Question Translation Training for Better Multilingual Reasoning

Wenhao Zhu

Overview

Question Translation Training for Better Multilingual Reasoning

Wenhao Zhu[1], **Shujian Huang**[1], **Fei Yuan**[2], **Shuaijie She**[1], **Jiajun Chen**[1], **Alexandra Birch**[3]
[1] National Key Laboratory for Novel Software Technology, Nanjing University
[2] Shanghai AI Lab [3] School of Informatics, University of Edinburgh
zhuwh@smail.nju.edu.cn, huangsj@nju.edu.cn, yuanfei@pjlab.org.cn, shesj@smail.nju.edu.cn
chenjj@nju.edu.cn, a.birch@ed.ac.uk

- TL'DR:
  - Goal: fine-tuning LLMs into strong **multilingual** reasoners.
  - Approach: utilizing the **translation task** to strengthen language alignment within the LLM and this way enables us to **transfer** its English proficiency to non-English scenarios.

Why Do We Choose This Topic?

- The lessons that I learned from my last year's works:
  - INK: Injecting kNN Knowledge in Nearest Neighbor Machine Translation.
    - Framework (interesting idea) is more important than implementation (performance score).
  - Multilingual Machine Translation with Large Language Models: Empirical Results and Analysis.
    - LLM's multilingual performance is far from satisfactory.
  - Extrapolating Large Language Models to Non-English by Aligning Languages.
    - Translation training can now impact other tasks .
