

xCoT: Cross-lingual Instruction Tuning for Cross-lingual Chain-of-Thought Reasoning

Linzheng Chai¹, Jian Yang^{1*}, Tao Sun¹, Hongcheng Guo¹, Jiaheng Liu¹,
Bing Wang¹, Xinnian Liang¹, Jiaqi Bai², Tongliang Li³, Qiyao Peng⁴, Zhoujun Li¹

¹The State Key Laboratory of Complex & Critical Software Environment, Beihang University

²Cyberspace Institute of Advanced Technology, Guangzhou University

³Beijing Information Science and Technology University

⁴School of New Media and Communication, Tianjin University
{chailinzheng, jiaya}@buaa.edu.cn

Abstract

Chain-of-thought (CoT) has emerged as a powerful technique to elicit reasoning in large language models and improve a variety of downstream tasks. CoT mainly demonstrates excellent performance in English, but its usage in low-resource languages is constrained due to poor language generalization. To bridge the gap among different languages, we propose a cross-lingual instruction fine-tuning framework (xCoT) to transfer knowledge from high-resource languages to low-resource languages. Specifically, the multilingual instruction training data (xCoT-INSTRUCT) is created to encourage the semantic alignment of multiple languages. We introduce cross-lingual in-context few-shot learning (xICL) to accelerate multilingual agreement in instruction tuning, where some fragments of source languages in examples are randomly substituted by their counterpart translations of target languages. During multilingual instruction tuning, we adopt the randomly online CoT strategy to enhance the multilingual reasoning ability of the large language model by first translating the query to another language and then answering in English. To further facilitate the language transfer, we leverage the high-resource CoT to supervise the training of low-resource languages with cross-lingual distillation. Experimental results demonstrate the superior performance of xCoT in reducing the gap among different languages, highlighting its potential to reduce the cross-lingual gap.

Introduction

Recent advancements in Large Language Models (LLMs) (Touvron et al. 2023; Touvron and Louis Martin 2023; Patel et al. 2023; OpenAI 2023; Bai et al. 2023; Hui et al. 2024) in natural language processing (NLP) have intensively engaged the interests of researchers. LLMs (Wei et al. 2022c; Zhang et al. 2023; Kojima et al. 2022) are further equipped with the chain-of-thought (CoT) technique to gain impressive performance in complex reasoning tasks, where LLMs first produce intermediate reasoning steps and infer the final answer.

However, existing studies related to the CoT methods are mainly constrained in high-resource languages (e.g. English) and deliver little consideration into multilingual sce-

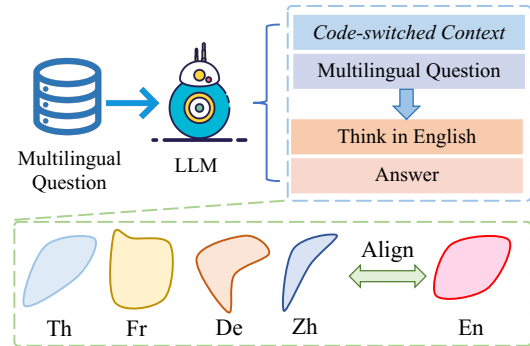


Figure 1: Illustration of xCoT. The cross-lingual instruction tuning is used to align representations of different languages.

narios. Recent works (Shi et al. 2023; Qin et al. 2023; Chen et al. 2023) endeavor to simply use prompt engineering to improve the language generalization ability of the model without any fine-tuning. These prompt-based methods ignore the potential of representation-based cross-lingual alignment derived from the cross-lingual supervised fine-tuning (cross-lingual SFT). Supervised fine-tuning has been shown to perform at a satisfactory level across various tasks, such as FLAN (Wei et al. 2022a) and InstructGPT (Ouyang et al. 2022). Therefore, *how to encourage cross-lingual alignment in supervised fine-tuning still requires further exploration.*

To minimize the gap among different languages, we propose a Cross-lingual Chain-of-Thought reasoning (xCoT) framework using cross-lingual supervised instruction fine-tuning. Specifically, we first construct the multilingual instruction training data (xCoT-INSTRUCT) by translating English to other languages. Then, we randomly substitute some fragments of source languages in examples by their counterpart translations of target languages. To transfer high-resource languages to low-resource languages, we mix the tokens of the source and target language in the same query to enable the LLMs to handle different languages. The code-switched examples and the query can be applied to cross-lingual in-context learning in supervised instruction tuning. During multilingual instruction tuning, we adopt

* Corresponding Author.

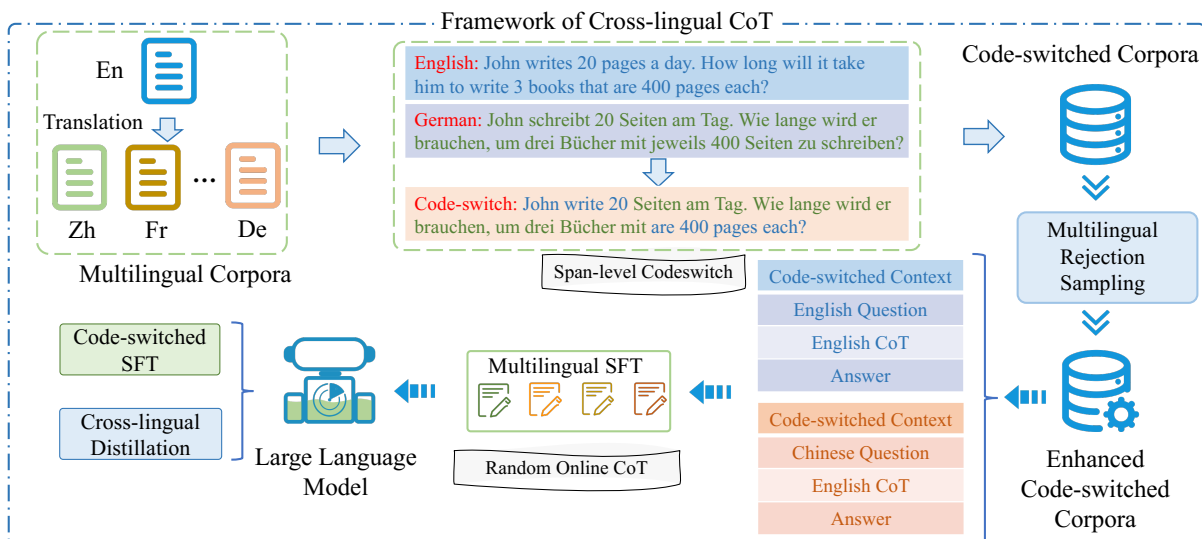


Figure 2: Overview of xCoT. The cross-lingual in-context few-shot learning encourages multilingual alignment in instruction tuning, where the query in the example is mixed with different language tokens. During multilingual instruction tuning, the randomly online CoT strategy is used to promote the multilingual reasoning ability of LLM and then answer in English. Finally, we leverage the high-resource CoT to supervise the training of low-resource languages with cross-lingual distillation.

the randomly online CoT strategy to enhance the multilingual reasoning ability of the large language model by first translating the query to another language and then answering in English. To further facilitate the language transfer, we leverage the high-resource CoT to supervise the training of low-resource languages with cross-lingual distillation. Experimental results on previous benchmarks demonstrate the superior performance of xCoT in reducing the gap among different languages, highlighting its potential to reduce the cross-lingual gap.

