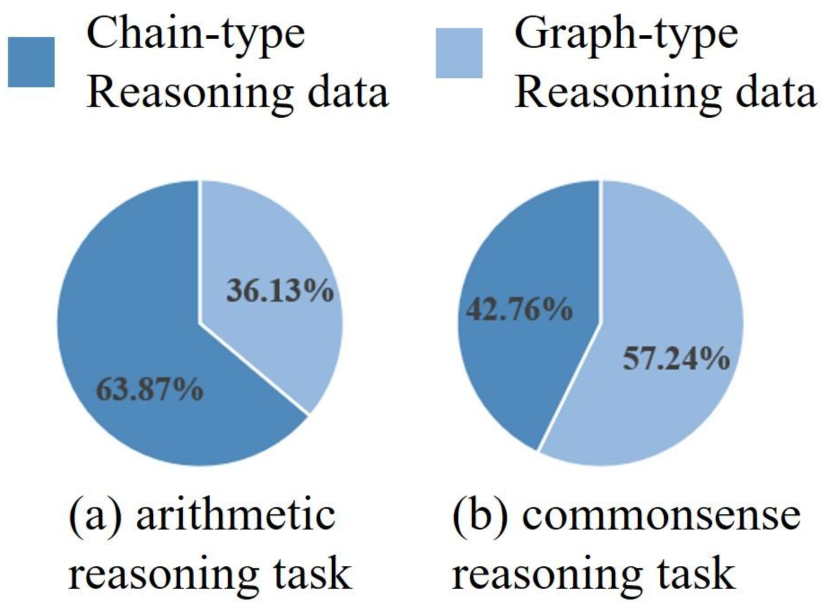


Fig 3: The proportion of chain-type and graph-type reasoning data in arithmetic and commonsense task

