

An Empirical Study of Multilingual Reasoning Distillation for Question Answering

Patomporn Payoungkhamdee¹, Peerat Limkonchotiwat^{2*},
Jinheon Baek³, Potsawee Manakul⁴, Can Udomcharoenchaikit¹,
Ekapol Chuangsuwanich⁵, Sarana Nutanong¹

¹School of Information Science and Technology, VISTEC ²AI Singapore ³KAIST
⁴SCB 10X ⁵Department of Computer Engineering, Chulalongkorn University
{patomporn.p_s21, canu_pro, snutanon}@vistec.ac.th, peerat@aisingapore.org
jinheon.baek@kaist.ac.kr, potsawee@scb10x.com, ekapolc@cp.eng.chula.ac.th

Abstract

Reasoning is one crucial capability in Large Language Models (LLMs), allowing them to perform complex tasks such as solving math problems and multi-step planning. While reasoning capability can emerge in larger models, smaller ones usually have to rely on distillation to transfer this capability from a larger model. However, recent efforts to distill reasoning capabilities have focused mainly on English, leaving multilingual distillation underexplored. To address this gap, this paper examines existing English reasoning distillation methods that utilize a variety of positive rationales in multilingual settings and proposes d-CoT-nR, which incorporates incorrect rationales as additional guidance. Empirical results from multilingual high-school examinations show that d-CoT-nR significantly surpasses the baseline, improving accuracy and the correctness of step-by-step reasoning.¹

1 Introduction

One potential capability of Large Language Models (LLMs) is reasoning, which allows them to analyze complex situations, draw inferences, and make predictions from the given input. This enhances their performance in tasks requiring understanding implicit relationships and generating coherent, informed responses (Wei et al., 2022b; Wang et al., 2023). For example, existing studies demonstrate that *Chain-of-Thought (CoT)* is beneficial for solving math problems (Cobbe et al., 2021) and problems demanding multi-step reasoning (Geva et al., 2021).

This reasoning capability remains challenging for smaller models (Wei et al., 2022a), making this capability out of reach to compute-constrained scenarios. Consequently, transferring reasoning

capabilities from larger to smaller models has gained attention recently.

A popular approach is to distill rationales from a larger teacher model (Magister et al., 2023; Hsieh et al., 2023; Ho et al., 2023), which has been shown to improve the CoT reasoning capabilities of smaller models in arithmetic, symbolic, and commonsense reasoning. Kang et al. (2023) further enhance the CoT reasoning process of smaller models by integrating relevant knowledge from external sources, compensating for their limited parameter knowledge. However, despite numerous studies on reasoning distillation, its applicability to multilingual setups remains underexplored.

To address this gap, this paper examines the effectiveness of existing reasoning distillation approaches for English in multilingual settings. We identify key factors and propose a novel set of distillation techniques to effectively distill smaller models for multilingual settings, as shown in Figure 1. Our research comprises three questions:

- (F1) *Native CoT distillation*: Is distillation using the same language as the question (native) more effective for smaller models than only English? While multilingual capabilities are now common, the choice of language for CoT distillation needs further research.
- (F2) *Diverse CoT (d-CoT)*: Is the diversity of rationales beneficial for CoT distillation in the multilingual setup? Diversifying rationales has been shown to enhance performance in monolingual settings (Ho et al., 2023). This paper verifies that diversification enhances accuracy in both seen and unseen languages.
- (F3) *Diverse CoT with negative rationales (d-CoT-nR)*: Can negative rationales further improve multilingual CoT distillation? Extending from our observation about the effectiveness of diverse CoT, we propose utilizing incorrect rationales (leading to incorrect outputs) to refine and delimit diversification.

^{*}Work was conducted while Peerat Limkonchotiwat was a PhD candidate at VISTEC

¹<https://github.com/calzone/lover/d-cot-nr>

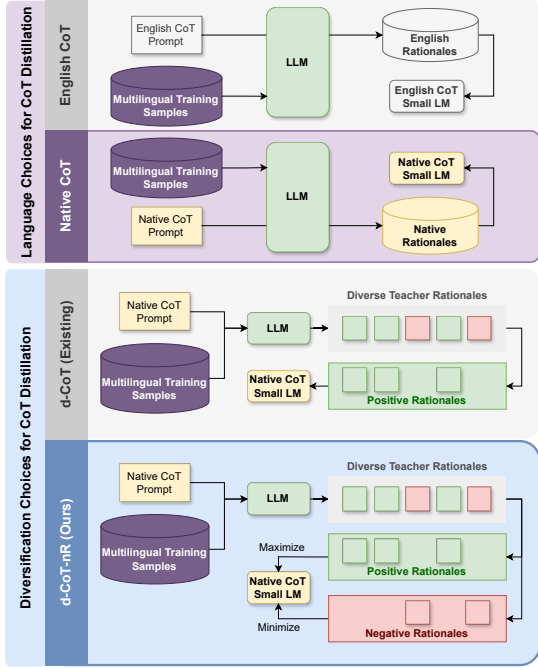


Figure 1: The key design decisions for multilingual CoT distillation are as follows: (Top) A comparison pipeline between English and native CoT distillation; (Bottom-greyed) d-CoT distillation incorporating diverse positive rationales; (Bottom-colored) Our d-CoT-nR distillation leveraging both positive and negative rationales.

To assess the multilingual reasoning distillation in knowledge-intensive tasks, we use multilingual high-school examinations (Hardalov et al., 2020) in multiple-choice questions. Surprisingly, for distillation of smaller models, native CoT consistently yields superior or comparable results to English CoT, contrary to the CoT prompting results in LLMs by Shi et al. (2022), which were done on larger models. Moreover, the results affirm the importance of the diversity of positive rationales in reasoning distillation. Finally, the inclusion of negative rationales significantly enhances model performance.

Our contributions are summarized as follows:

- To the best of our knowledge, this is the first work to provide an empirical study on strategies for multilingual reasoning distillation.
- We introduce d-CoT-nR, a strategy that utilizes negative rationales as additional guidance. The experimental results confirm that incorporating negative rationales along with positive ones during distillation is beneficial.
- We demonstrate that in small LMs distillation, native CoT consistently outperforms English

CoT, contrary to the existing trend in LLMs. However, in English-dominant pre-trained small models, the performance difference between native and English CoT becomes negligible.

2 Background

Rationale Generation. A dataset for multiple-choice question answering tasks can be denoted as $\mathcal{D} = \{(x_i, y_i)\}_{i=1}^N$, where x_i represents an input sequence comprising context, question, and answer choices (as illustrated in Figure 3), y_i is an answer in an alphabet letter, and N is the number of training samples. Following the work done in English (Magister et al., 2023; Hsieh et al., 2023), we can obtain the rationales from LLMs via CoT prompting (p_{CoT}) from the input as follows: $r_i = \text{LLM}(p_{\text{CoT}}, x_i)$. Then, we can use a collection of CoT training samples: $\mathcal{D}_{\text{CoT}} = \{(x_i, y_i, r_i)\}_{i=1}^N$, during distillation of smaller models below.