    - Targeting a rigorous evaluation task is important in this new era.

⇩

"**Question Translation Training for Better Multilingual Reasoning**"

improving multilingual performance + the rigorous reasoning task + two-step training framework

Multilingual Mathematical Reasoning

- Mathematical Reasoning
  - predicting the numerical answer based on the given question
  - answering with chain-of-thought usually gets more accurate prediction
- Shi et al. extend it to a multilingual task (mGSM).

Question: Olivia has $23. She bought five bagels for $3 each. How much money does she have left?
Answer: 8

English

Question: Olivia has $23. She bought five bagels for $3 each. How much money does she have left?
Step-by-Step Answer: 5 bagels for $3 each should cost 5 * 3 = 15 dollars. Olivia had $23 in the beginning, so now she has 23 - 15 = 8 dollars left. The answer is **8**.

English (answering with CoT)

Frage: Olivia hat 23 US-Dollar. Sie hat fünf Bagels für 3 US- Dollar pro Stück gekauft. Wie viel Geld hat sie übrig?
Antwort: 8

German

问题： 奥利维亚有 23 美元。她买了五个单价 3 美元的百吉饼。 她还剩多少钱？
解答: 8

Chinese

Figure from: Shi et al. Language Models Are Multilingual Chain-of-Thought Reasoners.

Unbalanced Multilingual Performance

- LLMs' performance on English questions is much higher than it on non-English questions.

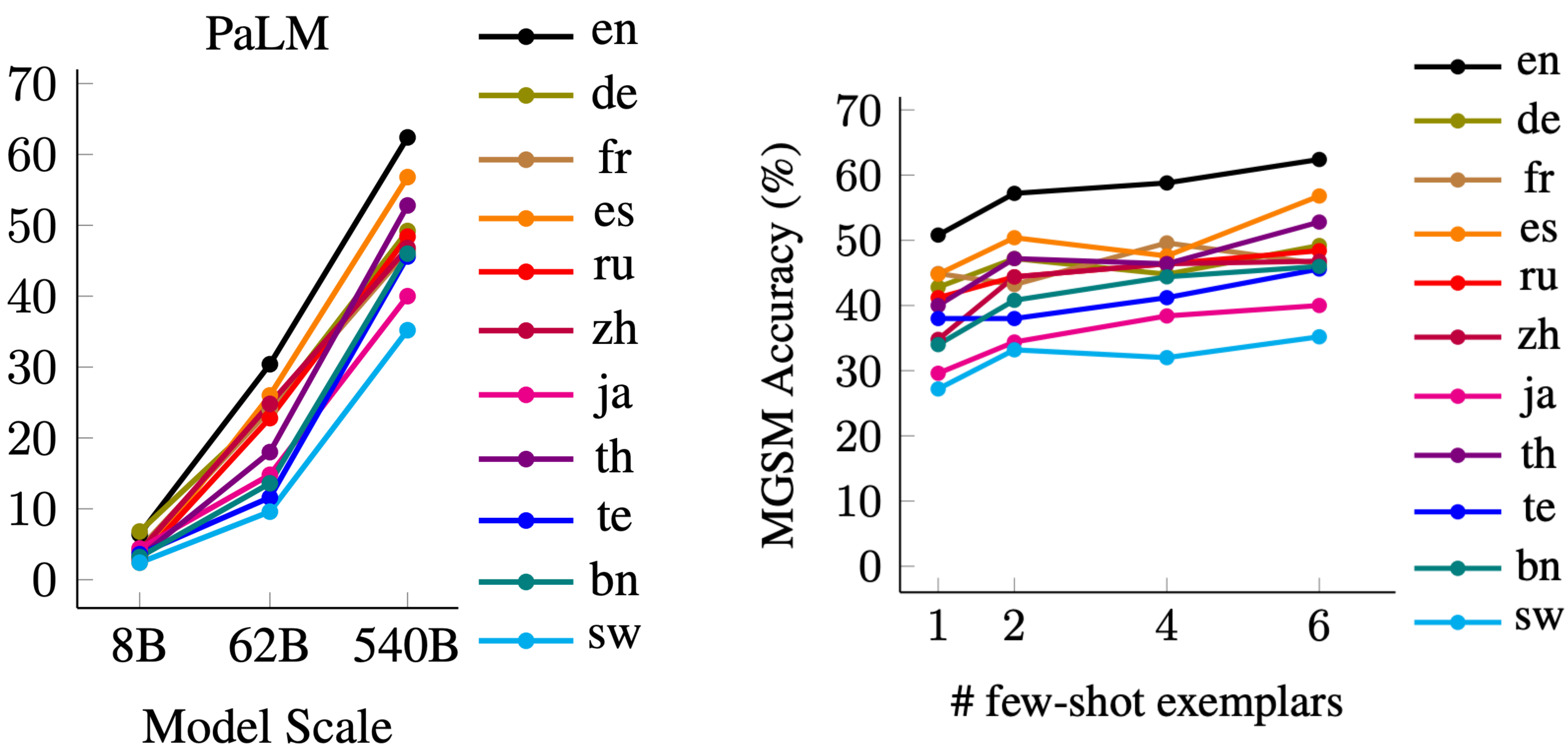
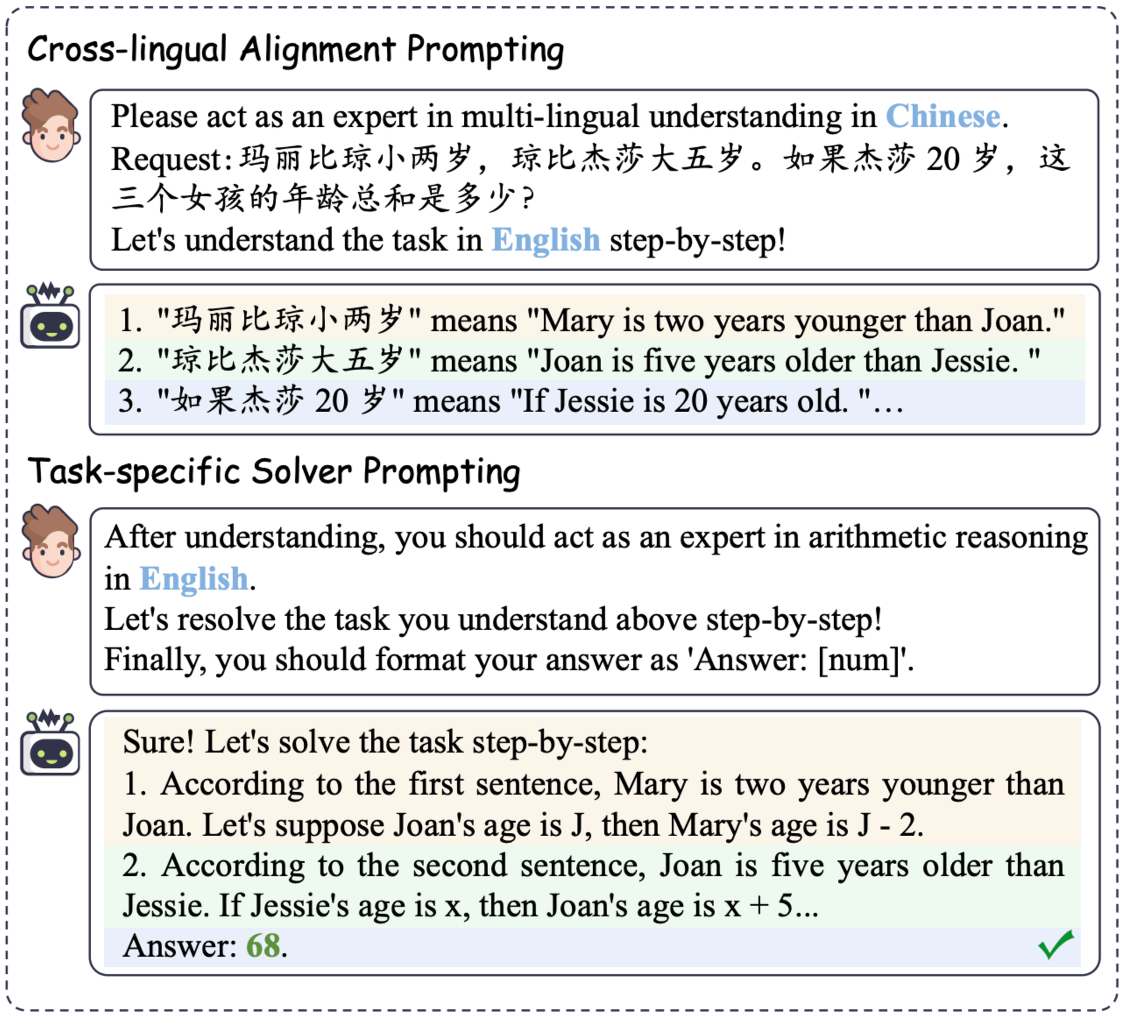