Extensive experiments of xCoT are evaluated on multilingual benchmarks MGSM of 11 languages and MSVAMP of 10 languages. The results demonstrate that our proposed method consistently achieves state-of-the-art performance across all languages, notably surpassing strong baseline by an average margin of 15%. The contributions in this work are summarized as follows: (1) We construct the multilingual instruction data to transfer knowledge of high-resource languages into low-resource languages. The training data is further augmented by cross-lingual in-context learning, where a piece of code-switched demonstration context and the current query are concatenated as the input for LLM. (2) During training, we propose the random online CoT (Random-CoT), which first randomly translates the query into other languages and then answers in English. (3) To align the representations of different languages, we propose cross-lingual knowledge to align the output distribution given the queries of different languages using Kullback–Leibler divergence.

Cross-lingual CoT Reasoning

Given the query $q = (q_1 \dots, q_m)$ of language L_i , the large language model (LLM) \mathcal{M} outputs the corresponding an-

swer $a = (a_1, \dots, a_n)$ of language L_j , where m and n are lengths of prompt and answer in a sample (q, a) . L_i and L_j are source and target language, where $L_{all} = \{L_k\}_{k=1}^K$ and K is number of languages. LLM further enhances the task performance by chain-of-thought reasoning, where the chain-of-thought examples of sequences $c = (c_1, \dots, c_t)$ are added into the exemplars of prompting. The high-quality rationales c comprised of a series of intermediate natural language reasoning steps provide helpful suggestions for the final output. Given multiple chain-of-thought examples as demonstrations and the original prompt q of the target language as a whole, the problem definition of cross-lingual CoT is described as:

$$P(a|q, c) = \prod_{j=1}^n P(a_j|a_{<j}; q, c, \mathcal{M}) \quad (1)$$

where q (question) and c (corresponding exemplars) are concatenated as a whole p to predict the answer denoted as $P(a|p)$. Driven by the CoT demonstrations c , the LLM first generates the intermediate steps and then outputs the final answer a .

xCoT

Model Overview

Figure 2 describes the overall framework of our method xCoT. Specifically, cross-lingual in-context few-shot learning (xICL) encourages multilingual alignment in instruction tuning, where the query in the example is mixed with different language tokens. During multilingual instruction tuning, the randomly online CoT strategy (Random-CoT) is used to promote the multilingual reasoning ability of LLM and then answer in English. Finally, we leverage the high-

Algorithm 1: Random Online CoT

Input: Multilingual Instruction Dataset: D ;
Multilingual LLM: \mathcal{M} ;
Maximum supervised fine-tuning step: T ;
Batch size: B ;
Target language set: $\mathcal{L}_{all} = \{L_k\}_{k=1}^K$;
Output: Fine-tuned LLM: \mathcal{M}

```
1  $t \leftarrow 0$ 
2 while  $t \leq T$ ; do
3   Random sampled batch  $\mathcal{B} \in D$ 
4   for  $k \leftarrow 1$  to  $B$ ; do
5      $(c^{L_i,j}, q^{L_i}, a^{L_j}) \leftarrow \mathcal{B}$ 
6      $L_k \sim U(\mathcal{L}_{all}) (L_k \neq L_j)$ 
7      $q^{L_k} \leftarrow \mathcal{M}([c^{L_i,j}, q^{L_i}])$ 
      // Translate  $q^{L_i} \rightarrow q^{L_k}$ 
8      $a^{L_j} \leftarrow \mathcal{M}([c^{L_i,j}, q^{L_i}, q^{L_k}])$ 
      // Answer in language  $L_j$ 
9      $\mathcal{B} \leftarrow \mathcal{B} \cup (q^{L_k}, q^{L_k}, a^{L_j})$ 
10  Optimize  $\mathcal{M}$  with  $\mathcal{B}$ 
11   $t \leftarrow t + 1$ 
12 return  $\mathcal{M}$ 
```

resource CoT to supervise the training of low-resource languages with cross-lingual distillation.

XCOT-INSTRUCT

Data Construction We create a new multilingual instruction dataset (XCOT-INSTRUCT) for cross-lingual chain-of-thought reasoning, which can be used as the training corpora for multilingual benchmarks, such as MGSM (Shi et al. 2023) and MSVAMP (Chen et al. 2023). We use the multilingual translator to expand the English instruction data GSM8K into other 10 languages, including German, French, Spanish, Russian, Chinese, Japanese, Thai, Telugu, Bengali, and Swahili. The instruction dataset of each language contains 7.4K samples, where we only translate the query into other languages and retain the response in English to facilitate the cross-lingual transfer. Finally, we obtain the multilingual instruction data $D = \{D^{L_k}\}_{k=1}^K$ and $(q^{L_i}, c^{L_i}, a^{L_j}) \in D^{L_i}$, where D^{L_i} is the SFT training data of language L_i and the number of the languages is K . D^{L_i} contains query q^{L_i} and response a^{L_j} with the corresponding context c^{L_i} . q^{L_i} is the query of source language and a^{L_j} is the response of the high-resource language (L_j is English in our work). $c^{L_i} = \{q_b^{L_i}, a_b^{L_j}\}_{b=1}^B$ is the context demonstration comprised of B queries of language L_i and the responses of L_j . For each language, we construct about 22K data context demonstration samples.

Cross-lingual Instruction Tuning Given the cross-lingual instruction corpora $D = \{D^{L_k}\}_{k=1}^K$, where D contains K languages and $L_{all} = \{L_k\}_{k=1}^K$. The LLM is jointly trained on the union of the multilingual corpora D :

$$\mathcal{L}_x = - \sum_{i=1}^K \mathbb{E}_{c^{L_i}, q^{L_i}, a^{L_j} \sim D^{L_i}} [\log P(a^{L_j} | q^{L_i}, c^{L_i}; \mathcal{M})] \quad (2)$$

where q^{L_i} is the query of the language L_i and a^{L_j} is the response of language L_j .

Cross-lingual In-context Learning (xICL)

To encourage cross-lingual alignment across different languages, we construct the code-switched query by replacing the spans of the source query with the counterparts of the target language.