General CoT Distillation. Similar to Magister et al. (2023); Hsieh et al. (2023); Ho et al. (2023), based on a triplet of (x_i, y_i, r_i) in \mathcal{D}_{CoT} from the teacher LLM, we train smaller models (parameterized with θ) to reason through the language modeling objective (Radford et al., 2018):

$$\mathcal{L}_{\text{CoT}} = -\frac{1}{B} \sum_{i=1}^B \log p_{\theta}(y_i, r_i | x_i), \quad (1)$$

where B is the batch size. Furthermore, we use the same training objective for the knowledge-augmented (KA) fine-tuning (FT) baseline without teacher rationales.

3 Methodology

This study explores the three key factors (F1-F3, illustrated in Figure 1) to enhance multilingual CoT distillation, which we formalize in this section.

Native CoT distillation. According to Shi et al. (2022), in multilingual settings, one can prompt a teacher model to generate a CoT in English (*English CoT*, $p_{\text{CoT}}^{\text{en}}$) irrespective of the language of the input. Yet, the English CoT may not capture nuanced meanings inherent to the original language. Alternatively, the teacher model can be instructed to use the native language of the input for reasoning (*Native CoT*, $p_{\text{CoT}}^{\text{nt}}$).

Diverse CoT (d-CoT) Distillation. To obtain a set of diverse rationales from the teacher, we use its stochasticity in generation as follows: $\{r_{ij}\}_{j=1}^D =$

LLM($p_{\text{CoT}}^L, \mathbf{x}_i; D$) where L is the choice of CoT language and D is the teacher rational diversity (Ho et al., 2023). It is worth noting that, since some of the generated rationales explicitly lead to the incorrect answer, similar to Zelikman et al. (2022), we use a heuristic approach to filter out those *negative rationales* (r_{ij}^-) and only keep *positive rationales* (r_{ij}^+). Then, **d-CoT** training samples can be written as $\mathcal{D}_{\text{d-CoT}} = \{(\mathbf{x}_i, \mathbf{y}_i, r_{ij}^+)\}_{i=1}^M$ where M is the number of positive rationales for all samples. We denote the final training objective as:

$$\mathcal{L}_{\text{d-CoT}} = -\frac{1}{B \cdot D} \sum_{j=1}^D \sum_{i=1}^B \log p_{\theta}(\mathbf{y}_i, r_{ij}^+ | \mathbf{x}_i) \quad (2)$$

Diverse CoT with negative rationales (d-CoT-nR) Distillation. Previous reasoning distillation studies (Ho et al., 2023; Kang et al., 2023) have exclusively employed the diversity of only positive rationales to guide the reasoning paths toward the correct answer. We hypothesize that incorporating negative rationales could assist in further refining the decoding search space. Specifically, there are multiple ways to train the model with negative rationales, as follows: (i) Unlikelihood objective (Welleck et al., 2020) (ii) Contrastive learning via CLICK (Zheng et al., 2023) (iii) Preference optimization using ORPO (Hong et al., 2024). Detailed descriptions are provided in Appendices A and B. We represent their objective function as $\mathcal{L}_{\text{CoT-nR}}$, and train the smaller model with d-CoT as follows:

$$\mathcal{L}_{\text{d-CoT-nR}} = \alpha \mathcal{L}_{\text{d-CoT}} + (1 - \alpha) \mathcal{L}_{\text{CoT-nR}}, \quad (3)$$

where α is the hyperparameter that decides the trade-off between d-CoT and CoT-nR.

4 Experimental Setup

Datasets. To validate the capability of smaller (distilled) models in knowledge-intensive reasoning tasks, we use the EXAMS dataset (Hardalov et al., 2020). This dataset consists of multilingual high-school multiple-choice examination questions and their relevant passages across various subjects, e.g., chemistry, biology, and history. Based on the background knowledge corpus, languages are categorized² into top-portion languages (TPL) and bottom-portion languages

²Based on Table 8 in Hardalov et al. (2020) where there are languages that has more than one million articles in the corpus and the rest are below.

(BPL). We use accuracy as the main metric while using McNemar’s test for the significant test.

Language Models. We use GPT3.5-turbo-1106 as the teacher model. For the student model, we experiment with mT5-base (580M) (Xue et al., 2021), XGLM (564M) (Lin et al., 2022), and Gemma-Instruct (2B) (Team et al., 2024)³. mT5 and XGLM are pre-trained on a multilingual corpus that compensated for language skewness, while the pretraining data for Gemma was predominately in English. Note that we report the average across three seeds for mT5 and XGLM in d-CoT and d-CoT-nR methods, while we report only one seed for Gemma. Additionally, the training details are described in Appendix C.

5 Experimental Results

5.1 English vs Native CoT Distillation

Our findings reveal that native CoT distillation is the most effective strategy for small language models. As shown in Table 1, native CoT distillation consistently outperforms the English counterpart across TPL, BPL, and overall accuracy for mT5 and XGLM. Similarly, for a stronger model such as Gemma, despite achieving high scores without CoT, native CoT distillation remains effective, demonstrating that native CoT distillation is effective in all settings examined. In addition, we observe that for a strong baseline (Gemma), English CoT distillation yields a similar improvement as native CoT distillation.

It is known that CoT prompting in English is more effective than in native languages in large models (Shi et al., 2022). However, we do not observe the same trend for CoT distillation for smaller models. This suggests that empirical results on CoT prompting may not lead to the same results for CoT distillation.

5.2 Diverse CoT (d-CoT) Distillation

Table 1 shows that diversification significantly improves distillation performance, which is consistent across TPL and BPL, as well as the three architectures. On average, native d-CoT is the best performer and statistically better than English d-CoT. However, in the case of the Gemma model, no statistically significant difference was found, with English d-CoT slightly outperforming native d-CoT. These results indicate that smaller models