Figure from: Shi et al. Language Models Are Multilingual Chain-Of-Thought Reasoners.

Related Work

- prompting close-source LLMs (translate-test)
    - The effectiveness of these prompting methods are not well-examined on open-source LLMs.



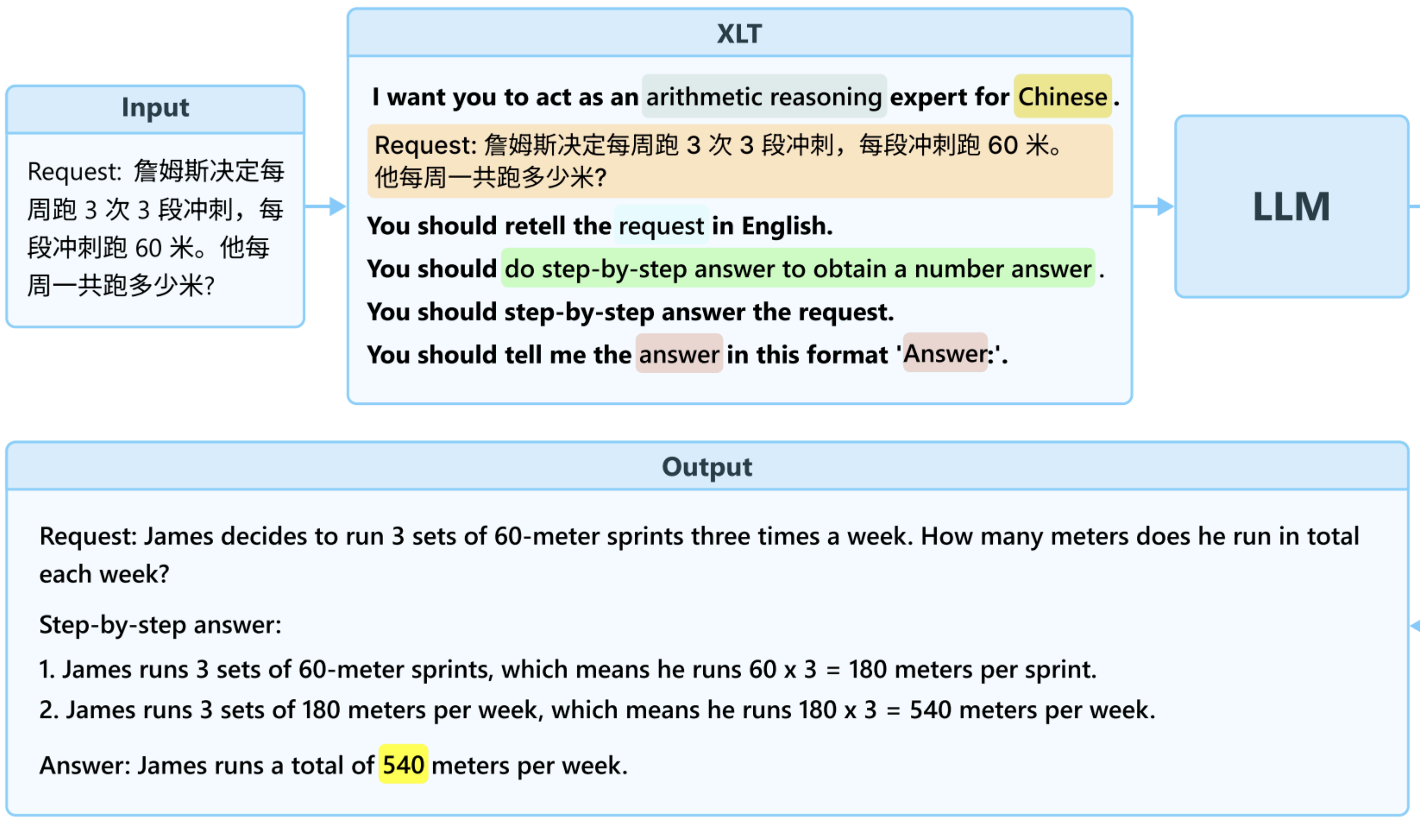


Figure from: Qin et al. Cross-lingual Prompting: Improving Zero-shot Chain-of-Thought Reasoning across Languages. & Huang et al. Not All Languages Are Created Equal in LLMs: Improving Multilingual Capability by Cross-Lingual-Thought Prompting.

Related Work

- instruction-tuning open-source LLMs (translate-training)
  - Translating English questions and CoT responses to non-English languages.
  - Combing multilingual data for instruction-tuning.

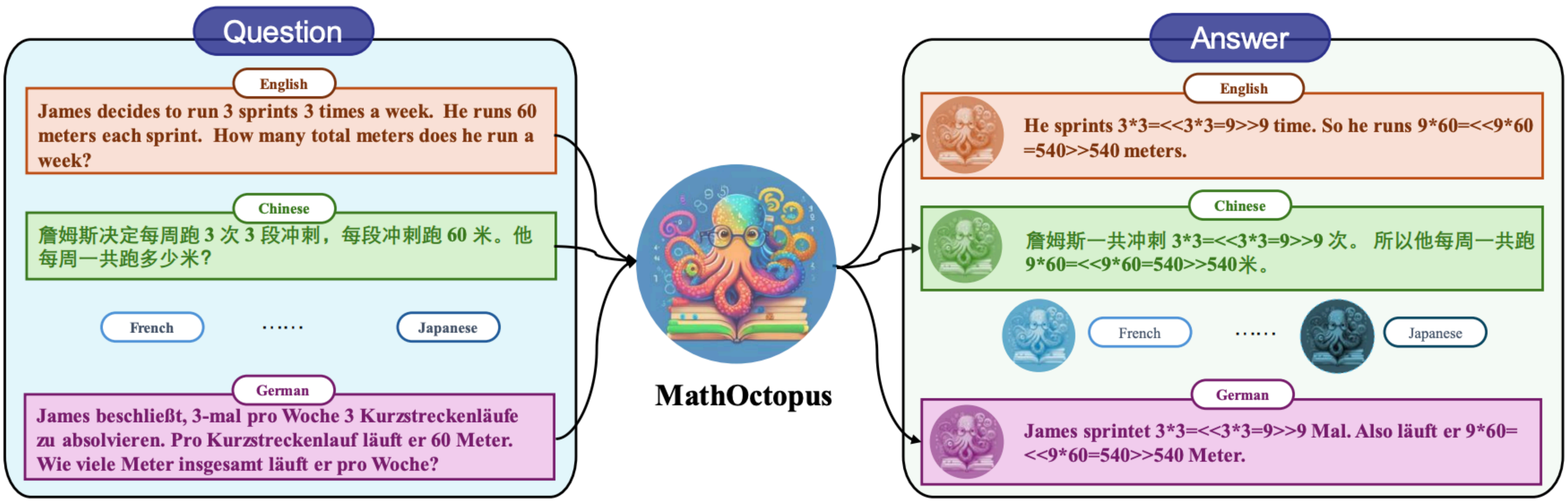
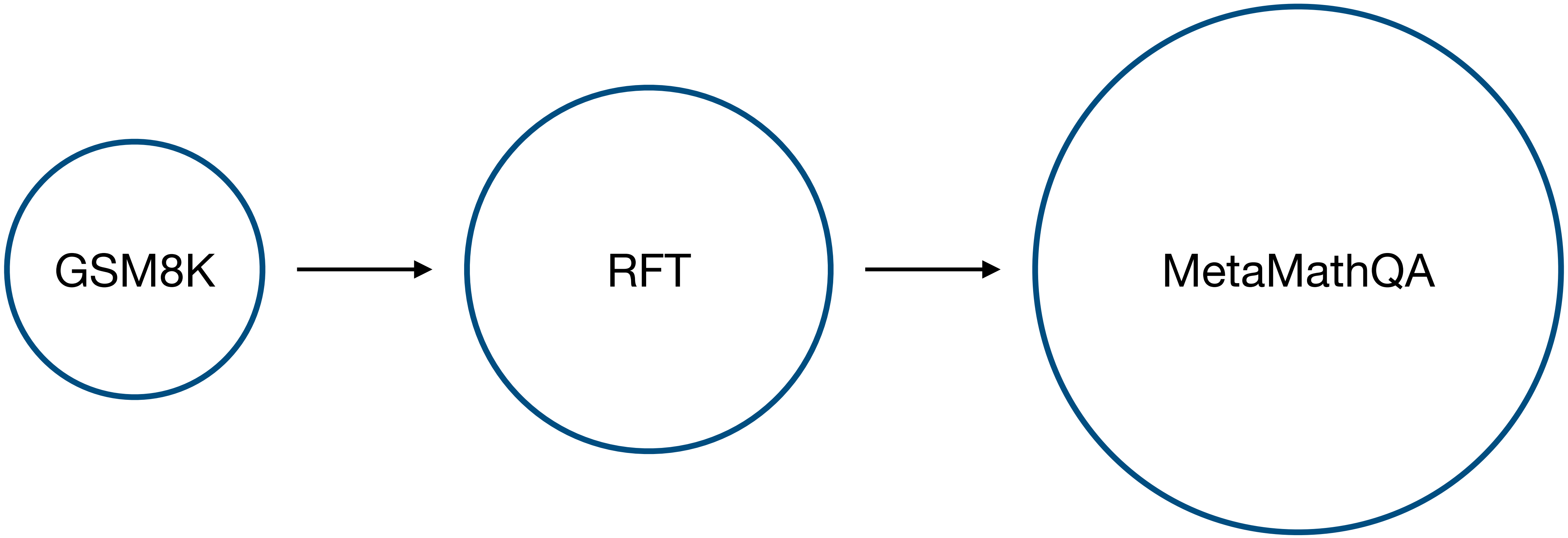


Figure from: Chen et al. Breaking Language Barriers in Multilingual Mathematical Reasoning: Insights and Observations.

Potential Drawbacks of Translate-training

- Translating English training data to numerous non-English languages incurs significant **translation cost**, especially considering the **constant addition** of large and complex instruction tuning sets.