Given a bilingual query (q^{L_i}, q^{L_j}) with the source language query $q^{L_i} = \{q_1^{L_i}, \dots, q_m^{L_i}\}$ of m tokens and the target translation $q^{L_j} = \{y_1^{L_j}, \dots, y_n^{L_j}\}$ of n tokens, we create the code-switched sequence $q^{L_i,j}$ by substituting the phrase $q_{u_1:u_2}^{L_i}$ with counterpart translation $q_{v_1:v_2}^{L_j}$, where $q_{v_1:v_2}^{L_j}$ is the target translation of source piece $q_{u_1:u_2}^{L_i}$. $q_{u_1:u_2}^{L_i}$ denotes the phrase in q^{L_i} from the u_1 -th token to the u_2 -th token ($1 \leq u_1 \leq u_2 \leq m$) and $q_{v_1:v_2}^{L_j}$ denotes the phrase in q^{L_j} from the v_1 -th token to the v_2 -th token ($1 \leq v_1 \leq v_2 \leq n$). For each phrase in code-switched in $q^{L_i,j}$, it comes from source phrase $q_{u_1:u_2}^{L_i}$ or target phrase $q_{v_1:v_2}^{L_j}$. The proportion of the source words in the code-switched sequence $q^{L_i,j}$ is denoted as α . $q_{i,j}^{L_i}$ contains $q^{L_i/j}$ (source sentence with target tokens) and $q^{L_j/i}$ (target sentence with source tokens).

Specifically, the code-switched sequence can be created in two ways: (1) $q^{L_i/j}$ (source sentence with target tokens): most tokens in $q^{L_i/j}$ derive from q^{L_i} , where some source phrases $q_{u_1:u_2}^{L_i}$ are substituted by their target counterpart phrases $q_{v_1:v_2}^{L_j}$ ($\alpha \geq 0.5$). (2) $q^{L_j/i}$ (target sentence with source tokens): most tokens in $q^{L_j/i}$ derive from q^{L_j} where some target phrases $q_{v_1:v_2}^{L_j}$ are substituted by their source counterpart phrases $q_{u_1:u_2}^{L_i}$ ($\alpha < 0.5$).

Random Online CoT(Random-CoT)

To force the model to understand the multilingual queries, we introduce the random online CoT, which first prompts the LLM to translate the query q^{L_i} to another language query q^{L_k} and then answer in a^{L_j} during the LLM tuning. Algorithm 1 describes the detail of Random-CoT, where given the training instance $(c^{L_i,j}, q^{L_i}, a^{L_j}) \in D$, we uniformly sample an intermediate language L_k ($L_k \neq L_i$) and prompt LLM first to translate q^{L_i} to q^{L_k} . Although L_j may belong to low-resource languages and the quality of q^{L_k} may be poor initially, our method still benefits from the translation signal of $q^{L_i} \rightarrow q^{L_k}$ by aligning the representations of different languages.

Cross-lingual Distillation (xDistill)

mSampling To further augment the cross-lingual instruction tuning, we refer to the RFT(Yuan et al. 2023), use the fine-tuned LLM \mathcal{M} to generate the synthetic response of the multilingual queries and then select correct reasoning paths as the augmented dataset D' . Finally, our model is trained on the original and augmented dataset $D \cup D'$.

From High-Resource to Low-Resource We use the high-resource samples to supervise the low-resource samples. Given the parallel high-resource sample $(c^{L_i,j}, q^{L_i}, a^{L_j})$

Method	Base Model	En	De	Fr	Es	Ru	Zh	Ja	Th	Te	Bn	Sw	Avg.
<i>Closed-Source Models</i>													
Native-CoT [†]	GPT-3.5	67.2	62.0	59.2	61.2	50.4	52.8	46.8	15.6	–	7.6	40.0	46.3
Native-CoT [†]	GPT-4	80.0	73.6	72.0	71.2	64.0	70.0	71.6	40.4	–	17.6	64.4	62.5
<i>Open-Source Models (7B)</i>													
Llama-2 [†] (Touvron and Louis Martin 2023)	Llama-2	43.2	37.2	34.4	32.4	28.0	22.4	15.2	4.8	–	3.2	5.2	22.6
RFT [†] (Yuan et al. 2023)	Llama-2	44.8	33.6	34.0	34.0	29.2	16.8	6.8	2.0	–	2.4	2.8	20.6
MAmmoTH [†] (Yue et al. 2023)	Llama-2	49.6	33.2	32.8	32.4	26.0	17.2	10.8	4.8	–	3.6	2.4	21.3
WizardMath [†] (Luo et al. 2023)	Llama-2	47.6	30.4	30.4	34.8	30.8	22.4	24.0	4.0	–	2.0	3.4	23.0
MathOctopus [†] (Chen et al. 2023)	Llama-2	54.8	43.6	38.0	45.2	48.4	45.2	35.6	36.4	–	33.2	38.4	41.9
xCoT	Bloom	30.0	30.4	28.8	32.4	33.6	30.0	29.6	28.4	28.4	33.2	26.8	30.1
xCoT	Llama-2	48.4	47.2	49.6	48.8	50.0	50.0	50.0	49.2	42.8	40.4	48.4	47.7
<i>Open-Source Models (13B)</i>													
LLama-2 [†] (Touvron and Louis Martin 2023)	Llama-2	50.4	42.8	40.8	45.2	39.2	32.8	25.2	6.8	–	6.0	7.6	29.7
RFT [†] (Yuan et al. 2023)	Llama-2	52.0	38.4	44.8	46.8	41.6	33.6	26.4	4.4	–	3.2	3.6	29.5
MAmmoTH [†] (Yue et al. 2023)	Llama-2	56.4	45.6	39.6	50.0	36.8	31.2	19.2	5.2	–	3.6	1.6	28.9
WizardMATH [†] (Luo et al. 2023)	Llama-2	52.8	40.4	42.0	45.6	34.4	28.0	22.0	5.6	–	6.4	5.6	28.4
MathOctopus [†] (Chen et al. 2023)	Llama-2	51.6	49.2	49.6	53.2	47.6	51.2	39.6	46.0	–	42.0	46.0	47.6
xCoT	Llama-2	54.4	52.4	46.4	54.8	56.8	54.0	49.6	50.0	47.2	50.0	51.6	51.5

Table 1: Multilingual evaluation results on the MGSM benchmark. †: Results from (Chen et al. 2023), for MathOctopus, we uniformly report the performance under xRFT and parallel-training settings.