³<https://huggingface.co/google/gemma-2b-it>

Method	TPL				BPL							Average		
	it	pl	vi	pt	sr	hu	tr	bg	hr	mk	sq	TPL	BPL	All
Random	26.0	25.0	25.0	25.0	26.2	27.7	23.1	25.0	26.7	25.0	25.0	25.3	25.5	25.4
mT5 (580M)														
FT with KA	34.9	25.8	25.3	32.2	32.1	29.6	30.2	34.9	34.6	28.4	33.8	29.5	31.9	31.1
KA + English CoT	61.3	58.6	49.2	59.4	59.1	60.1	61.9	71.9	64.2	70.6	61.3	57.1	64.2	61.6
KA + Native CoT	66.5*	65.2*	52.5*	63.5	62.1*	63.8*	64.7*	74.7*	68.3*	75.0*	64.5*	61.9*	67.6*	65.5*
KA + English d-CoT	72.0[‡]	70.7[‡]	56.4[‡]	73.1[‡]	65.0[‡]	73.4[‡]	72.3[‡]	79.4[‡]	72.6[‡]	81.4[‡]	71.6[‡]	68.0[‡]	73.7[‡]	71.6[‡]
KA + Native d-CoT	71.6[‡]	70.6[‡]	57.6^{*‡}	74.1[‡]	66.9^{*‡}	73.9[‡]	72.9[‡]	79.9^{*‡}	73.5^{*‡}	80.7^{*‡}	72.5^{*‡}	68.4^{*‡}	74.3^{*‡}	72.2^{*‡}
XGLM (564M)														
FT with KA	31.1	27.7	26.7	21.6	30.1	27.8	26.3	29.8	29.6	28.8	25.2	26.8	28.2	27.7
KA + English CoT	46.9	45.8	36.0	38.7	46.0	42.7	46.2	54.6	49.8	51.3	45.2	41.8	48.0	45.7
KA + Native CoT	52.7*	52.7*	42.8*	45.6*	49.3	50.7*	45.6	60.9*	56.0*	57.4*	49.7*	48.4*	52.8*	51.2*
KA + English d-CoT	57.7[‡]	59.4[‡]	48.6[‡]	54.9[‡]	54.3[‡]	52.8[‡]	56.8[‡]	68.7[‡]	60.1[‡]	66.4[‡]	59.0[‡]	55.1[‡]	59.7[‡]	58.1[‡]
KA + Native d-CoT	59.2^{*‡}	60.5^{*‡}	49.9^{*‡}	58.6^{*‡}	57.1^{*‡}	58.3^{*‡}	58.6[‡]	70.5^{*‡}	62.7[‡]	69.0^{*‡}	62.6[‡]	57.1^{*‡}	62.7^{*‡}	60.7^{*‡}
Gemma-Instruct (2B)														
FT with KA	50.6	69.7	51.3	65.4	63.6	69.0	66.8	75.8	57.7	75.4	53.8	59.3	66.0	63.6
KA + English CoT	66.7	67.0	51.9	63.1	62.6	64.2	63.1	74.3	68.9	77.2	66.9	62.2	68.2	66.0
KA + Native CoT	67.0	66.5	51.3	65.0	63.9	66.7	65.3	74.7	70.3	77.9	68.6	62.5	69.6*	67.0*
KA + English d-CoT	69.2	72.0[‡]	57.1[‡]	70.2[‡]	68.6[‡]	69.5[‡]	69.9[‡]	77.4[‡]	71.5[‡]	78.8	72.8[‡]	67.1[‡]	72.7[‡]	70.6[‡]
KA + Native d-CoT	69.3	71.3[‡]	55.7[‡]	70.8[‡]	67.7[‡]	69.4[‡]	70.0[‡]	78.1[‡]	72.0	77.8	70.5	66.8[‡]	72.2[‡]	70.2[‡]
Teacher														
Zero-shot English CoT	70.3	66.8	54.2	64.6	68.5	59.9	69.5	73.0	72.6	68.2	13.9	64.0	60.8	62.0
Zero-shot Native CoT	69.3	61.0	50.7	67.4	64.9	57.8	51.0	68.9	71.1	64.5	19.1	62.1	56.8	58.7

Table 1: Accuracy of English and native CoT Distillations. We use "*" to indicate that native CoT is statistically significantly better than English CoT and vice versa and "‡" to indicate a statistically significant improvement when comparing d-CoT and CoT in the specified language(s).

benefit from further guidance in terms of rationale diversity and distillation in the target language.

5.3 Diverse CoT with Negative Rationales (d-CoT-nR) Distillation

In this study, we explore the potential of integrating d-CoT-nR with various training objectives to identify the most effective approach. Note that, given the computational costs involved, we first conduct experiments using smaller variants of encoder-decoder (mT5) and decoder-only (XGLM) architectures to identify the optimal training configuration. Our goal is to determine the most effective d-CoT-nR approach that enhances the performance of both architectures relative to d-CoT. Once the most effective d-CoT-nR configuration is established, we then apply this configuration to Gemma, the more capable base model in this study.

As shown in Table 2, the findings indicate that the unlikelihood approach significantly enhances the overall accuracy, improving performance not only across different languages but also across various architectures. Additionally, our results show that CLICK, a contrastive learning approach, is not beneficial in our setup. Furthermore, we found that although the ORPO methodology shows a notable performance enhancement in the XGLM model by increasing accuracy from 60.7% to 66.3% and surpassing the teacher model, it does not achieve comparable improvements in the mT5

model. Furthermore, empirical results reveal that utilizing the unlikelihood approach in the Gemma model for English reasoning leads to an average accuracy improvement of 0.7 percentage points. Therefore, we recommend the simplest approach, namely the unlikelihood method.

6 Additional Analysis

6.1 Unseen Languages

We extend our investigation to determine if small models trained on reasoning distillation exhibit cross-lingual capabilities for performing CoT reasoning in unseen languages. Table 3 indicates that, in most cases, CoT distillation enhances accuracy in unseen languages. Our results conform with previous studies (Hu et al., 2020) that when multilingual models are aligned in the pre-training step, learning in one language can be transferred across languages in the same model.

The significant performance improvement observed in d-CoT distillation implies that incorporating diversity can enhance generalization. Additionally, in the context of native reasoning for models such as mT5 and XGLM, our proposed d-CoT-nR method, using the unlikelihood objective, outperforms other approaches and demonstrates robustness in handling unseen languages. However, for the Gemma model, no improvement was observed when applying English d-CoT-nR to unseen languages; thus, future investigation could