GSM8K: Cobbe et al. Training Verifiers to Solve Math Word Problems.
RFT: Yuan et al. Scaling Relationship on Learning Mathematical Reasoning with Large Language Models.
MetaMathQA: Yu et al. Metamath: Bootstrap Your Own Mathematical Questions for Large Language Models.

Potential Drawbacks of Translate-training

- It is hard for the translation engine to **accurately translate** lengthy, logical texts containing mathematical symbols in chain-of-thought responses.

```
[English Response]
Half of the wallet's price is $100 / 2 = $50.
Betty's grandparents gave her $15 * 2 = $30.
In total, Betty has $50 + $15 + $30 = $95.
So she still needs $100 - $95 = $5 which her parents plan to give her for her
birthday.

[Chinese Response]
钱包价格的一半是100美元 / 2 = 50美元。
贝蒂的祖父母给了她15美元 * 2 = 30美元。
这意味着，贝蒂还需要50美元 - 30美元 - 15美元 = 5美元。

[Chinese Response (English Meaning)]
Half the wallet price is $100/2=$50.
Betty's grandparents gave her $15*2=$30.
This means that Betty still needs $50 - $30 - $15 = $5.
```

Example: missing reasoning step in the translated response

Potential Drawbacks of Translate-training

- It is hard for the translation engine to **accurately translate** lengthly, logical texts containing mathematical symbols in chain-of-thought responses.

```
[English Response]
The total time he takes to walk to work 3 times a week is 2*3 = 6 hours
The total time he takes to walk to and from work 3 times a week is 6+6 = 12 hours
The total time he takes to bike to work 2 times a week is 2*1 = 2 hours
The total time he takes to bike to and from work 2 times a week is 2+2 = 4 hours
The total time he takes to go to and from work in a week is 12+4 = 16 hours

[Chinese Response]
Roque每周走路去上班需要2*3=6小时。
Roque每周骑自行车去上班需要1*2=2小时。
他每周去上班的总时间是6+2=8小时。
他每周回家的总时间也是8小时。
因此，Roque每周去上班和回家总共需要8+8=16小时。

[Chinese Response (English Meaning)]
It takes Roque 2*3=6 hours to walk to work every week.
It takes Roque 1*2=2 hours to ride a bicycle to work every week.
The total time he goes to work every week is 6+2=8 hours.
The total time he spends home every week is also 8 hours.
Therefore, it takes Roque a total of 8+8=16 hours per week to get to work and
back home.
```

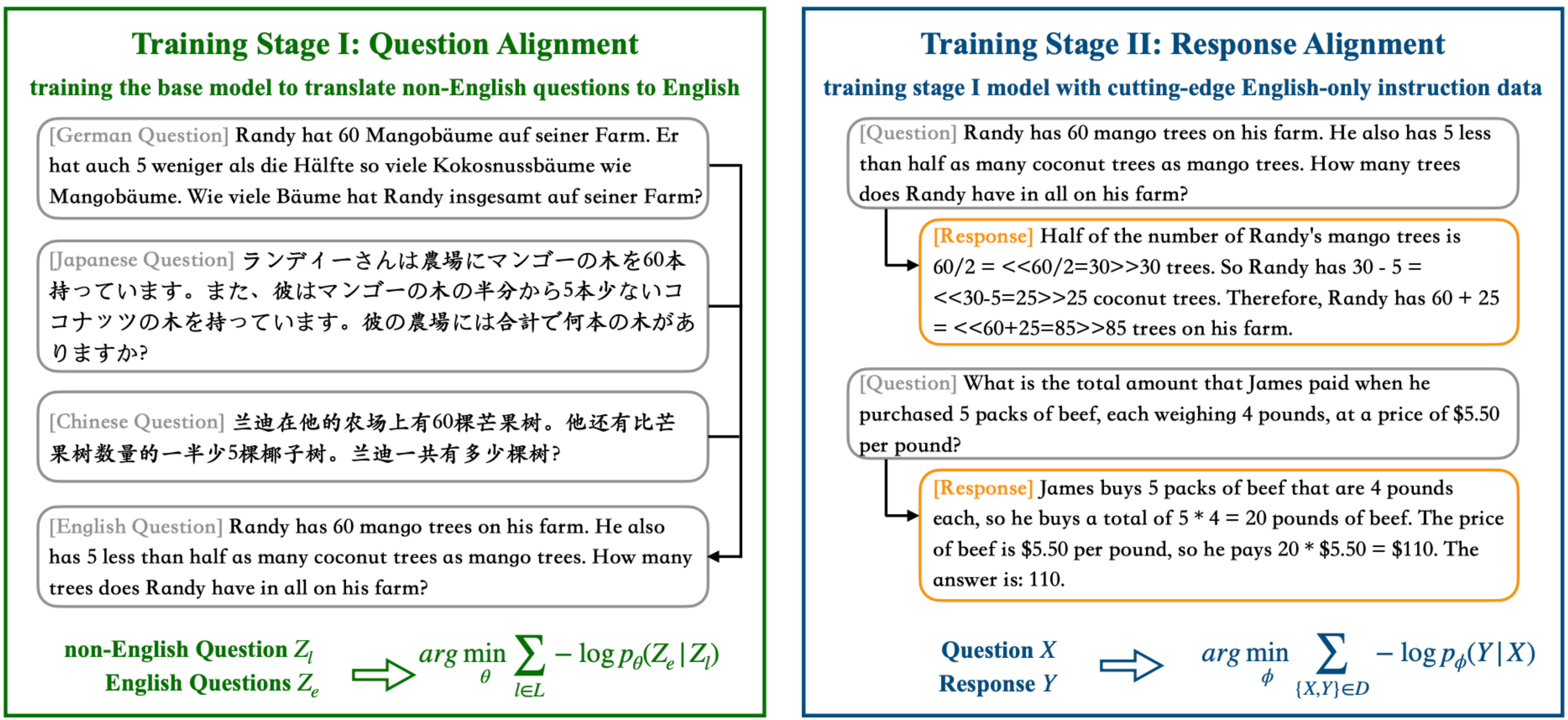
Example: distinct reasoning paths between English response and translated response

How Would the Ideal Solution Look Like?

- Training process
  - using cutting-edge English instruction data
  - not requiring translating English instruction data into non-English
- Inference process
  - The model is able to solve non-English problems with its English expertise.

Our Training Framework

- Intuition: make LLM to understand non-English question in a English way and transferring LLM's proficiency in English to non-English scenarios.



Implementation Details

Below is an instruction that describes a task, paired with an input that provides further context.

Write a response that appropriately completes the request.

Instruction:

Translate the following sentences from Chinese to English. 兰迪在他的农场上有60棵芒果树。他还有比芒果树数量的一半少5棵椰子树。兰迪一共有多少棵树?