Method	Base Model	En	De	Fr	Es	Ru	Zh	Ja	Th	Bn	Sw	Avg.
<i>Closed-Source Models</i>												
Native-CoT [†]	GPT-3.5	81.2	73.9	78.2	74.6	70.9	78.4	74.0	46.0	14.4	68.4	66.0
Native-CoT [†]	GPT-4	80.1	78.1	83.9	81.5	77.9	78.9	74.8	68.1	31.2	75.7	73.0
<i>Open-Source Models (7B)</i>												
Llama-2 [†] (Touvron and Louis Martin 2023)	Llama-2	38.8	39.0	39.1	39.2	39.1	35.2	31.6	18.2	11.5	17.2	30.9
RFT [†] (Yuan et al. 2023)	Llama-2	42.7	40.8	41.5	42.5	39.5	34.9	33.9	16.9	7.7	14.9	31.5
MAmmoTH [†] (Yue et al. 2023)	Llama-2	45.1	39.6	39.9	42.9	33.7	26.8	26.7	6.3	4.3	4.2	27.0
WizardMath [†] (Luo et al. 2023)	Llama-2	48.5	39.2	37.7	44.8	37.4	36.3	37.9	17.0	16.1	10.3	32.5
MathOctopus [†] (Chen et al. 2023)	Llama-2	46.8	43.1	45.3	44.5	42.1	43.2	43.2	40.5	32.8	42.3	42.4
xCoT	Llama-2	47.3	44.3	42.8	44.0	43.9	43.4	43.7	39.8	38.0	41.7	42.9

Table 2: Multilingual evaluation results on the MSVAMP benchmark. †: Results from (Chen et al. 2023), for MathOctopus, we uniformly report the performance under xRFT and parallel-training settings.

and low-resource sample $(c^{L_{k,j}}, q^{L_k}, a^{L_j})$, the model separately predict the target distribution $P_{high} = p(a^{L_j} | c^{L_{i,j}}, q^{L_i})$ and $P_{low} = p(a^{L_j} | c^{L_{k,j}}, q^{L_k})$. Since q^{L_i} and q^{L_k} are semantically equal, we can use distribution P_{high} to supervise P_{low} in token level:

$$\mathcal{L}_d = -\frac{1}{n} \sum_{t=1}^n [P_{high}^t \log P_{low}^t] \quad (3)$$

where P_{high}^t and P_{low}^t is the t -token distribution in answer. Through the token-level cross-lingual supervision, we transfer the knowledge to low-resource languages.

Experiments

Cross-lingual Supervised Fine-tuning

For each question in the dataset, we randomly select 2 other questions and corresponding answers as the context. We set $0 < \alpha < 1$ with a 0.8 replacement threshold to perform the code-switch operation on the question in context. For the selection of language in the codeswitch operation, we use English as the source language, and the language corresponding to the question is used as the target language. For

the English question, we randomly select another language as the target language.

We implement our model based on Llama-2-7B, Llama-2-13B, and Bloom-7b1. We finetune these models with 3 epochs and use a cosine scheduler with a learning rate of $2e-5$ and set 3% warm up. For cross-lingual distillation, the weight β of the distillation loss is set to 0.3.

Evaluation

To comprehensively assess the cross-lingual proficiency of xCoT, we evaluate the method using the MGSM (Shi et al. 2023) benchmark, which extends the English GSM8K (Cobbe et al. 2021) dataset into ten typologically varied languages through the manual translation of problems. To conduct a thorough and wide-ranging evaluation of the multilingual mathematical problem-solving skills, we also evaluate our method on MSVAMP (Chen et al. 2023), originating from the SVAMP (Patel, Bhattamishra, and Goyal 2021) dataset. This dataset incorporates mathematical problems in 10 different languages, initially translated using machine translation and subsequently refined through careful

k	Llama-2-7B					Llama-2-13B					Avg.
	De	Fr	Es	Ru	Zh	De	Fr	Es	Ru	Zh	
10	1.68	1.67	1.68	1.80	1.69	1.98	1.96	1.98	1.98	1.96	7.22
20	2.58	2.56	2.59	2.66	2.59	2.95	2.95	2.97	2.97	2.94	11.02
30	3.21	3.24	3.21	3.35	3.22	3.70	3.71	3.70	3.73	3.69	13.93
50	4.32	4.29	4.34	4.47	4.35	4.93	4.93	4.91	4.72	4.91	18.43

Table 3: Distinct reasoning paths of each language with different sampling times. Different reasoning paths per question are generated by different SFT models with different k .

human review and correction for accuracy and nuance. Our method is evaluated on MGSM and MSVAMP with the accuracy metric. In the experiments, we report the accuracy of all methods.

Baselines

xCoT are mainly compared with: (1) close-source LLM GPT-3.5, GPT-4; (2) open-source models Llama-2 and Bloom. xCoT primarily conduct experiments based on Llama-2 and compare with other Llama-2-based methods RFT, MathOctopus, MAMmoTH, WizardMath, etc. Furthermore, we select Bloom as the base model to explore the performance of xCoT when combined with multilingual LLM.

Main Results

MGSM Table 1 presents the results of our method and previous baselines on MGSM of 11 languages, including En, De, En, Fr, Es, Ru, Zh, Ja, Th, Te, Bn, Sw. Compared to the open-source baseline Llama-2, MAMmoTH (Yue et al. 2023) trained with a union of chain-of-thought (CoT) and program-of-thought (PoT) gains strong improvement. Our method significantly outperforms the previous strong baseline MAMmoTH by an average of points. It can prove that our method can leverage cross-lingual in-context learning (xICL), random online CoT (Random-CoT), and cross-lingual distillation (xDistill) to encourage alignment across different languages.

MSVAMP Table 2 compares the performance of our method with previous relevant methods on MSVAMP of 10 languages. The recent strong multilingual baseline MathOctopus beats the previous baselines MAMmoth and WizardMath with the help of the multilingual instruction dataset MGSM8KInstruct. Further, our proposed method gains the best performance of 42.9 points in 7B level across all languages, demonstrating that our proposed framework strengthens transferability from the high-resource languages to all other languages.

Analysis

Ablation Study To verify the effectiveness of each module in our method, we conduct an ablation study by gradually adding modules. In addition, Llama-2-7B+Translate serves as the baseline model, fine-tuned only on GSM8K. At the time of testing, MGSM data was first translated into English using Google Translate. The multilingual LLM Llama-7B is first trained on the multilingual corpora xCoT-INSTRUCT,

where the model is denoted as ①. Compared to the initial model ①, the model ② with the code-switched context in multilingual tuning gains the improvement of +4.7 points on average, which shows the usage of xICL in encouraging alignment across different languages. Then, the model ③ is further enhanced with mSampling by a large margin +5.8 points, where the model generates the multilingual responses and chooses correct reasoning paths as the augmented dataset. During multilingual tuning, our method adopts Random-CoT to first translate the query to another language and then answer in English. For the output distribution, the high-resource distribution is used to supervise the low-resource distribution (xDistill). Putting them all together, we obtain the final model xCoT (⑤) with 47.7 points. Table 4 summarizes the results of the ablation study of cross-lingual transfer in different parts, which emphasizes the effectiveness of cross-lingual transfer that can gradually improve performance in different aspects. Compared with the Google Translate method, xCoT has significant improvements.