Method	TPL				BPL				Average					
	it	pl	vi	pt	sr	hu	tr	bg	hr	mk	sq	TPL	BPL	All
Random	26.0	25.0	25.0	25.0	26.2	27.7	23.1	25.0	26.7	25.0	25.0	25.3	25.5	25.4
mT5 (580M)														
KA + Native d-CoT	71.6	70.6	57.6	74.1	66.9	73.9	72.9	79.9	73.5	80.7	72.5	68.4	74.3	72.2
KA + Native d-CoT-nR (Unlikelihood)	72.5 [◊]	71.3 [◊]	58.3	73.9	67.4	74.8 [◊]	73.3	80.6 [◊]	73.1	81.2 [◊]	72.6 [◊]	69.0 [◊]	74.7 [◊]	72.6 [◊]
KA + Native d-CoT-nR (Click)	71.9	70.2	57.7	73.9	66.0	73.7	72.5	79.8	72.5	80.8	72.4	68.4	74.0	71.9
KA + Native d-CoT-nR (ORPO)	71.8	70.6	57.5	73.1	66.0	74.1	72.9	79.3	73.3	80.5	72.3	68.2	74.1	71.9
XGLM (564M)														
KA + Native d-CoT	59.2	60.5	49.9	58.6	57.1	58.3	58.6	70.5	62.7	69.0	62.6	57.1	62.7	60.7
KA + Native d-CoT-nR (Unlikelihood)	65.3 [◊]	66.6 [◊]	53.1 [◊]	63.9 [◊]	62.2 [◊]	65.5 [◊]	63.5 [◊]	76.7 [◊]	68.2 [◊]	75.4 [◊]	68.7 [◊]	62.2 [◊]	68.6 [◊]	66.3 [◊]
KA + Native d-CoT-nR (Click)	44.5	45.7	37.6	44.6	44.4	43.6	42.8	53.1	48.2	50.7	44.7	43.1	46.8	45.5
KA + Native d-CoT-nR (ORPO)	66.2 [◊]	65.7 [◊]	53.2 [◊]	65.3 [◊]	62.6 [◊]	65.6 [◊]	63.2 [◊]	76.2 [◊]	68.1 [◊]	75.0 [◊]	68.4 [◊]	62.6 [◊]	68.4 [◊]	66.3 [◊]
Gemma-Instruct (2B)														
KA + English d-CoT	69.2	72.0	57.1	70.2	68.6	69.5	69.9	77.4	71.5	78.8	72.8	67.1	72.7	70.6
KA + English d-CoT-nR (Unlikelihood)	69.8	73.4	57.2	71.7	67.2	70.5	72.2 [◊]	77.4	71.6	78.7	74.8	68.0	73.2	71.3
Teacher														
Zero-shot English CoT	70.3	66.8	54.2	64.6	68.5	59.9	69.5	73.0	72.6	68.2	13.9	64.0	60.8	62.0
Zero-shot Native CoT	69.3	61.0	50.7	67.4	64.9	57.8	51.0	68.9	71.1	64.5	19.1	62.1	56.8	58.7

Table 2: Performance comparison between d-CoT and d-CoT-nR. "◊" denotes the statistically significant improvement when comparing d-CoT-nR and d-CoT in the specified language(s).

explore this setup across additional models.

Method	Languages					All
	de	fr	es	ar	lt	
mT5-base						
FT with KA	30.5	34.9	35.3	34.2	43.0	35.6
+ English CoT	58.2	62.3	52.8	72.2	82.1	65.5
+ Native CoT	58.1	67.3*	59.6*	76.3*	85.0*	69.3*
+ English d-CoT	72.2 [‡]	75.3 [‡]	68.1 [‡]	81.2 [‡]	89.5 [‡]	77.3 [‡]
+ Native d-CoT	72.4 [‡]	77.4 ^{‡*}	70.1 ^{‡*}	80.5 [‡]	89.2 [‡]	77.9 [‡]
+ Native d-CoT-nR	74.6 [◊]	78.3	72.9 [◊]	82.0 [◊]	88.7	79.3 [◊]
XGLM						
FT with KA	32.9	35.5	31.5	25.8	27.7	30.7
+ English CoT	46.4	49.4	41.3	51.2	59.4	49.5
+ Native CoT	44.5	48.7	47.7	50.5	64.4*	51.2
+ English d-CoT	54.7 [‡]	58.4 [‡]	45.0	68.3 [‡]	73.1 [‡]	59.9 [‡]
+ Native d-CoT	40.3	60.7 ^{‡*}	47.1*	68.4 [‡]	75.2 ^{‡*}	58.3 [‡]
+ Native d-CoT-nR	46.0	61.9 [◊]	44.3	76.2 [◊]	83.5 [◊]	62.4 [◊]
Gemma-Instruct (2B)						
FT with KA	65.2	48.7	53.6	80.4	72.3	64.1
+ English CoT	65.3	67.0	58.7	73.7	84.7	69.9
+ Native CoT	64.6	65.1	61.3	74.2	84.5	69.9
+ English d-CoT	72.6	74.8*	67.7 [‡]	81.1 [‡]	89.2 [‡]	77.1 [‡]
+ Native d-CoT	67.4	72.0 [‡]	63.0	78.1 [‡]	84.8	73.1 [‡]
+ English d-CoT-nR	69.5	69.2	67.2	79.4	88.5	74.8
Teacher						
Zero-shot English CoT	59.1	63.2	66.0	60.7	62.9	62.4
Zero-shot Native CoT	58.4	61.0	62.1	48.2	59.7	57.9

Table 3: Accuracy in unseen languages. Symbols have the same meanings as previously defined.

6.2 Reasoning Quality

While our empirical results demonstrate that CoT distillation improves accuracy in both seen and unseen languages, it remains unclear whether this performance gain is attributed to enhanced reasoning. To investigate this further, we analyzed step-by-step correctness following the methodology of Zeng et al. (2024) with the expert model (GPT-4o-2024-05-13); further details are provided in Appendix D.

According to Table 4, the analysis reveals that the d-CoT method increases overall reasoning

accuracy from 38.2% to 49.4%. The introduction of d-CoT-nR further improves accuracy to 50.1%. However, for unseen languages, the accuracy of d-CoT-nR is slightly lower than that of d-CoT by 0.8 percentage points.

It is noteworthy that the teacher model achieves an accuracy of 62.6%, indicating that d-CoT-nR not only improves accuracy but also enhances the quality of reasoning, demonstrating potential benefits for knowledge-intensive tasks.

Method	Rationale Correctness (%)			
	TPL	BPL	All	Unseen
KA + Native CoT	32.5	40.6	38.2	34.5
KA + Native d-CoT	42.0	52.4	49.4	40.0
KA + Native d-CoT-nR	45.1	52.1	50.1	39.2
Teacher	58.1	64.5	62.6	65.2

Table 4: Average correctness of the rationales producing by mT5 and the teacher model (GPT-3.5-turbo-1106).

7 Conclusion

This paper examines the effectiveness of reasoning distillation techniques in multilingual settings, focusing on small language models. Our empirical results show that native CoT distillation consistently outperforms the English counterpart across multiple metrics and languages, significantly improving models like mT5 and XGLM. Furthermore, diversifying rationales (d-CoT) improves performance across all languages and model architectures. Finally, we showcase that the proposed d-CoT-nR approach, incorporating both positive and negative rationales, is the best performer, significantly improving accuracy and robustness for unseen languages in most cases.