2 Research Goal

Improve accuracy and stability in complex reasoning tasks

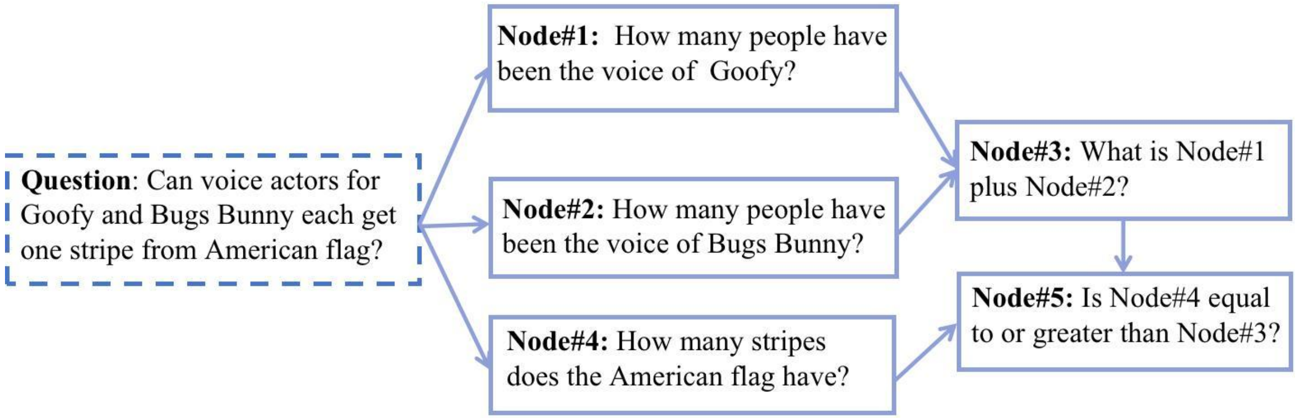


Fig 4: An example of Question Decomposition Meaning Representation(QDMR) Graph

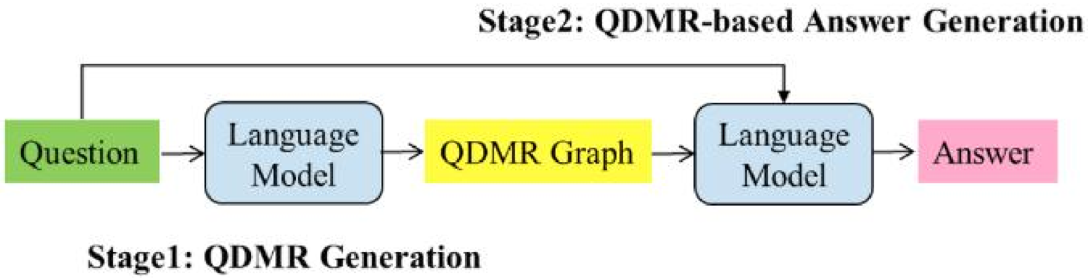


Fig 5: Overall Architecture

3 Research Method

Question Decomposition Meaning Representation Graph Generation

(a) Model Input

There are some demonstration examples for QDMR graph generation prompting are to omitted in there. The example include original question similar to Figure 5(a) as input and QDMR graph for original question similar to Figure 5(b) as output.

Q: Can voice actors for Goofy and Bugs Bunny each get one stripe from American flag?

A: Please refer to the above example. You need to generate a specific Question Decomposition Meaning Representation Graph for the given Question, which describes the complete logic of Question solving. In the Question Decomposition Meaning Representation Graph, nodes are different sub-questions. By indicating the parent node, the relationship between different nodes, i.e. the relationship between different sub-questions, can be explained.

You need to pay attention to the following points:

(1) **The Question Decomposition should be as fine-grained as possible.**

(2) For different sub-questions, **overlapping between diffrent sub-questions should be avoided.**

(3) When you generate a new sub-question(that is, a sub node), you should **identify all the preorder sub-questions related to it and place them in the Parent information.**

(4) During the process of Question Decomposition, you can **set new variables and equations at any time to help solve the Question.**

(b) Model Output

QDMR:

Node#1: How many people have been the voice of Goofy?; **Parent: {}**

Node#2: How many people have been the voice of Bugs Bunny?; **Parent: {}**

Node#3: What is Node#1 plus Node#2?; **Parent: {Node#1,Node#2}**

Node#4: How many stripes does the American flag have?; **Parent: {}**

Node#5: Is Node#4 equal to or greater than Node#3?; **Parent: {Node#3,Node#4}**

(c) Question Decomposition Meaning Representation Graph

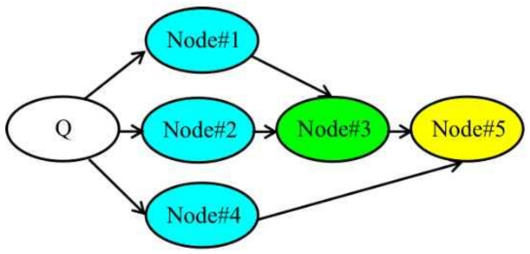


Fig 6: Example of generating QDMR graph through in-context learning

3 Research Method

QDMR-based Answer Generation

(a) Model Input

There are some demonstration examples for QDMR-based answer generation prompting are to omitted in there. The example include original question and QDMR graph similar to Figure 6(a) as input and QDMR-based answer similar to Figure 6(b) as output.

Q: Can voice actors for Goofy and Bugs Bunny each get one stripe from American flag?

QDMR:
Node#1: How many people have been the voice of Goofy?; **Parent: {}**
Node#2: How many people have been the voice of Bugs Bunny?; **Parent: {}**
Node#3: What is Node#1 plus Node#2?; **Parent: {Node#1,Node#2}**
Node#4: How many stripes does the American flag have?; **Parent: {}**
Node#5: Is Node#4 equal to or greater than Node#3?; **Parent: {Node#3,Node#4}**
A: Please refer to the above example. You need to use the <<Question Decomposition Meaning Representation Graph>> to generate a specific <<Answer>> for the <<Question>>.
You need to pay attention to the following points:
(1) You should solve the sub-question in the topological order of the <<Question Decomposition Meaning Representation Graph>>, that is sequentially solve sub-question 1, 2, 3, ...(Node#1,Node#2,Node#3,...).
(2) When solving each sub-question[Node#i], you should rely on its parent information. Repeat the solution result of the related pre-order sub-questions to limit the solution space of the sub-question[Node#i] to their relevant contextual environment.
(3) In the <<Question Decomposition Meaning Representation Graph>>, the solution result to the last sub-question is the answer to the original <<Question>>.