Cross-lingual Prompting To trigger the cross-lingual potential capability of LLM, we introduce xICL to force the model to understand the multilingual queries and align their representations. To advance multilingual agreement in instruction tuning, we randomly replace some fragments of source languages in examples with their counterpart translations of target languages for cross-lingual in-context few-shot learning (xICL). Table 6 shows the results of xCoT with English context, native context, and code-switched context on different backbones. The query mixed with different language tokens brings significant improvement in various languages.

Cross-lingual Reasoning Path Our multilingual instruction data is augmented by multilingual sampling, where the fine-tuned LLM generates the response and selects the correct path. Table 3 shows that different languages have a similar number of reasoning paths, which proves that using the cross-lingual CoT successfully transfers reasoning patterns from one language to another language. xCoT can accumulate all reasoning paths to improve the model performance.

Multilingual Representations We randomly select 250 parallel queries with their 2-shot examples of each language in xCoT-INSTRUCT and visualize their representations (Maaten and Hinton 2008) of the last Llama decoder layer in Figure 3 using our multilingual model fine-tuned on xCoT-INSTRUCT and the multilingual baseline. The first hidden state of the layers is adopted as the sentence representation. Compared to Figure 3(a) of the baseline, different languages become closer and more likely to overlap with each other in Figure 3(b) of our method, demonstrating that our method effectively aligns representations of different languages to the shared space.

Understanding and Reasoning in English After the cross-lingual SFT with Random-CoT, xCoT chooses the high-resource language (English) as the auxiliary language to understand and answer the non-English question. In Figure 4, our method uses “Let’s think step by step in *En-*

ID	Method	En	De	Fr	Es	Ru	Zh	Ja	Th	Te	Bn	Sw	AVG
-	Llama-2-7B + Translate	42.4	39.2	36.8	39.2	40.0	39.6	38.0	32.0	36.0	39.2	40.4	38.4
①	Llama-2-7B	38.4	37.2	37.6	39.2	37.6	30.4	29.6	26.8	13.2	23.2	18.4	30.1
②	+ <i>xCL</i>	45.6	38.4	33.6	36.4	40.4	37.2	32.0	32.0	26.0	30.0	32.0	34.8
③	+ <i>mSampling</i>	50.8	41.2	42.0	44.4	42.0	40.4	38.0	40.8	35.6	34.8	37.6	40.6
④	+ <i>Random-CoT</i>	48.0	50.4	46.0	48.4	46.8	51.2	46.4	47.2	41.6	41.2	46.8	46.7
⑤	+ <i>xDistill</i>	48.4	47.2	49.6	48.8	50.0	50.0	50.0	49.2	42.8	40.4	48.4	47.7

Table 4: Ablation study on MGSM based on Llama-2-7b. ⑤ is the final model of our method(xCoT).

Random-CoT Direction	En	De	Fr	Es	Ru	Zh	Ja	Th	Te	Bn	Sw	Avg.
$L_{all} \rightarrow En$	50.8	41.2	42.0	44.4	42.0	40.4	38.0	40.8	35.6	34.8	37.6	40.6
$L_{low} \rightarrow En \cup L_{high} \rightarrow L_{high}$	50.8	45.6	49.6	46.4	49.6	41.2	42.4	42.8	37.6	38.8	42.8	44.3
$L_{all} \rightarrow L_{all}$	46.0	46.4	45.2	50.0	48.0	50.4	47.6	43.2	40.8	40.4	45.2	45.7
$L_{all} \rightarrow L_{high}$	48.4	47.2	49.6	48.8	50.0	50.0	50.0	49.2	42.8	40.4	48.4	47.7

Table 5: Study on translation direction of Random-CoT. L_{all} , L_{high} and L_{low} denote all languages, high-resource languages, and low-resource languages, respectively.

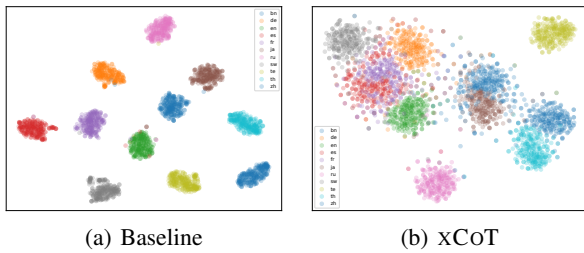


Figure 3: (a) and (b) are representations of Llama-7B and our method from the last decoder layer. Each color denotes one language (11 languages in MGSM).

“*glish*” to answer the English question. For the non-English question, we adopt “Let’s think the question in $\{Language\}$ and then think step by step in *English*”, where $\{Language\}$ can be high-resource languages in SFT tuning but we set $\{Language\}$ as English during inference stage. To effectively transfer knowledge from high-resource to low-resource languages, we force LLM to understand the query in English and then think in English.

Analysis in Random-CoT To facilitate the alignment among different languages, the question of language L_{i_1} is first translated into another language L_{i_2} in SFT tuning. Given the query of language L_{i_1} , we can translate another language L_{i_2} ($L_{i_1} \neq L_{i_2}$). L_{all} denotes all languages; L_{high} contains En, De, Fr, Es, Ru, Zh, Ja; L_{low} contains Th, Te, Bn, Sw. For example, the strategy “ $L_{all} \rightarrow L_{high}$ ” denotes that the $L_{i_1} \in L_{all} \wedge L_{i_2} \in L_{high}$. Table 5 shows the results of our method with different Random-CoT strategies and the strategy “ $L_{all} \rightarrow L_{high}$ ” gets the best performance, which can be attributed the language transfer from high-resource to low-resource languages.

Data Efficiency Figure 5 plots the multilingual evaluation results of xCoT with different SFT data sizes. We observe that our method with nearly 20% SFT data can still beat the strong baseline Llama-2-7B, which can attributed to the mutual reinforcement of multiple languages.

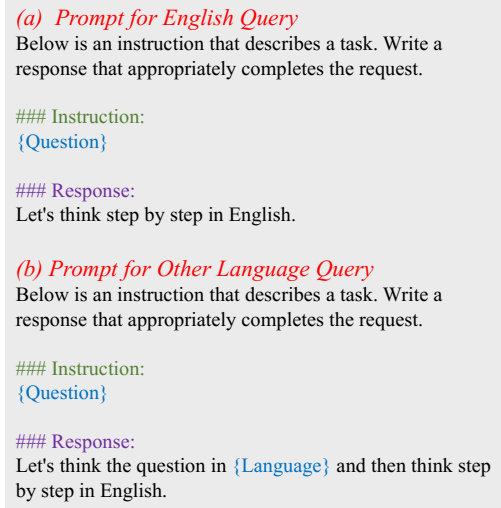


Figure 4: The prompt of Random-CoT.