Response:

Randy has 60 mango trees on his farm. He also has 5 less than half as many coconut trees as mango trees. How many trees does Randy have in all on his farm?

Implementation Details

Below is an instruction that describes a task, paired with an input that provides further context.

Write a response that appropriately completes the request.

Instruction:

Randy has 60 mango trees on his farm. He also has 5 less than half as many coconut trees as mango trees. How many trees does Randy have in all on his farm?

Response:

Half of the number of Randy's mango trees is 60/2 = <<60/2=30>>30 trees. So Randy has 30 - 5 = <<30-5=25>>25 coconut trees. Therefore, Randy has 60 + 25 = <<60+25=85>>85 trees on his farm.

Experiment Setting

- Base LLM
  - LLaMA2-7B/13B
- Dataset:
  - We use multilingual questions in GSM8KInstruct for question alignment.
  - We use MetaMathQA for response alignment.
  - We use mGSM (IND) and mSVAMP (OOD) for evaluation.

	Dataset	Usage	#Lang	Size
augmented GSM8K →	METAMATHQA	Training	1	395,000
translated GSM8K →	GSM8KINSTRUCT	Training	10	73,559
	MGSM	Evaluation	10	2,500
	MSVAMP	Evaluation	10	10,000

Experiment Setting

- Baselines
  - SFT, RFT, MAmmoTH, WizardMath, MetaMath (instruction-tuned with English instruction data)
  - MathOctopus (instruction-tuned with multilingual instruction data)
  - MonoReason (our reproduction of MetaMath)
  - MultiReason (our reproduction of MathOctopus)

Experiments

- Main results
  - Can question alignment stage brings improvement?
  - Can question alignment stage help us to beat translate-training baseline?
- Analysis
  - What if we use other translation data for stage 1 training?
  - What if we switch the order of two training stages?
  - What if we perform multi-task training instead of multi-stage training?
  - What is the benefit of achieving language alignment on answer consistency?
  - What is the connection between our approach with translate-testing?

Main Results

- Question alignment stage (QAlign) enables LLM's proficiency in English to be transferred to non-English tasks.

System (7B)	Bn	Th	Sw	Ja	Zh	De	Fr	Ru	Es	En	Avg.
SFT† (Touvron et al., 2023)	3.2	4.8	5.2	15.2	22.4	37.2	34.4	28.0	32.4	43.2	22.6
RFT† (Yuan et al., 2023)	2.4	2.0	2.8	6.8	16.8	33.6	34.0	29.2	34.0	44.8	20.6
MAmmoTH† (Yue et al., 2023)	3.6	4.8	2.4	10.8	17.2	33.2	32.8	26.0	32.4	49.6	21.3
WizardMath† (Luo et al., 2023)	2.0	4.0	3.4	24.0	22.4	30.4	30.4	30.8	34.8	47.6	23.0
MathOctopus† (Chen et al., 2023)	28.8	34.4	39.2	36.0	38.4	44.8	43.6	39.6	42.4	52.4	40.0
MetaMath (Yu et al., 2023)	6.4	4.0	3.2	39.2	38.8	56.8	52.8	47.2	58.0	63.2	37.0
MultiReason	26.8	36.0	36.8	33.2	42.4	42.8	40.8	42.4	42.8	47.2	39.1
MonoReason	7.6	5.6	5.2	34.0	45.2	54.0	**56.8**	51.6	58.8	65.5	38.4
QAlign + MonoReason (Ours)	**32.4**	**39.6**	**40.4**	**44.0**	**48.4**	**54.8**	**56.8**	**52.4**	**59.6**	**68.0**	**49.6**

System (13B)	Bn	Th	Sw	Ja	Zh	De	Fr	Ru	Es	En	Avg.
SFT† (Touvron et al., 2023)	6.0	6.8	7.6	25.2	32.8	42.8	40.8	39.2	45.2	50.4	29.7
RFT† (Yuan et al., 2023)	3.2	4.4	3.6	26.4	33.6	38.4	44.8	41.6	46.8	52.0	29.5
MAmmoTH† (Yue et al., 2023)	3.6	5.2	1.6	19.2	31.2	45.6	39.6	36.8	50.0	56.4	28.9
WizardMath† (Luo et al., 2023)	6.4	5.6	5.6	22.0	28.0	40.4	42.0	34.4	45.6	52.8	28.3
MathOctopus† (Chen et al., 2023)	35.2	46.8	42.8	43.2	48.8	44.4	48.4	47.6	48.0	53.2	45.8
MetaMath (Yu et al., 2023)	11.6	6.4	7.6	42.8	49.2	64.8	65.2	63.6	65.2	67.2	44.4
MultiReason	37.6	42.2	44.0	43.2	53.6	47.6	54.0	48.0	54.8	56.4	48.1
MonoReason	12.4	11.2	6.4	42.0	46.0	**64.0**	**62.4**	61.6	64.8	68.4	43.9
QAlign+ MonoReason (Ours)	**38.4**	**49.6**	**46.0**	**52.4**	**59.2**	62.0	**62.4**	**64.4**	**67.2**	**69.2**	**57.1**

mGSM

Main Results

- After question alignment, our fine-tuned LLM surpasses the translate-training baseline (MathOctopus/MultiReason) by a large margin.

System (7B)	Bn	Th	Sw	Ja	Zh	De	Fr	Ru	Es	En	Avg.
SFT† (Touvron et al., 2023)	3.2	4.8	5.2	15.2	22.4	37.2	34.4	28.0	32.4	43.2	22.6
RFT† (Yuan et al., 2023)	2.4	2.0	2.8	6.8	16.8	33.6	34.0	29.2	34.0	44.8	20.6
MAmmoTH† (Yue et al., 2023)	3.6	4.8	2.4	10.8	17.2	33.2	32.8	26.0	32.4	49.6	21.3
WizardMath† (Luo et al., 2023)	2.0	4.0	3.4	24.0	22.4	30.4	30.4	30.8	34.8	47.6	23.0
MathOctopus† (Chen et al., 2023)	28.8	34.4	39.2	36.0	38.4	44.8	43.6	39.6	42.4	52.4	40.0
MetaMath (Yu et al., 2023)	6.4	4.0	3.2	39.2	38.8	56.8	52.8	47.2	58.0	63.2	37.0
MultiReason	26.8	36.0	36.8	33.2	42.4	42.8	40.8	42.4	42.8	47.2	39.1
MonoReason	7.6	5.6	5.2	34.0	45.2	54.0	**56.8**	51.6	58.8	65.5	38.4
QAlign + MonoReason (Ours)	**32.4**	**39.6**	**40.4**	**44.0**	**48.4**	**54.8**	**56.8**	**52.4**	**59.6**	**68.0**	**49.6**