Limitations

The main evaluation method utilized in this work is multiple-choice exam questions. This assessment format provides a structured and controlled way to assess the capabilities of a large language model (LLM). However, this is a substantial departure from real-world knowledge-intensive tasks, such as document analysis, strategy development, content recommendation, and creative writing. While the proposed method narrows the performance gap in multiple-choice question answering, it may not correspond to enhanced user experiences in real-world tasks. Further research is needed to derive an assessment method that mimics the real-world conditions in tasks stated above without sacrificing the structured and controlled benefits we receive from multiple-choice exams.

References

- Karl Cobbe, Vineet Kosaraju, Mohammad Bavarian, Mark Chen, Heewoo Jun, Lukasz Kaiser, Matthias Plappert, Jerry Tworek, Jacob Hilton, Reiichiro Nakano, Christopher Hesse, and John Schulman. 2021. [Training verifiers to solve math word problems](#).
- Mor Geva, Daniel Khashabi, Elad Segal, Tushar Khot, Dan Roth, and Jonathan Berant. 2021. [Did aristotle use a laptop? a question answering benchmark with implicit reasoning strategies](#). *Transactions of the Association for Computational Linguistics*, 9:346–361.
- Momchil Hardalov, Todor Mihaylov, Dimitrina Zlatkova, Yoan Dinkov, Ivan Koychev, and Preslav Nakov. 2020. [Exams: A multi-subject high school examinations dataset for cross-lingual and multilingual question answering](#). In *Conference on Empirical Methods in Natural Language Processing*.
- Namgyu Ho, Laura Schmid, and Se-Young Yun. 2023. [Large language models are reasoning teachers](#). In *Proceedings of the 61st Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*, pages 14852–14882, Toronto, Canada. Association for Computational Linguistics.
- Jiwoo Hong, Noah Lee, and James Thorne. 2024. [Orpo: Monolithic preference optimization without reference model](#).
- Cheng-Yu Hsieh, Chun-Liang Li, Chih-kuan Yeh, Hootan Nakhost, Yasuhisa Fujii, Alex Ratner, Ranjay Krishna, Chen-Yu Lee, and Tomas Pfister. 2023. [Distilling step-by-step! outperforming larger language models with less training data and smaller model sizes](#). In *Findings of the Association for Computational Linguistics: ACL 2023*, pages 8003–8017, Toronto, Canada. Association for Computational Linguistics.
- Junjie Hu, Sebastian Ruder, Aditya Siddhant, Graham Neubig, Orhan Firat, and Melvin Johnson. 2020. [XTREME: A massively multilingual multi-task benchmark for evaluating cross-lingual generalisation](#). In *Proceedings of the 37th International Conference on Machine Learning*, volume 119 of *Proceedings of Machine Learning Research*, pages 4411–4421. PMLR.
- Minki Kang, Seanie Lee, Jinheon Baek, Kenji Kawaguchi, and Sung Ju Hwang. 2023. [Knowledge-augmented reasoning distillation for small language models in knowledge-intensive tasks](#). In *Thirty-seventh Conference on Neural Information Processing Systems*.
- Xi Victoria Lin, Todor Mihaylov, Mikel Artetxe, Tianlu Wang, Shuohui Chen, Daniel Simig, Myle Ott, Naman Goyal, Shruti Bhosale, Jingfei Du, Ramakanth Pasunuru, Sam Shleifer, Punit Singh Koura, Vishrav Chaudhary, Brian O’Horo, Jeff Wang, Luke Zettlemoyer, Zornitsa Kozareva, Mona Diab, Veselin Stoyanov, and Xian Li. 2022. [Few-shot learning with multilingual generative language models](#). In *Proceedings of the 2022 Conference on Empirical Methods in Natural Language Processing*, pages 9019–9052, Abu Dhabi, United Arab Emirates. Association for Computational Linguistics.
- Ilya Loshchilov and Frank Hutter. 2019. [Decoupled weight decay regularization](#). In *International Conference on Learning Representations*.
- Lucie Charlotte Magister, Jonathan Mallinson, Jakub Adamek, Eric Malmi, and Aliaksei Severyn. 2023. [Teaching small language models to reason](#). In *Proceedings of the 61st Annual Meeting of the Association for Computational Linguistics (Volume 2: Short Papers)*, pages 1773–1781, Toronto, Canada. Association for Computational Linguistics.
- Alec Radford, Karthik Narasimhan, Tim Salimans, and Ilya Sutskever. 2018. Improving language understanding by generative pre-training.
- Freda Shi, Mirac Suzgun, Markus Freitag, Xuezhi Wang, Suraj Srivats, Soroush Vosoughi, Hyung Won Chung, Yi Tay, Sebastian Ruder, Denny Zhou, Dipanjan Das, and Jason Wei. 2022. [Language models are multilingual chain-of-thought reasoners](#). *CoRR*, abs/2210.03057.
- Gemma Team, Thomas Mesnard, Cassidy Hardin, Robert Dadashi, Surya Bhupatiraju, Shreya Pathak, Laurent Sifre, Morgane Rivière, Mihir Sanjay Kale, Juliette Love, Pouya Tafti, Léonard Hussonot, Pier Giuseppe Sessa, Aakanksha Chowdhery, Adam Roberts, Aditya Barua, Alex Botev, Alex Castro-Ros, Ambrose Slone, Amélie Héliou, Andrea Tacchetti, Anna Bulanova, Antonia Paterson, Beth Tsai, Bobak Shahriari, Charline Le Lan, Christopher A. Choquette-Choo, Clément Crepy, Daniel Cer,