Cases Study Given the queries of different languages, our method prompts LLM to first consider the multilingual query in English and then answer in English. Table 7 shows examples of the Spanish and German baseline. We observe that Llama-2 tends to generate incorrect answers for non-English queries. For the German Question “Jimmy hat 2 \$ mehr als doppelt so viel Geld wie Ethel. Wenn Ethel 8 \$ hat, wie viel Geld hat dann Jimmy im Moment?”, our method first thinks the non-English query in English “Question: Jimmy has \$2 more than twice the money Ethel has. If Ethel has \$8, how much money does Jimmy have now?”. and then answer in English. It proves that our method can align both query and response across different languages.

Related Work

Large Language Models

Large language models (LLMs) have shown great power in NLP tasks, and as the scale of the model gets larger, LLMs emerge with surprising capabilities (Touvron and

Base Model	CoT	En	De	Fr	Es	Ru	Zh	Ja	Th	Te	Bn	Sw	AVG
Bloom-7B	En-Context	32.8	31.2	30.4	30.0	34.0	29.6	28.0	27.6	27.2	30.4	28.4	29.9
	Native-Context	32.8	30.0	25.6	32.0	31.2	30.0	25.2	28.8	25.6	30.0	30.4	29.2
	Codeswitch-Context	30.0	30.4	28.8	32.4	33.6	30.0	29.6	28.4	28.4	33.2	26.8	30.1
Llama-2-7B	En-Context	50.0	46.4	50.0	47.2	47.6	50.4	47.6	44.8	42.0	39.2	48.4	46.7
	Native-Context	50.0	48.4	44.0	48.8	47.6	46.4	45.6	45.6	42.0	41.2	44.4	45.8
	Codeswitch-Context	48.4	47.2	49.6	48.8	50.0	50.0	50.0	49.2	42.8	40.4	48.4	47.7
Llama-2-13B	En-Context	50.8	50.4	48.0	50.4	55.6	48.4	48.0	54.0	44.8	45.6	49.2	49.5
	Native-Context	50.8	52.0	49.2	49.2	56.0	50.0	49.6	52.8	46.0	45.2	48.8	49.9
	Codeswitch-Context	50.8	51.6	49.2	50.8	56.0	50.0	48.8	54.0	44.8	46.8	50.0	50.2

Table 6: The results of xCoT with different context prompting strategies.

Spanish Question	Dan planta 3 rosales. Cada rosal tiene 25 rosas. Cada rosa tiene 8 espinas. ¿Cuántas espinas hay en total?
Llama-2	There are $3 \times 25 = 75$ roses. There are $75 \times 8 = 600$ thorns. There are $600 \times 2 = 1200$ forks. The answer is 1200. ✗
xCoT	Question: Dan plants 3 rose bushes. Each rose bush has 25 roses. Each rose has 8 thorns. How many thorns are there in total? Let's think step by step in English. There are $25 \times 3 = 75$ roses. So there are $75 \times 8 = 600$ thorns. The answer is 600. ✓
German Question	Jimmy hat 2 \$ mehr als doppelt so viel Geld wie Ethel. Wenn Ethel 8 \$ hat, wie viel Geld hat dann Jimmy im Moment?
Llama-2	Jimmy's money is $2 * 8 = 16$. The answer is 16. ✗
xCoT	Question: Jimmy has \$2 more than twice the money Ethel has. If Ethel has \$8, how much money does Jimmy have now? Let's think step by step in English. Twice the money Ethel has is $2 * 8 = 16$ dollars. Jimmy has 2 more than twice the money Ethel has, which means he has $16 + 2 = 18$ dollars. The answer is 18. ✓

Table 7: Cases study on Llama-2 and our proposed method within 7B-level on MGSM.

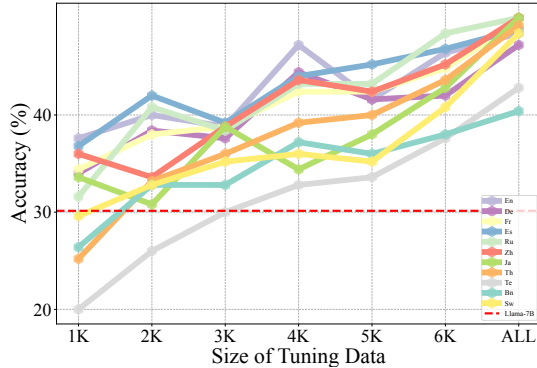


Figure 5: Multilingual evaluation results on MGSM with different SFT data size.

Louis Martin 2023; Wei et al. 2022b; Du et al. 2022; Guo et al. 2023), such as following human instructions, in-context learning, and reasoning complex tasks. Wei et al. (2022c); Kojima et al. (2022) found that LLM can solve complex problems efficiently by chain-of-thought prompting strategy. Some works (Yuan et al. 2023; Shao et al. 2024; Yue et al. 2023; Yu et al. 2023) further improve the LLM's reasoning capabilities by fine-tuning. In this paper, we focus on using cross-lingual corpus to enhance the cross-lingual reasoning capabilities of LLM.

Cross-lingual Transfer

Cross-lingual transfer pertains to utilizing labeled data from a resource language to address the challenge of insufficient

labeled data in the target language. Previous works (Conneau and Lample 2019; Conneau et al. 2020; Yang et al. 2020; Ma et al. 2020; Yang et al. 2023) demonstrate that pre-trained models trained on multi-lingual data proficiently perform cross-lingual transfer tasks. These multi-lingual pre-trained models have found extensive application across various downstream NLP tasks, such as translation (Tan et al. 2019; Yang et al. 2022b; Gaschi et al. 2023; Yang et al. 2022c), summarization (Bhattacharjee et al. 2023; Wang et al. 2023), information extraction (Zhou et al. 2022; Yang et al. 2022a; Li et al. 2023; Wu et al. 2020), spell checking (Bexte et al. 2022; Sharma et al. 2023; Li et al. 2022, 2024). Many LLMs (Scao et al. 2022; Muennighoff et al. 2022) are trained on multilingual data, endowing them with strong cross-linguistic abilities.

Conclusion

We propose a cross-lingual instruction fine-tuning framework (xCoT) to address the disparity by encouraging alignment among different languages. A cross-lingual instruction dataset is created to align reasoning capability across various languages semantically. Our method incorporates cross-lingual in-context Learning to trigger the cross-lingual alignment. During instruction tuning, we adopt random online CoT, which prompts LLM to translate the query into different languages and subsequently provide an English response. To further promote language transfer, we leverage a high-resource CoT to guide low-resource CoT training with cross-lingual distillation. Comprehensive evaluations demonstrate xCoT effectively reduces the multilingual linguistic gap, highlighting its potential as a robust solution.

Acknowledgments

This work was supported in part by the National Natural Science Foundation of China (Grant Nos. 62276017, 62406033, U1636211, 61672081), and the State Key Laboratory of Complex & Critical Software Environment (Grant No. SKLCCSE-2024ZX-18).