System (13B)	Bn	Th	Sw	Ja	Zh	De	Fr	Ru	Es	En	Avg.
SFT† (Touvron et al., 2023)	6.0	6.8	7.6	25.2	32.8	42.8	40.8	39.2	45.2	50.4	29.7
RFT† (Yuan et al., 2023)	3.2	4.4	3.6	26.4	33.6	38.4	44.8	41.6	46.8	52.0	29.5
MAmmoTH† (Yue et al., 2023)	3.6	5.2	1.6	19.2	31.2	45.6	39.6	36.8	50.0	56.4	28.9
WizardMath† (Luo et al., 2023)	6.4	5.6	5.6	22.0	28.0	40.4	42.0	34.4	45.6	52.8	28.3
MathOctopus† (Chen et al., 2023)	35.2	46.8	42.8	43.2	48.8	44.4	48.4	47.6	48.0	53.2	45.8
MetaMath (Yu et al., 2023)	11.6	6.4	7.6	42.8	49.2	64.8	65.2	63.6	65.2	67.2	44.4
MultiReason	37.6	42.2	44.0	43.2	53.6	47.6	54.0	48.0	54.8	56.4	48.1
MonoReason	12.4	11.2	6.4	42.0	46.0	**64.0**	**62.4**	61.6	64.8	68.4	43.9
QAlign+ MonoReason (Ours)	**38.4**	**49.6**	**46.0**	**52.4**	**59.2**	62.0	**62.4**	**64.4**	**67.2**	**69.2**	**57.1**

mGSM

Main Results

- Our fine-tuned LLMs also exhibit better robustness on the out-of-domain test set.

System (7B)	Bn	Th	Sw	Ja	Zh	De	Fr	Ru	Es	En	Avg.
SFT† (Touvron et al., 2023)	11.5	18.2	17.2	31.6	35.2	39.0	39.1	39.1	39.2	38.8	30.9
RFT† (Yuan et al., 2023)	7.7	16.9	14.9	33.9	34.9	40.8	41.5	39.5	42.5	42.7	31.3
MAmmoTH† (Yue et al., 2023)	4.3	6.3	4.2	26.7	26.8	39.6	39.9	33.7	42.9	45.1	26.3
WizardMath† (Luo et al., 2023)	16.1	17.0	10.3	37.9	36.3	39.2	37.7	37.4	44.8	48.5	32.5
MathOctopus† (Chen et al., 2023)	31.8	39.3	43.4	41.1	42.6	48.4	50.6	46.9	49.4	50.7	44.1
MetaMath (Yu et al., 2023)	14.2	17.8	16.5	53.2	53.1	61.4	60.7	58.9	61.2	65.5	46.3
MultiReason	27.6	36.5	42.4	40.9	43.2	44.3	46.7	42.3	45.5	48.0	41.3
MonoReason	15.0	17.1	15.4	51.9	54.4	60.9	62.2	59.3	**63.3**	**65.5**	46.2
QAlign + MonoReason (Ours)	**41.7**	**47.7**	**54.8**	**58.0**	**55.7**	**62.8**	**63.2**	**61.1**	**63.3**	65.3	**57.2**

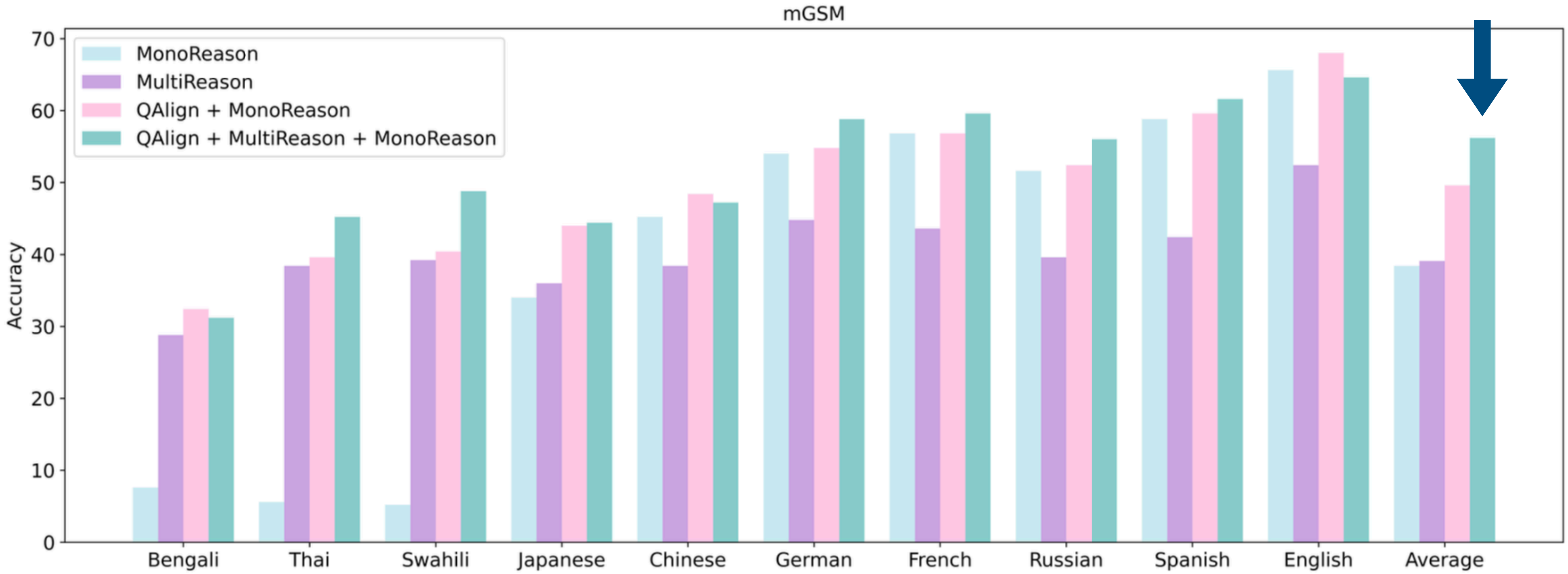
System (13B)	Bn	Th	Sw	Ja	Zh	De	Fr	Ru	Es	En	Avg.
SFT† (Touvron et al., 2023)	13.9	23.4	19.8	41.8	43.3	46.2	47.8	47.8	46.1	50.9	38.1
RFT† (Yuan et al., 2023)	12.2	24.8	19.4	42.4	42.3	45.1	45.2	46.5	45.6	47.1	37.1
MAmmoTH† (Yue et al., 2023)	5.0	13.7	12.9	42.2	47.7	52.3	53.8	50.7	53.9	53.4	38.6
WizardMath† (Luo et al., 2023)	13.7	16.3	12.5	29.5	37.0	48.7	49.4	43.8	49.4	56.3	35.7
MathOctopus† (Chen et al., 2023)	35.2	41.2	46.8	39.2	52.0	47.2	48.0	45.6	53.2	56.4	46.5
MetaMath (Yu et al., 2023)	14.6	15.7	17.4	57.0	56.6	67.3	64.7	63.7	65.9	67.7	49.1
MultiReason	35.0	41.3	44.6	49.9	48.1	53.3	53.2	51.6	52.5	54.5	48.4
MonoReason	20.6	20.5	19.1	57.0	58.8	68.4	**68.1**	**67.5**	**68.9**	**68.9**	51.8
QAlign + MonoReason (Ours)	**49.2**	**55.5**	**55.2**	**64.3**	**63.8**	**69.5**	**68.1**	66.4	66.4	67.6	**62.6**

mSVAMP

Main Results

- Incorporating multilingual supervised data into our framework can achieve a higher ceiling for multilingual performance.