(b) Model Output

Answer:
Node#1: How many people have been the voice of Goofy?; Parent: {}
According to common knowledge, ##Six people have been the voice of Goofy.##
Node#2: How many people have been the voice of Bugs Bunny?; Parent: {}
According to common knowledge, ##Seven people have been the voice of Bugs Bunny.##
Node#3: What is Node#1 plus Node#2?; Parent: {Node#1,Node#2}
Given <<Six people have been the voice of Goofy.>> and <<Seven people have been the voice of Bugs Bunny.>>, ##Thirteen people have been the voice of Goofy and Bugs Bunny. ##
Node#4: How many stripes does the American flag have?; Parent: {}
According to common knowledge, ##The American flag has Thirteen stripes on it. ##
Node#5: Is Node#4 equal to or greater than Node#3?; Parent: {Node#3,Node#4}
Given <<The American flag has Thirteen stripes on it. >> and <<Thirteen people have been the voice of Goofy and Bugs Bunny. >>, ##The number of stripes owned by American flag is same as the number of people have been the voice of Goofy and Bugs Bunny.##
Therefore, the final answer is true.

(c) QDMR-based Answer Generation

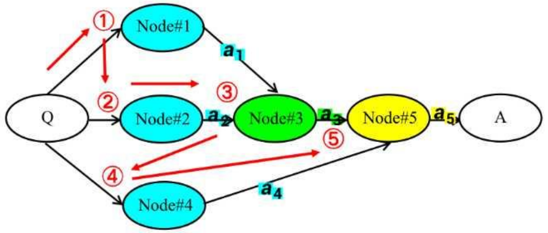


Fig 7: Example of generating answer based on the QDMR graph through in-context learning