- Daphne Ippolito, David Reid, Elena Buchatskaya, Eric Ni, Eric Noland, Geng Yan, George Tucker, George-Christian Muraru, Grigory Rozhdestvenskiy, Henryk Michalewski, Ian Tenney, Ivan Grishchenko, Jacob Austin, James Keeling, Jane Labanowski, Jean-Baptiste Lespiau, Jeff Stanway, Jenny Brennan, Jeremy Chen, Johan Ferret, Justin Chiu, Justin Mao-Jones, Katherine Lee, Kathy Yu, Katie Millican, Lars Lowe Sjoesund, Lisa Lee, Lucas Dixon, Machel Reid, Maciej Mikula, Mateo Wirth, Michael Sharman, Nikolai Chinaev, Nithum Thain, Olivier Bachem, Oscar Chang, Oscar Wahltinez, Paige Bailey, Paul Michel, Petko Yotov, Rahma Chaabouni, Ramona Comanescu, Reena Jana, Rohan Anil, Ross McIlroy, Ruibo Liu, Ryan Mullins, Samuel L Smith, Sebastian Borgeaud, Sertan Girgin, Sholto Douglas, Shree Pandya, Siamak Shakeri, Soham De, Ted Klimentko, Tom Hennigan, Vlad Feinberg, Wojciech Stokowiec, Yu hui Chen, Zafarali Ahmed, Zhitao Gong, Tris Warkentin, Ludovic Peran, Minh Giang, Clément Farabet, Oriol Vinyals, Jeff Dean, Koray Kavukcuoglu, Demis Hassabis, Zoubin Ghahramani, Douglas Eck, Joelle Barral, Fernando Pereira, Eli Collins, Armand Joulin, Noah Fiedel, Evan Senter, Alek Andreev, and Kathleen Kenealy. 2024. [Gemma: Open models based on gemini research and technology](#).
- Xuezhi Wang, Jason Wei, Dale Schuurmans, Quoc V Le, Ed H. Chi, Sharan Narang, Aakanksha Chowdhery, and Denny Zhou. 2023. [Self-consistency improves chain of thought reasoning in language models](#). In *The Eleventh International Conference on Learning Representations*.
- Jason Wei, Yi Tay, Rishi Bommasani, Colin Raffel, Barret Zoph, Sebastian Borgeaud, Dani Yogatama, Maarten Bosma, Denny Zhou, Donald Metzler, Ed H. Chi, Tatsunori Hashimoto, Oriol Vinyals, Percy Liang, Jeff Dean, and William Fedus. 2022a. [Emergent abilities of large language models](#). *Transactions on Machine Learning Research*. Survey Certification.
- Jason Wei, Xuezhi Wang, Dale Schuurmans, Maarten Bosma, Ed H. Chi, Quoc Le, and Denny Zhou. 2022b. [Chain of thought prompting elicits reasoning in large language models](#). *CoRR*, abs/2201.11903.
- Sean Welleck, Ilia Kulikov, Stephen Roller, Emily Dinan, Kyunghyun Cho, and Jason Weston. 2020. [Neural text generation with unlikelihood training](#). In *International Conference on Learning Representations*.
- Linting Xue, Noah Constant, Adam Roberts, Mihir Kale, Rami Al-Rfou, Aditya Siddhant, Aditya Barua, and Colin Raffel. 2021. [mT5: A massively multilingual pre-trained text-to-text transformer](#). In *Proceedings of the 2021 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies*, pages 483–498, Online. Association for Computational Linguistics.
- Eric Zelikman, Yuhuai Wu, Jesse Mu, and Noah Goodman. 2022. [STar: Bootstrapping reasoning with reasoning](#). In *Advances in Neural Information Processing Systems*.
- Zhongshen Zeng, Pengguang Chen, Shu Liu, Haiyun Jiang, and Jiaya Jia. 2024. [Mr-gsm8k: A meta-reasoning benchmark for large language model evaluation](#).
- Chujie Zheng, Pei Ke, Zheng Zhang, and Minlie Huang. 2023. [Click: Controllable text generation with sequence likelihood contrastive learning](#). In *Findings of the Association for Computational Linguistics: ACL 2023*, pages 1022–1040, Toronto, Canada. Association for Computational Linguistics.

A Sample Generation for d-CoT-nR

To generate training samples for d-CoT-nR distillation, we extend each sample of d-CoT by randomly adding negative rationales from the same given input. For a training sample that has no negative rationales, we augment them by stochastically switching the final answer into an incorrect choice. Hence, we can represent the training data for d-CoT-nR as, $\mathcal{D}_{\text{d-CoT-nR}} = \{(\mathbf{x}_i, \mathbf{y}_i, \mathbf{r}_i^+, \mathbf{r}_i^-)\}_{i=1}^M$.

B Candidate Methods to utilize the negative rationales

Unlikelihood. Directly apply unlikelihood on the negative rationales. It is a simple yet most straightforward way to *degenerate* undesired sequence of tokens. Welleck et al. (2020) defined the unlikelihood loss as:

$$\mathcal{L}_{\text{Unlikelihood}} = -\frac{1}{B} \sum_{i=1}^B \log(1 - p_{\theta}(\mathbf{y}_i, \mathbf{r}_i^- | \mathbf{x}_i)) \quad (4)$$

CLICK. To encourage the likelihood rank of a positive rationale to surpass a negative rationale given the same input, a contrastive learning method known as CLICK (Zheng et al., 2023) can be used to marginalize over sequence likelihood. The objective $\mathcal{L}_{\text{CLICK}}$ is defined as follows:

$$-\frac{1}{B} \sum_{i=1}^B \max\{0, \gamma + \log p_{\theta}(\mathbf{y}_i, \mathbf{r}_i^- | \mathbf{x}_i) - \log p_{\theta}(\mathbf{y}_i, \mathbf{r}_i^+ | \mathbf{x}_i)\} \quad (5)$$

where γ is the margin hyperparameter.

ORPO. Given the favorable positive rationale and unwanted negative rationale, alternatively, the problem can be seen as preference optimization (PO). We apply ORPO (Hong et al., 2024) that adopt the concept of odds ratio disregarding the reward model. Formally, the ORPO loss in our setup can be written as

$$\mathcal{L}_{\text{ORPO}} = -\frac{1}{B} \sum_{i=1}^B \log \sigma \left(\log \frac{\text{odds}_{\theta}(\mathbf{y}_i, \mathbf{r}_i^+ | \mathbf{x}_i)}{\text{odds}_{\theta}(\mathbf{y}_i, \mathbf{r}_i^- | \mathbf{x}_i)} \right). \quad (6)$$

C Training Details

C.1 Fine-tuning and Inference

Similar to (Kang et al., 2023), we fine-tune all models using a learning rate of 5e-5, a batch size of 32, and a warm-up ratio of 0.1 for 3 epochs, utilizing the AdamW optimizer (Loshchilov and Hutter, 2019). However, for the XGLM in English and native CoT distillation, we extend the training duration to 20 epochs due to their

slower convergence relative to other configurations. Furthermore, we implement a classifier dropout rate of 0.1 across all strategies within the mT5 model.

For the inference phase, following Kang et al. (2023), we adopt the self-consistency inference method as described by Wang et al. (2023), setting the number of candidates to 10 for all experiments. The final answer was determined by the majority vote among these candidates.

C.2 Hyperparameters Tuning in d-CoT-nR Distillation

We perform a grid search to find the optimal parameters on development set of the dataset using hyper-parameters listed in the Table 5. The best performing hyperparameters on the development set is reported in Table 6. In addition, we further experiment the optimal hyperparameters with 3 different seeds for the study.

Methods	α	γ
Unlikelihood	0.9, 0.95, 0.99	-
CLICK	0.9, 0.8, 0.7	15, 20
ORPO	0.9, 0.8, 0.7	-

Table 5: Hyperparameter configurations.

Methods	Unlikelihood	CLICK	ORPO
mT5	$\alpha = 0.9$	$\alpha = 0.8, \gamma = 20$	$\alpha = 0.8$
XGLM	$\alpha = 0.99$	$\alpha = 0.9, \gamma = 15$	$\alpha = 0.9$

Table 6: Optimal hyperparameters.