+6.6%



Analysis

- We explore language alignment with other language directions, types and domains of data and confirm our intuition that in fact X-En questions perform best.

Data	Direction	mGSM Non-En	mGSM En	mSVAMP Non-En	mSVAMP En
Question	X→En	47.6	68.0	56.5	65.3
Question	En→X	36.2	68.0	48.3	64.4
Response	X→En	46.4	67.2	52.1	64.9
Response	En→X	42.8	68.0	49.0	63.9
Flores-101	X→En	36.3	68.0	46.8	65.4

Analysis

- When we reverse the training order, we observe that the model also tends to translate non-English questions instead of answering them.

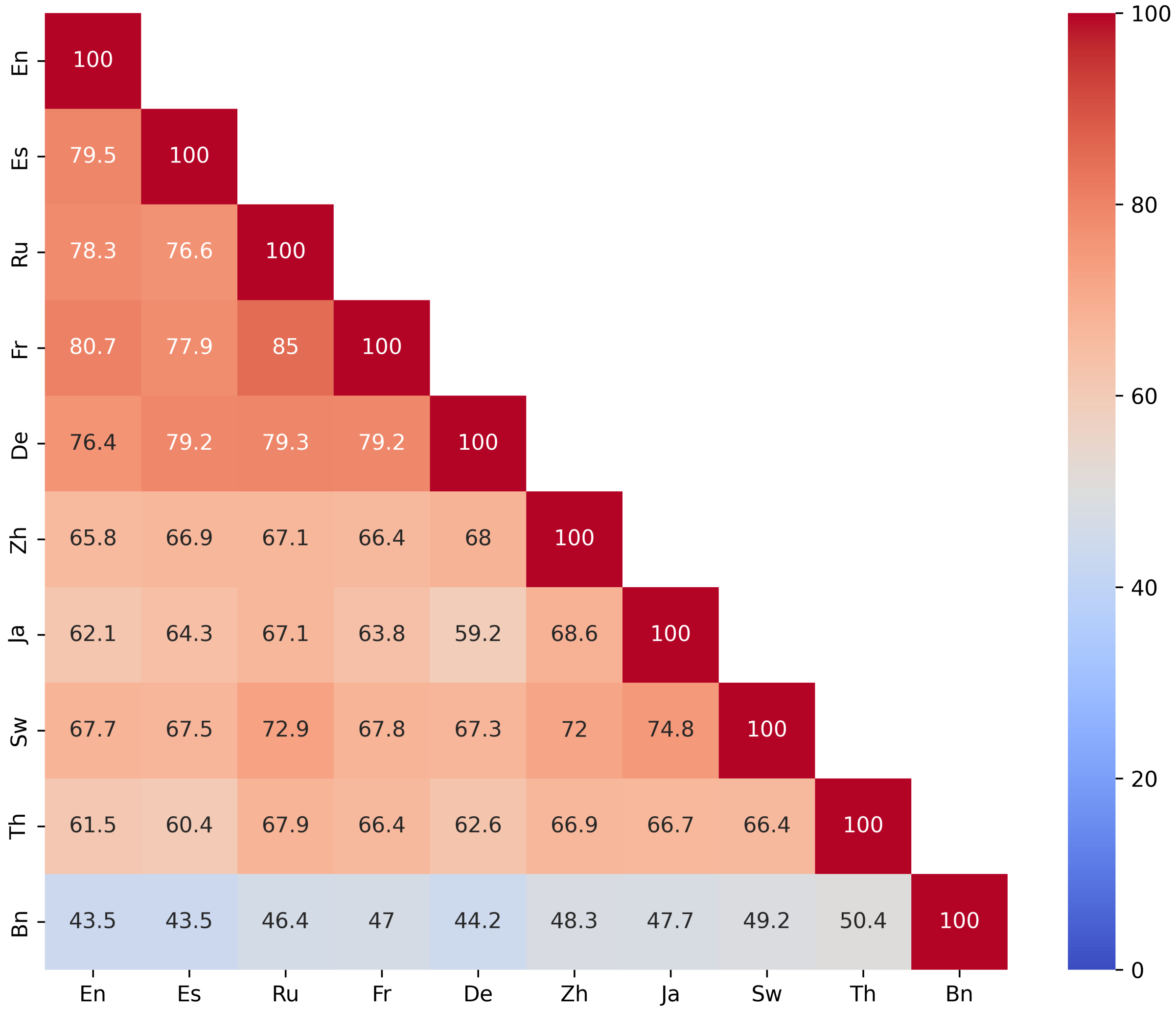
System	Bn	Th	Sw	Ja	Zh	De	Fr	Ru	Es	En	Avg.
QAlign -> MonoReason	32.4	39.6	40.4	44.0	48.4	54.8	56.8	52.4	59.6	68.0	49.6
MonoReason -> QAlign	2.0	2.0	2.8	2.0	2.0	2.0	1.6	2.0	2.0	2.8	2.1

- When we combine the data for multi-task training, we also observe that the model tends to translate non-English questions rather than answering them.