4 Model Performance

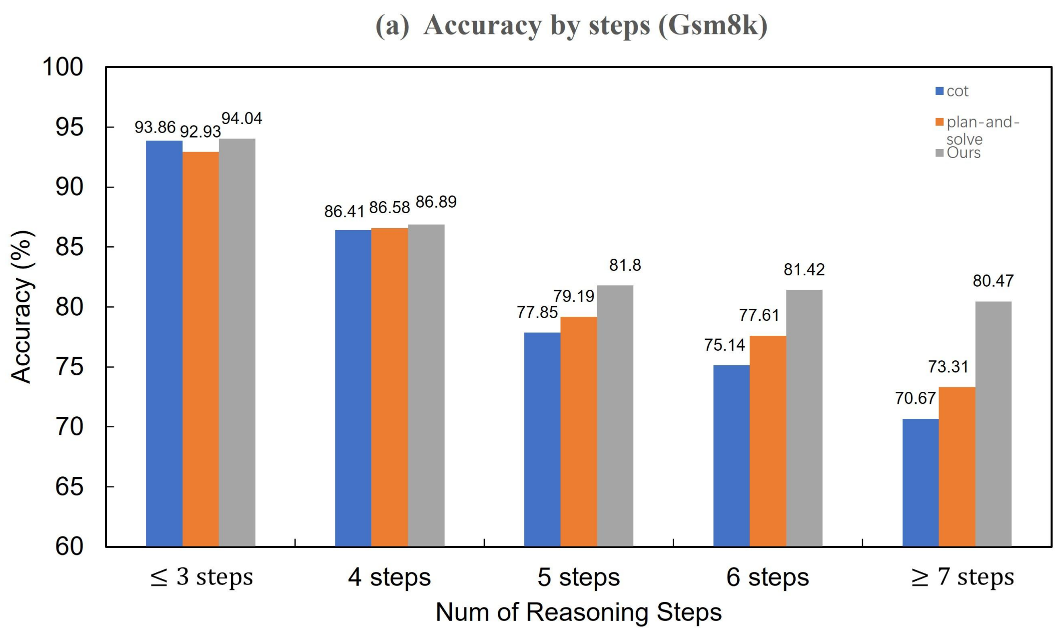
Accuracy Analysis and Comparisons

Model	Prompt	Dataset						Average
		AQuA	Gsm8k	MultiArith	SingleEq	Musique	HotpotQA	
text-davinci-002	Cot	0.3743	0.5838	0.9212	0.8571	0.6783	0.8123	0.7045
	PS	**0.4185**	0.6587	**0.9567**	0.8970	0.6854	0.8331	0.7416
	Ours	0.4181	**0.6737**	0.9474	**0.9148**	**0.7005**	**0.8475**	**0.7503**
text-davinci-003	Cot	0.3426	0.7079	0.9434	0.9288	0.6916	0.8519	0.7444
	PS	0.4002	0.7164	0.9506	0.9403	0.7073	**0.8874**	0.7670
	Ours	**0.4194**	**0.7520**	**0.9510**	**0.9591**	**0.7351**	0.8853	**0.7837**
GPT3.5-Turbo	Cot	0.4672	0.8312	0.9677	0.9611	0.6132	0.8586	0.7832
	PS	0.4440	0.8385	**0.9821**	0.9795	0.6467	0.8706	0.7936
	Ours	**0.4688**	**0.8613**	0.9735	**0.9826**	**0.6683**	**0.8997**	**0.8090**

Table 1: Within the experimental setting of in-context learning, accuracy comparison between our method, Chain-of-Thought prompting(Cot) and Plan-and-Solve prompting(PS) on multiple reasoning datasets. The best results are boldfaced.

4 Model Performance

Accuracy Analysis and Comparisons



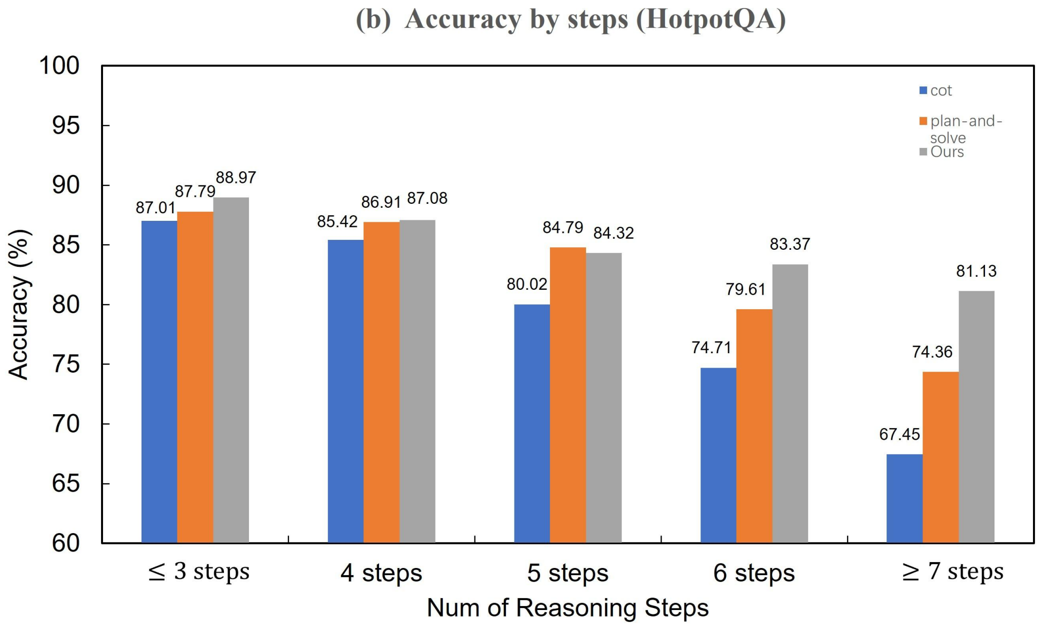


Fig 8: Accuracy (%) of our method compared with chain of thought and plan-and solve prompting , broken down by the number of reasoning steps required in the expected solution.