C.3 Computing Resources

We trained the mT5 and XGLM models on an NVIDIA Tesla V100 GPU (32GB VRAM) over a period of 3 days, depending on the specific methods employed, with the inference phase for the test set requiring approximately 1-2 days. The Gemma model was fine-tuned on a single NVIDIA A100 GPU (80GB VRAM) for 3-4 days, with the test set inference phase extending to about 5 days. In total, it takes around 3,000 GPU hours on V100 and 840 GPU hours on A100. The associated training costs and performance improvements are detailed in Table 7.

D Reasoning Quality Assessment

To evaluate the accuracy of the reasoning steps, we selected identical sub-samples from the test set across various methods from a student model

Method	Estimated Training Time (Hours)	Accuracy (%)
mT5		
KA	2.77	25.4
CoT	6.55	65.5 (+40.1)
d-CoT	17.72	72.2 (+6.6)
d-CoT-nR (Wellick)	35.47	72.6 (+0.4)
d-CoT-nR (Click)	35.52	71.9 (-0.3)
d-CoT-nR (ORPO)	35.55	71.9 (-0.3)
XGLM		
KA	5.28	27.7
CoT	34.96	51.2 (+23.5)
d-CoT	34.13	60.7 (+9.5)
d-CoT-nR (Wellick)	70.18	66.3 (+5.6)
d-CoT-nR (Click)	70.47	45.5 (-15.2)
d-CoT-nR (ORPO)	70.73	66.3 (+5.6)

Table 7: Estimated training time for each approach of mT5 and XGLM.

(mT5). To minimize the study costs, we sampled 400 identical test sample IDs, achieving approximately a 95% confidence level with a 5% margin of error from a total of 13,510 test samples. Additionally, we ensured that the distribution of language-subjects in the sub-samples reflected the entire test set.

Regarding a test sample, a small model will generate a certain number of candidate rationales. We adopt the expert (GPT-4o-2024-05-13) prompting template from Zeng et al. (2024) to assess the step-by-step correctness of the student rationales and average the percentage of correct rationale across the identical samples.

E Demonstration

The demonstration of solving identical problem (Chemistry in Turkish language as shown in Figure 3) using step-by-step reasoning via various strategies on mT5 model, along with expert assessments, is illustrated in Figure 5, 6, 7 and, 8 for CoT, d-CoT, d-CoT-nR and teacher, respectively. Furthermore, we use ChatGPT for the translation provided in the figures.

F Generating Teacher Rationales

We use GPT-3.5-turbo-1106 as a teacher model. The system prompt utilized in our experiments is demonstrates in Figure 2 where [language] set to English for English CoT. For native CoT generation, we used the predefined language specified in the dataset metadata for each sample. The generation hyperparameters were consistent across all experiments, with $D = 10$, a temperature of 0.7, and a maximum token limit of 512.

System Prompt

The following are multiple-choice question and the relevant context in [subject] subject. Solve them in a step-by-step fashion and output a single option as the final answer in [language] language.

User Prompt

Context: [Context]
 Question: [Question]
 (A) [A text]
 (B) [B text]
 (C) [C text]
 (D) [D text]

Figure 2: CoT prompt template for the teacher model.

Context: IA grubu elementleri, Ca, Sr, Ba gibi aktif metallerin su ile reaksiyonu sonucunda hidrojen gazı elde edilir. Amonyak çözeltisinin sodyum hipoklorit ile etkileşmesinden elde edilir. $Cu(I)Cl$ çözeltisi genellikle $CuCl$ çözeltisinin çok fazla bakır metali yardımıyla indirgenmesi sonucunda elde edilir. Rüzgârdan elde edilecek elektrikle suyun hidroliz edilmesi sonucunda; su, oksijen ve hidrojen elementlerine ayrılarak çok ucuz bir yolla hidrojen elde edilmiş olacaktır. Doğal agrega, kum ocaklarından, dere yataklarından ya da deniz kıyısından elde edilir. Hidrojen gazının yakılması sonucunda ortama sadece su/su buharı çıkar. Elementlerinden de elde edilen hidrojen sülfür laboratuvarlarda demir sülfür üzerine hidrojen klorür etki ettirmekle elde edilir. Sulu çözeltisinin hidrojen peroksit ile oksidasyonu, tellurat iyonunu verir. Küçük molekül, (genellikle su, amonyak veya hidrojen klorid) elimine edilir. Ufalanma, insan eliyle veya makineler yardımıyla yapılacak olursa kırmızı taş veya sadece kırma denilen agrega elde edilir. Hidroklorik asit, hidrojen klorür adlı maddenin, suda çözülmesiyle elde edilir. NaBr, sodyum hidroksit in hidrojen bromür ile tepkimesi ile elde edilir. Makineleri dolaşan soğutucu su tatlı su olup eksilmesi halinde takviye edilir. Örneğin metandan bir hidrojen çıkarılırsa metil (CH), etandan bir hidrojen çıkarılırsa etil (CH) elde edilir. Genellikle sikloheksan ile bir karışım halinde elde edilir. Hidrokarbon yakıtın gas molekülleri anot yüzeyinde su buharı ile birlikte adsorblanır ve hidrojen atomları, etkin bir şekilde sıyrılarak elektrolite absorblanır. Potasyum elektroliz yöntemiyle elde edilen ilk metaldir. Elde edilen kireç, agrega ile karıştırılarak harç olarak kullanılmaktaydı. Alkali metal polisüfitler, bir sülfür çözeltisinin işlenmesi ile elde edilir, örneğin sodyum sülfat, elemental kükürt: Bazı durumlarda, bu anyonlar organik çözücülerde eriyebilen organik tuzlar olarak elde edilir. Demir dışı metaller genellikle elektroliz yoluyla rafine edilir. Bu NaBr'nin sulu çözeltisinin kolin gazı ile reaksiyonu ile elde edilir: Sodyum bromür endüstride oldukça kullanışlı bir malzemedir. Propilenin, su ve sülfat asidiyle hidrojenlendirilmesinden elde edilir. Su cazibeyle değil çoğunlukla pompaj ile elde edilir. Oksijen genel olarak beş farklı işlemle elde edilir: Hava ayrıştırma (ASU), basınç salımlı adsorpsiyon (PSA), vakum basınç salımlı adsorpsiyon (VPSA), elektroliz ile su ayrıştırma ve membran teknolojisi. Bu mekanik dönme hareketi sonucunda alternatörlerde elektrik elde edilir.

Question: Kalium hidroksid su çözeltisinin elektroliz süreci sonucunda katoda elektrodasında ne elde edilir:
(A) Kalium iyonları. (B) gas agrega halinde hidrojen . (C) Gas agrega halinde oksijen. (D) elementar helde kalium.