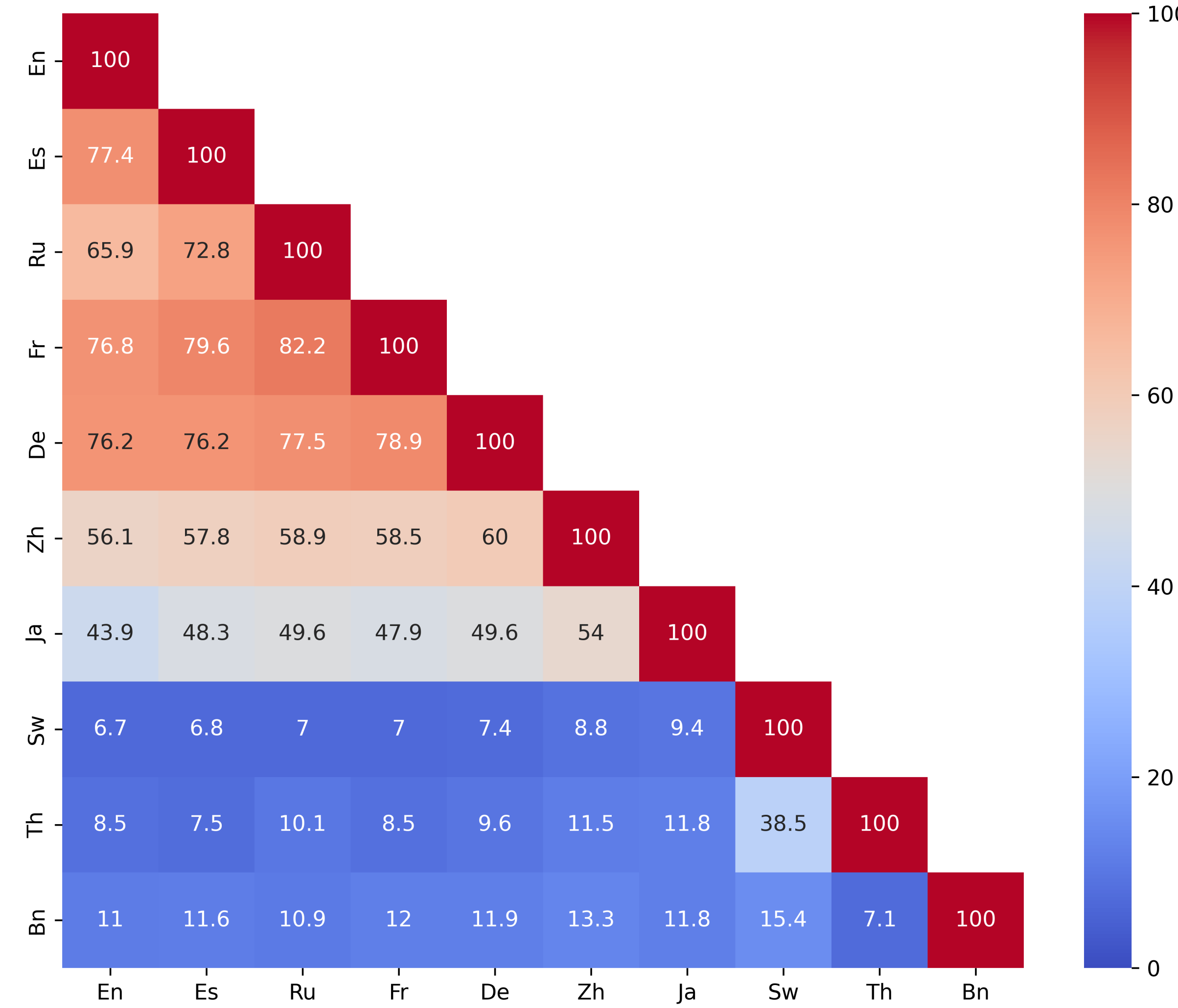
System	Bn	Th	Sw	Ja	Zh	De	Fr	Ru	Es	En	Avg.
QAlign -> MonoReason	32.4	39.6	40.4	44.0	48.4	54.8	56.8	52.4	59.6	68.0	49.6
QAlign + MonoReason	2.8	2.4	2.8	2.8	2.4	5.2	3.6	3.6	8.0	67.6	10.1

Analysis

- Another advantage of establishing language alignment is the improvement it brings to the consistency of predicted answers against multilingual queries.



w. QAlign



w.o. QAlign

Question (English)	Gold Answer
There are 27 unicorns left in the world. One third of them are in the Scottish Highlands. Two thirds of the Scottish unicorns are female. How many female Scottish unicorns are there?	6
Chain-of-Thought Response	**Prediction**
[MonoReason] One third of the unicorns are in the Scottish Highlands, so there are 27/3 = 9 unicorns in the Highlands. Two thirds of the Scottish unicorns are female, so there are 9 * 2/3 = 6 female Scottish unicorns. The answer is: 6	6 ✓
[QAlign + MonoReason] One third of the unicorns are in the Scottish Highlands, so there are 27/3 = 9 unicorns in the Highlands. Two thirds of the unicorns in the Highlands are female, so there are 9 * 2/3 = 6 female unicorns in the Highlands. Therefore, there are 6 female Scottish unicorns. The answer is: 6	6 ✓
Question (Chinese)	**Gold Answer**
世界上剩下27只独角兽，其中三分之一在苏格兰高地。 苏格兰高地的独角兽的三分之二是雌性。苏格兰雌性独角兽有多少只?	6
Chain-of-Thought Response	**Prediction**
[MonoReason] If three-fourths of the remaining 27 one-horned animals are in the Scottish Highlands, then there are 27 * 3/4 = 21 one-horned animals in the Scottish Highlands. If two-thirds of the one-horned animals in the Scottish Highlands are female, then there are 21 * 2/3 = 14 female one-horned animals in the Scottish Highlands. Therefore, there are 14 female one-horned animals in the Scottish Highlands. The answer is: 14	14 ✗
[QAlign + MonoReason] If there are 27 remaining unicorns in the world, and one-third of them are in the Scottish Highlands, then there are 27/3 = 9 unicorns in the Scottish Highlands. If two-thirds of the unicorns in the Scottish Highlands are female, then there are 9 * 2/3 = 6 female unicorns in the Scottish Highlands. Therefore, there are 6 female unicorns in the Scottish Highlands. The answer is: 6	6 ✓

Question Alignment vs. Translate-Test

- Translate-test approach does not always help, it requires the LLM to have great proficiency in both translation and reasoning.
- For our aligned MonoReason model, direct inference and translate-test achieves similar performance.

Method	mGSM		mSVAMP	
	Non-En	En	Non-En	En
MonoReason (7B)				
Direct Inference	35.4	65.5	47.6	68.9
Translate-test	30.8	-	42.3	-
QAlign+MonoReason (7B)				
Direct Inference	47.6	68.0	56.5	65.3
Translate-test	46.6	-	56.6	-

Effects of Using Different English Instruction Data

- Under different English instruction data, the incorporation of a question alignment stage always boost LLM's non-English performance.

Supervision	QAlign	mGSM Non-En	mGSM En	mSVAMP Non-En	mSVAMP En
GSM8K	✗	18.8	43.6	33.6	47.2
GSM8K	✓	26.3	41.6	36.8	47.0
MetaMathQA	✗	35.4	65.6	44.4	65.3
MetaMathQA	✓	47.6	68.0	56.5	65.3

Conclusion

- We present a novel question alignment fine-tuning step which performs targeted language alignment for best use of the LLMs English reasoning abilities.
- We fine-tune open-source LLMs, LLaMA2- 7B/13B into strong multilingual reasoners, which beat the translate-training baseline by a large-margin.
- Model, code and data have all been open-sourced:
  - arXiv: https://arxiv.org/abs/2401.07817
  - Github: https://github.com/NJUNLP/QAlign
  - Huggingface: https://huggingface.co/Wenhao97

Thank You!