References

- Bai, J.; Bai, S.; Chu, Y.; Cui, Z.; Dang, K.; Deng, X.; Fan, Y.; Ge, W.; Han, Y.; Huang, F.; Hui, B.; Ji, L.; Li, M.; Lin, J.; Lin, R.; Liu, D.; Liu, G.; Lu, C.; Lu, K.; Ma, J.; Men, R.; Ren, X.; Ren, X.; Tan, C.; Tan, S.; Tu, J.; Wang, P.; Wang, S.; Wang, W.; Wu, S.; Xu, B.; Xu, J.; Yang, A.; Yang, H.; Yang, J.; Yang, S.; Yao, Y.; Yu, B.; Yuan, H.; Yuan, Z.; Zhang, J.; Zhang, X.; Zhang, Y.; Zhang, Z.; Zhou, C.; Zhou, J.; Zhou, X.; and Zhu, T. 2023. Qwen Technical Report. *CoRR*.
- Bexte, M.; Laarmann-Quante, R.; Horbach, A.; and Zesch, T. 2022. LeSpell-A Multi-Lingual Benchmark Corpus of Spelling Errors to Develop Spellchecking Methods for Learner Language. In *Proc. of LREC*, 697–706.
- Bhattacharjee, A.; Hasan, T.; Ahmad, W. U.; Li, Y.; Kang, Y.; and Shahriyar, R. 2023. CrossSum: Beyond English-Centric Cross-Lingual Summarization for 1, 500+ Language Pairs. In *Proc. of ACL*, 2541–2564.
- Chen, N.; Zheng, Z.; Wu, N.; Shou, L.; Gong, M.; Song, Y.; Zhang, D.; and Li, J. 2023. Breaking Language Barriers in Multilingual Mathematical Reasoning: Insights and Observations. *CoRR*.
- Cobbe, K.; Kosaraju, V.; Bavarian, M.; Chen, M.; Jun, H.; Kaiser, L.; Plappert, M.; Tworek, J.; Hilton, J.; Nakano, R.; Hesse, C.; and Schulman, J. 2021. Training Verifiers to Solve Math Word Problems. *CoRR*.
- Conneau, A.; Khandelwal, K.; Goyal, N.; Chaudhary, V.; Wenzek, G.; Guzmán, F.; Grave, E.; Ott, M.; Zettlemoyer, L.; and Stoyanov, V. 2020. Unsupervised Cross-lingual Representation Learning at Scale. In *ACL 2020*, 8440–8451.
- Conneau, A.; and Lample, G. 2019. Cross-lingual Language Model Pretraining. In *NeurIPS 2019*, 7057–7067.
- Du, Z.; Qian, Y.; Liu, X.; Ding, M.; Qiu, J.; Yang, Z.; and Tang, J. 2022. GLM: General Language Model Pretraining with Autoregressive Blank Infilling. In *Proc. of ACL*, 320–335.
- Gaschi, F.; Fontaine, X.; Rastin, P.; and Toussaint, Y. 2023. Multilingual Clinical NER: Translation or Cross-lingual Transfer? In *Proceedings of the 5th Clinical Natural Language Processing Workshop, ClinicalNLP@ACL 2023, Toronto, Canada, July 14, 2023*, 289–311.
- Guo, H.; Yang, J.; Liu, J.; Yang, L.; Chai, L.; Bai, J.; Peng, J.; Hu, X.; Chen, C.; Zhang, D.; Shi, X.; Zheng, T.; Zheng, L.; Zhang, B.; Xu, K.; and Li, Z. 2023. OWL: A Large Language Model for IT Operations. *CoRR*.
- Hui, B.; Yang, J.; Cui, Z.; Yang, J.; Liu, D.; Zhang, L.; Liu, T.; Zhang, J.; Yu, B.; Lu, K.; et al. 2024. Qwen2. 5-coder technical report. *arXiv preprint arXiv:2409.12186*.
- Kojima, T.; Gu, S. S.; Reid, M.; Matsuo, Y.; and Iwasawa, Y. 2022. Large Language Models are Zero-Shot Reasoners. In *NeurIPS*.
- Li, Y.; Huang, S.; Zhang, X.; Zhou, Q.; Li, Y.; Liu, R.; Cao, Y.; Zheng, H.; and Shen, Y. 2023. Automatic Context Pattern Generation for Entity Set Expansion. *TKDE 2023*, 12458–12469.
- Li, Y.; Xu, Z.; Chen, S.; Huang, H.; Li, Y.; Ma, S.; Jiang, Y.; Li, Z.; Zhou, Q.; Zheng, H.-T.; and Shen, Y. 2024. Towards Real-World Writing Assistance: A Chinese Character Checking Benchmark with Faked and Misspelled Characters. In *Proc. of ACL*, 8656–8668.
- Li, Y.; Zhou, Q.; Li, Y.; Li, Z.; Liu, R.; Sun, R.; Wang, Z.; Li, C.; Cao, Y.; and Zheng, H.-T. 2022. The Past Mistake is the Future Wisdom: Error-driven Contrastive Probability Optimization for Chinese Spell Checking. In *Proc. of ACL Findings*, 3202–3213.
- Luo, H.; Sun, Q.; Xu, C.; Zhao, P.; Lou, J.; Tao, C.; Geng, X.; Lin, Q.; Chen, S.; and Zhang, D. 2023. Wizardmath: Empowering mathematical reasoning for large language models via reinforced evol-instruct. *arXiv preprint arXiv:2308.09583*.
- Ma, S.; Yang, J.; Huang, H.; Chi, Z.; Dong, L.; Zhang, D.; Awadalla, H. H.; Muzio, A.; Eriguchi, A.; Singhal, S.; Song, X.; Menezes, A.; and Wei, F. 2020. XLM-T: Scaling up Multilingual Machine Translation with Pretrained Cross-lingual Transformer Encoders. *CoRR*.
- Maaten, L. v. d.; and Hinton, G. 2008. Visualizing data using t-SNE. *JMLR*, 2579–2605.
- Muennighoff, N.; Wang, T.; Sutawika, L.; Roberts, A.; Biderman, S.; Scao, T. L.; Bari, M. S.; Shen, S.; Yong, Z.-X.; Schoelkopf, H.; et al. 2022. Crosslingual generalization through multitask finetuning. *arXiv preprint arXiv:2211.01786*.
- OpenAI. 2023. GPT-4 Technical Report. *CoRR*.
- Ouyang, L.; Wu, J.; Jiang, X.; Almeida, D.; Wainwright, C. L.; Mishkin, P.; Zhang, C.; Agarwal, S.; Slama, K.; Ray, A.; Schulman, J.; Hilton, J.; Kelton, F.; Miller, L.; Simens, M.; Askell, A.; Welinder, P.; Christiano, P. F.; Leike, J.; and Lowe, R. 2022. Training language models to follow instructions with human feedback. In *NeurIPS*.
- Patel, A.; Bhattamishra, S.; and Goyal, N. 2021. Are NLP Models really able to Solve Simple Math Word Problems? In *Proc. of NAACL*, 2080–2094.
- Patel, A.; Li, B.; Rasooli, M. S.; Constant, N.; Raffel, C.; and Callison-Burch, C. 2023. Bidirectional Language Models Are Also Few-shot Learners. In *Proc. of ICLR*.
- Qin, L.; Chen, Q.; Wei, F.; Huang, S.; and Che, W. 2023. Cross-lingual Prompting: Improving Zero-shot Chain-of-Thought Reasoning across Languages. In *Proc. of EMNLP*, 2695–2709.
- Scao, T. L.; Fan, A.; Akiki, C.; Pavlick, E.; Ilić, S.; Hesslow, D.; Castagné, R.; Luccioni, A. S.; Yvon, F.; Gallé, M.; et al. 2022. Bloom: A 176b-parameter open-access multilingual language model. *arXiv preprint arXiv:2211.05100*.
- Shao, H.; Qian, S.; Xiao, H.; Song, G.; Zong, Z.; Wang, L.; Liu, Y.; and Li, H. 2024. Visual cot: Advancing multi-modal language models with a comprehensive dataset and benchmark for chain-of-thought reasoning. In *Proc. of NeurIPS*.