Figure 3: An example input in their native language.

Context: Hydrogen gas is produced through the reaction of IA group elements, such as Ca, Sr, and Ba, with water. It is also obtained from the interaction of an ammonia solution with sodium hypochlorite. Cu(I)Cl solution is typically obtained by reducing CuCl solution with an excess of copper metal. Through the electrolysis of water using electricity generated from wind, water is split into oxygen and hydrogen elements, yielding hydrogen in a very cost-effective manner. Natural aggregate is obtained from sand quarries, riverbeds, or seashores. The combustion of hydrogen gas releases only water/water vapor into the environment. Hydrogen sulfide, produced from its elements, is obtained in laboratories by reacting hydrogen chloride with iron sulfide. The oxidation of its aqueous solution with hydrogen peroxide yields the tellurate ion. A small molecule (usually water, ammonia, or hydrogen chloride) is eliminated. Crushing, whether done manually or with machinery, produces aggregate known as red stone or simply crushed stone. Hydrochloric acid is obtained by dissolving hydrogen chloride in water. NaBr is produced by reacting sodium hydroxide with hydrogen bromide. Fresh water, used as a coolant in machines, is replenished if depleted. For example, removing one hydrogen from methane yields methyl (CH₃), and removing one hydrogen from ethane yields ethyl (C₂H₅). It is usually obtained in a mixture with cyclohexane. Hydrocarbon fuel gas molecules are adsorbed on the anode surface along with water vapor, and hydrogen atoms are effectively stripped and absorbed into the electrolyte. Potassium was the first metal obtained by electrolysis. The lime obtained is mixed with aggregate to be used as mortar. Alkali metal polysulfides are obtained by processing a sulfide solution, such as sodium sulfate with elemental sulfur. In some cases, these anions are obtained as organic salts soluble in organic solvents. Non-ferrous metals are usually refined by electrolysis. This NaBr is obtained by reacting its aqueous solution with chlorine gas; sodium bromide is a highly useful industrial material. It is obtained from the hydrogenation of propylene with water and sulfuric acid. Water is typically obtained not by gravity but by pumping. Oxygen is generally obtained through five different processes: air separation (ASU), pressure swing adsorption (PSA), vacuum pressure swing adsorption (VPSA), water electrolysis, and membrane technology. This mechanical rotational motion generates electricity in alternators.

Question: What is obtained at the cathode electrode as a result of the electrolysis process of the potassium hydroxide aqueous solution?

(A) Potassium ions. (B) Hydrogen in gaseous form. (C) Oxygen in gaseous form. (D) Elemental potassium.

Figure 4: The example input *translated* into English language.

KA + Native CoT

Verilen metinde, kaliyum hidroksid su çözeltilisinin elektroliz süreci sonucunda katoda elektrodasında ne elde edilir: (C) Gas agrega halinde oksijen. => C

(Translated)

In the given text, what is obtained at the cathode electrode as a result of the electrolysis process of the potassium hydroxide aqueous solution: (C) Oxygen in gaseous form. => C

Expert Assessment

Step-by-step Evaluation:

1. The student was asked about the products at the cathode during the electrolysis of potassium hydroxide (KOH) in aqueous solution.
2. In the process of electrolysis of KOH, water is electrolyzed rather than KOH itself.
3. At the cathode (negative electrode), reduction occurs, and typically, hydrogen ions (H⁺) are reduced to hydrogen gas (H₂), not oxygen.
4. The student selected option C - Gas agrega halinde oksijen, which is incorrect terminology for cathode products. Gas agrega halinde oksijen (oxygen gas) is actually produced at the anode (positive electrode).

Final Judgement: wrong

Figure 5: The output produced by the model trained using native CoT from the given identical test example (Figure 3).

KA + Native d-CoT

1. Kaliyum hidroksid su çözeltilisinin elektroliz süreci sonucunda katoda elektrodasında hidrojen elde edilir. 2. Doğru cevap: (B) gas agrega halinde hidrojen. => B

(Translated)

1. Hydrogen is obtained at the cathode electrode as a result of the electrolysis process of the potassium hydroxide aqueous solution. 2. Correct answer: (B) Hydrogen in gaseous form. => B

Expert Assessment

Step-by-step Evaluation:

1. The student correctly identified that Kaliyum hidroksid (potassium hydroxide) in water undergoes electrolysis.
2. They correctly stated that during this process, hydrogen is produced at the cathode.
3. The answer choice selected by the student is (B) gas agrega halinde hidrojen, which translates to gaseous hydrogen, aligning with their correct reasoning.

Final Judgement: correct

Figure 6: The output produced by the model trained using native d-CoT from the given identical test example (Figure 3).

KA + Native d-CoT-nR

1. Kalium hidroksid, su çözeltisinin elektroliz süreci sonucunda katoda elektrodasında elde edilir. 2. Kalium hidroksid su çözeltisinin elektroliz süreci sonucunda katoda elektrodasında hidrojen elde edilir. 3. Dolayısıyla, doğru cevap (B) gas agrega halinde hidrojen'dir. => B

(Translated)

1. Potassium hydroxide is obtained at the cathode electrode as a result of the electrolysis process of the aqueous solution. 2. Hydrogen is obtained at the cathode electrode as a result of the electrolysis process of the potassium hydroxide aqueous solution. 3. Therefore, the correct answer is (B) Hydrogen in gaseous form. => B

Expert Assessment**Step-by-step Evaluation:**

1. The student starts by reiterating the question, which is unnecessary but does not contain errors, except it should be clearer that this is a re-statement of the question.
2. The student correctly states that during the electrolytic process of potassium hydroxide solution, hydrogen is produced at the cathode.
3. Based on the correctly stated reasoning in step 2, the student arrives at the correct solution, which is (B) gas aggregate hydrogen.

Final Judgement: correct

Figure 7: The output produced by the model trained using native d-CoT-nR from the given identical test example (Figure 3).

Teacher

Elektroliz süreci sonucunda katoda elektrodasında hidrojen gazı elde edilir.
Dolayısıyla, doğru cevap: (B) gas agrega halinde hidrojen.

(Translated)

Hydrogen gas is obtained at the cathode electrode as a result of the electrolysis process.
Therefore, the correct answer is: (B) Hydrogen in gaseous form.

Expert Assessment

Step-by-step Evaluation: The student correctly identifies that during the electrolytic process involving potassium hydroxide (KOH) solution, hydrogen gas is produced at the cathode. The student's choice of (B) gas aggregate hydrogen aligns with the correct understanding that hydrogen gas will indeed form at the cathode due to the reduction of water.

Final Judgement: correct

Figure 8: Output generated by the teacher (GPT3.5-turbo-1106) from the given identical test example (Figure 3).