4 Model Performance

Accuracy Analysis and Comparisons

Model	Prompt	AQuA	Gsm8k	MultiArith	SingleEq	Musique	HotpotQA	Average
		Dataset						
Llama2-7B	ICL-Cot	0.2638	0.2669	0.6876	0.6561	0.3598	0.4111	0.4409
	ICL-PS	0.2874	0.2775	0.7109	0.6501	0.3950	0.3921	0.4522
	ICL-Ours	0.2969 (38%)	0.2861 (44%)	0.7325 (86%)	0.6782 (74%)	0.3988 (64%)	0.4370 (58%)	0.4716
	SFT-Ours	**0.3110** (42%)	**0.3035** (50%)	**0.7378** (90%)	**0.7003** (90%)	**0.4146** (84%)	**0.4554** (76%)	**0.4871**
Chatglm2-6B	ICL-Cot	0.2717	0.2835	0.7002	0.6765	0.3789	0.3699	0.4468
	ICL-PS	0.3150	0.3078	0.6942	0.6802	0.3907	0.3780	0.4610
	ICL-Ours	0.3254 (36%)	0.3269 (40%)	0.7230 (74%)	0.7037 (66%)	0.4017 (60%)	0.4140 (56%)	0.4825
	SFT-Ours	**0.3465** (48%)	**0.3362** (56%)	**0.7391** (82%)	**0.7285** (78%)	**0.4337** (72%)	**0.4378** (74%)	**0.5036**

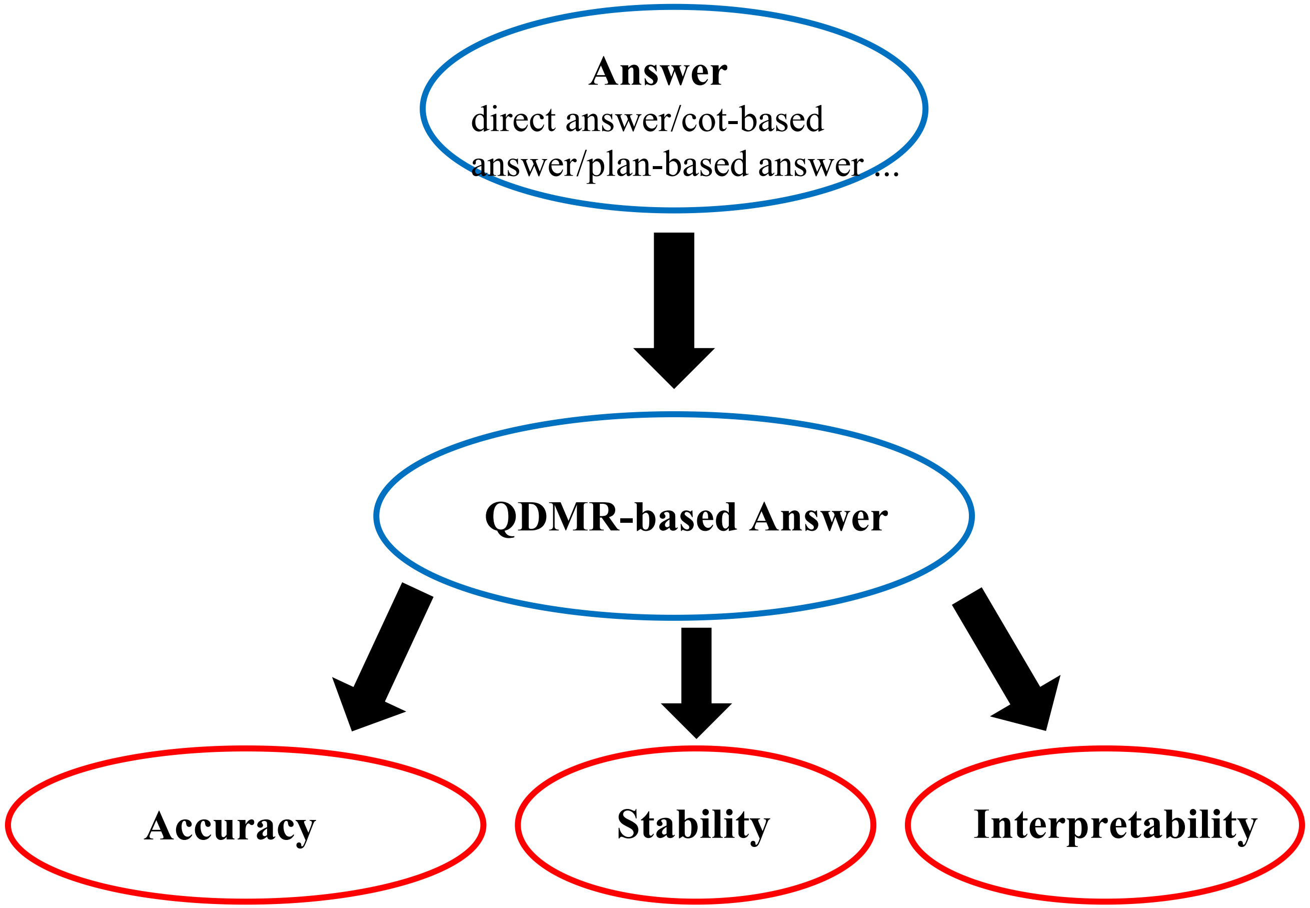
Table 2: Comparison of supervised fine-tuning and in-context learning in producing QDMR graph and related effects on the final reasoning results. The best results are boldfaced. Tip: The first line in ICT-Ours and SFT-Ours represents the Accuracy, the second line in ICT-Ours and SFT-Ours represents quality evaluation of the QDMR graph generated under the current settings.