- Sharma, S.; Valls-Vargas, J.; King, T. H.; Guerin, F.; and Arora, C. 2023. Contextual multilingual spellchecker for user queries. In *Proc. of SIGIR*, 3395–3399.
- Shi, F.; Suzgun, M.; Freitag, M.; Wang, X.; Srivats, S.; Vosoughi, S.; Chung, H. W.; Tay, Y.; Ruder, S.; Zhou, D.; Das, D.; and Wei, J. 2023. Language models are multilingual chain-of-thought reasoners. In *Proc. of ICLR*.
- Tan, X.; Chen, J.; He, D.; Xia, Y.; Qin, T.; and Liu, T. 2019. Multilingual Neural Machine Translation with Language Clustering. In *Proc. of EMNLP*, 963–973.
- Touvron, H.; Lavril, T.; Izacard, G.; Martinet, X.; Lachaux, M.; Lacroix, T.; Rozière, B.; Goyal, N.; Hambro, E.; Azhar, F.; Rodriguez, A.; Joulin, A.; Grave, E.; and Lample, G. 2023. LLaMA: Open and Efficient Foundation Language Models. *CoRR*.
- Touvron, H.; and Louis Martin, e. 2023. Llama 2: Open Foundation and Fine-Tuned Chat Models. *CoRR*.
- Wang, J.; Meng, F.; Liang, Y.; Zhang, T.; Xu, J.; Li, Z.; and Zhou, J. 2023. Understanding Translationese in Cross-Lingual Summarization. In *Proc. of EMNLP Findings*, 3837–3849.
- Wei, J.; Bosma, M.; Zhao, V. Y.; Guu, K.; Yu, A. W.; Lester, B.; Du, N.; Dai, A. M.; and Le, Q. V. 2022a. Finetuned Language Models are Zero-Shot Learners. In *Proc. of ICLR*.
- Wei, J.; Tay, Y.; Bommasani, R.; Raffel, C.; Zoph, B.; Borgeaud, S.; Yogatama, D.; Bosma, M.; Zhou, D.; Metzler, D.; Chi, E. H.; Hashimoto, T.; Vinyals, O.; Liang, P.; Dean, J.; and Fedus, W. 2022b. Emergent Abilities of Large Language Models. *TMLR*.
- Wei, J.; Wang, X.; Schuurmans, D.; Bosma, M.; Ichter, B.; Xia, F.; Chi, E. H.; Le, Q. V.; and Zhou, D. 2022c. Chain-of-Thought Prompting Elicits Reasoning in Large Language Models. In *NeurIPS*.
- Wu, Q.; Lin, Z.; Karlsson, B. F.; Huang, B.; and Lou, J. 2020. UniTrans : Unifying Model Transfer and Data Transfer for Cross-Lingual Named Entity Recognition with Unlabeled Data. In *Proc. of IJCAI*, 3926–3932.
- Yang, J.; Huang, S.; Ma, S.; Yin, Y.; Dong, L.; Zhang, D.; Guo, H.; Li, Z.; and Wei, F. 2022a. CROP: Zero-shot Cross-lingual Named Entity Recognition with Multilingual Labeled Sequence Translation. In *Findings of EMNLP 2022*, 486–496.
- Yang, J.; Ma, S.; Dong, L.; Huang, S.; Huang, H.; Yin, Y.; Zhang, D.; Yang, L.; Wei, F.; and Li, Z. 2023. GanLM: Encoder-Decoder Pre-training with an Auxiliary Discriminator. In *Proc. of ACL*, 9394–9412.
- Yang, J.; Ma, S.; Zhang, D.; Wu, S.; Li, Z.; and Zhou, M. 2020. Alternating Language Modeling for Cross-Lingual Pre-Training. In *Proc. of AAAI*, 9386–9393.
- Yang, J.; Yin, Y.; Ma, S.; Zhang, D.; Li, Z.; and Wei, F. 2022b. High-resource Language-specific Training for Multilingual Neural Machine Translation. In *Proc. of IJCAI*, 4461–4467.
- Yang, J.; Yin, Y.; Ma, S.; Zhang, D.; Wu, S.; Guo, H.; Li, Z.; and Wei, F. 2022c. UM4: Unified Multilingual Multiple Teacher-Student Model for Zero-Resource Neural Machine Translation. In *Proc. of IJCAI*, 4454–4460.
- Yu, L.; Jiang, W.; Shi, H.; Yu, J.; Liu, Z.; Zhang, Y.; Kwok, J. T.; Li, Z.; Weller, A.; and Liu, W. 2023. MetaMath: Bootstrap Your Own Mathematical Questions for Large Language Models. *CoRR*.
- Yuan, Z.; Yuan, H.; Li, C.; Dong, G.; Tan, C.; and Zhou, C. 2023. Scaling relationship on learning mathematical reasoning with large language models. *arXiv preprint arXiv:2308.01825*.
- Yue, X.; Qu, X.; Zhang, G.; Fu, Y.; Huang, W.; Sun, H.; Su, Y.; and Chen, W. 2023. Mammoth: Building math generalist models through hybrid instruction tuning. *arXiv preprint arXiv:2309.05653*.
- Zhang, Z.; Zhang, A.; Li, M.; Zhao, H.; Karypis, G.; and Smola, A. 2023. Multimodal Chain-of-Thought Reasoning in Language Models. *CoRR*.
- Zhou, R.; Li, X.; Bing, L.; Cambria, E.; Si, L.; and Miao, C. 2022. ConNER: Consistency Training for Cross-lingual Named Entity Recognition. In *Proc. of EMNLP*, 8438–8449.