4 Model Performance

Table 3: Examples of QDMR Graph produced by Llama2-7B Model in In-context Learning and Supervised Fine-tuning Settings.

Dataset	Question	QDMR Graph Generated Through In-context Learning	QDMR Graph Generated Through Supervised Fine-tuning
Musique	What is the occupation of the father of the actor that plays Will Graham on teh series 'Hannibal'?	**Node#1:** Who is the father of the actor that plays #4?;Parent:{#4} **Node#2:** #1 is the occupation of whom?;Parent:{#1} **Node#3:** Who plays #4?;Parent:{#4}	**Node#1:** who plays will graham on hannibal?; Parent:{} **Node#2:** #1 >>father?; Parent:{#1} **Node#3:** What job did #2 have?; Parent:{#2}
GSM8K	In a dance class of 20 students, 20% enrolled in contemporary dance, 25% of the remaining enrolled in jazz dance, and the rest enrolled in hip-hop dance.What percentage of the entire students enrolled in hip-hop dance?	**Node#1:** How many students enrolled in hip-hop dance?;Parent:{} **Node#2:** How many students enrolled in contemporary dance?;Parent:{} **Node#3:** How many students enrolled in jazz dance?;Parent:{} **Node#4:** What percentage of the entire students enrolled in hip-hop dance?; Parent:{#1,#2,#3}	**Node#1:** How many students are enrolled in contemporary dance?;Parent:{} **Node#2:** How many students are left after contemporary dance?;Parent:{#1} **Node#3:** How many students are enrolled in jazz dance?;Parent:{#2} **Node#4:** How many students enrolled in hip-hop dance?;Parent: {#2,#3} **Node#5:** What percentage of the entire students are enrolled in hip-hop dance?;Parent:{#4}

5 Conclusions



5 Conclusions

Research limitations

- The QDMR graph is generated using either prompt-driven techniques or fine-tuning language models. However, the current production process of the QDMR graph suffers from a lack of fine-grained guidance signals. This deficiency prevents the system from resolving issues such as deadlocks, irrelevant paths, or loops within the graph. In our subsequent research, we plan to explore the integration of procedural supervision with reinforcement learning techniques, including Reward Learning from Human Feedback (RLHF), to provide more precise supervisory signals.
- The experiment was conducted exclusively within the contexts of mathematical reasoning and commonsense reasoning. In the future, we aim to expand our research to include more diverse and complex tasks and scenarios. Moreover, we will delineate the process of problem understanding and resolution into two distinct phases, utilizing external tools to perform actual operations, aiming to reduce the occurrence of hallucinations.

Acknowledgements

- Prof. Tong Xu and Prof. Wenbin Jiang
- Research partners

Foundation:

- National Natural Science Foundation of China
- USTC Research Funds of the Double First-Class Initiative
- Fundamental Research Funds for the Central Universities

Thank you for your attention!