Chain-of-Quizzes: Pedagogy-inspired Example Selection in In-Context-Learning

Yiquan Wu[1,2,4*], **Anlai Zhou**[1*], **Yuhang Liu**[3], **Yifei Liu**[3]
Adam Jatowt[4], **Weiming Lu**[1†], **Jun Xiao**[1], **Kun Kuang**[1,2†]

[1]College of Computer Science and Technology, Zhejiang University, Hangzhou, China, [2]AI&Law Lab, Zhejiang University,
[3]School of Software Technology, Zhejiang University, [4]University of Innsbruck, Innsbruck, Austria
{wuyiquan, zhouanlai, liuyuhang, liuyifei, luwm, kunkuang}@zju.edu.cn,
adam.jatowt@uibk.ac.at, junx@cs.zju.edu.cn

Abstract

In-context learning (ICL) has emerged as a powerful tool for enhancing large language models (LLMs) in addressing downstream tasks. In this paper, we explore the vital task of example selection in ICL by mimicking the human learning process. We propose a Chain-of-Quizzes (CoQ) framework inspired by educational theories such as Bruner's Spiral Learning and Mastery Learning theory. Specifically, our framework employs the LLMs to answer the quiz (question in the example) to sift 'good' examples, combines these examples iteratively with the increasing complexity, and utilizes a final exam to gauge the combined example chains. Our extensive experiments on diverse reasoning datasets show the proposed approach outperforms baseline models. These findings underscore the framework's potential for future research. The code and data will be made available here[1].

1 Introduction

With the scaling up the model size and corpus size (Devlin et al., 2018; Radford et al., 2019; Brown et al., 2020; Chowdhery et al., 2023), large language models (LLMs) have brought innovation to many fields (Adeshola and Adepoju, 2023; Cheng et al., 2023; Wu et al., 2023b). One of the most important abilities of LLMs, in-context learning (ICL), is a paradigm that allows language models to learn downstream tasks given only a few demonstrative examples (Dong et al., 2022).

In this paper, we delve into the essential task of example selection within ICL. Since ICL assumes that "LLMs can deduce many things from a few cases", like chain-of-thought that mimics the reasoning process of humans (Chu et al., 2023), we propose to rethink the selection of examples in analogy to the human learning process as shown in Figure 1. Specifically, we investigate two key questions: **1) What is a 'good' example for a learner (LLM)?** and **2) How to efficiently combine these good examples?**

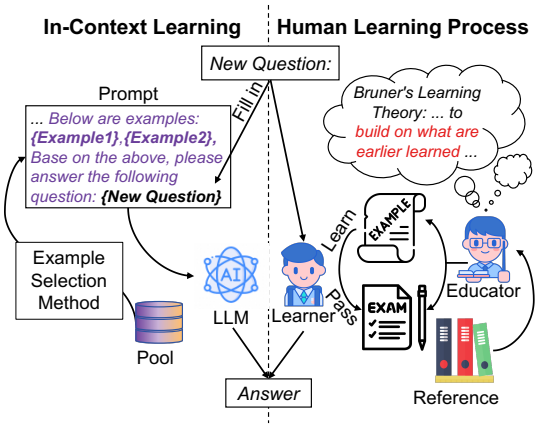

Figure 1: An illustration of the two learning processes, our motivation is to propose an example selection method for in-context learning (left part) by simulating the human learning process (right part).

Regarding the first question, in the human learning process, after acquiring several good examples, learners' ability of question-answering can be improved. The example here is considered not as knowledge itself but rather as a quiz to help the learner become familiar with the question type and activate relevant knowledge, implying that it must be understood by the learner easily. Therefore, we consider an example as 'good' for an LLM if it can successfully respond to the question[2] posed in the example, and we will only use good examples in the prompts.

For the second question, we propose a Chain-of-Quizzes (CoQ) framework inspired by two educational theories: Bruner's Spiral Learning theory and Mastery Learning theory[3]. Bruner's Spiral

[*] Equal contribution.
[†] Corresponding authors.
[1]https://github.com/anlaiJoe/Chain-of-Quizzes

[2]For simplicity, we focus on QA tasks in the current work.
[3]Here we use simple words like 'example' and 'question' to represent complex pedagogical concepts.

Findings of the Association for Computational Linguistics ACL 2024, pages 10136–10142

Learning theory proposes that learning should be built upon what learners have previously learned, and should gradually increase in complexity of examples (Takaya, 2008). The Mastery Learning theory suggests that the learners should master the current examples before answering the new question (Block et al., 1971). Specifically, the CoQ framework engages the LLM in answering the quiz (question in the example) to sift 'good' examples. It incrementally increases the complexity of the quiz based on the model's performance when combining the examples into a chain, and employs a final exam to evaluate the effectiveness of each example chain. Our framework is further refined using majority voting to account for the inherent randomness in example selection.

Extensive experiments conducted across various datasets demonstrate that our approach yields superior results compared to baseline models. These initial findings highlight the framework's significant potential for future research.

2 Related Work

2.1 In-Context Learning

With the development of deep learning, especially the LLMs, significant progress has been made in many tasks (Zhang et al., 2022a, 2023a,b; Wu et al., 2020, 2022, 2024, 2023a; Zhang et al., 2023c, 2024b,a). In-context learning (ICL) is a method where LLMs leverage provided context (e.g., examples) to perform tasks without explicit retraining (Brown et al., 2020; Min et al., 2021; Chen et al., 2022; Wei et al., 2023; Wu et al., 2023b). Recent work in ICL has predominantly concentrated on the design of prompt templates, such as Chain-of-Thought (CoT) and self-consistency (Wei et al., 2022; Zhang et al., 2023d; Wang et al., 2022b). Some studies have explored example selection, which is another important direction of ICL, the existing methods generally rely on retrieval mechanisms for each test instance to select relevant examples (Qin et al., 2023; Mavromatis et al., 2023; Li and Qiu, 2023). Zhang et al. (2022c) uses a grouping approach to ensure the diversity of examples, and Fu et al. (2022) leverages examples with more reasoning steps. (Ye et al., 2023) employ a Conditional Determinantal Point Process (DDP) for joint probability modeling of the demonstration examples. (Zhang et al., 2022b) proposes an identifying generalizable policies-based demonstration selecting strategy. (Rubin et al., 2022) utilizes LLM to assess the quality of demonstrations. (Li and Qiu, 2023) uses a filter-then-search method to tackle the enumerating challenge. (Chang and Jia, 2023) adopts a scoring approach to address the issue of demonstration selection. In this work, mimicking the human learning process, we propose a framework that selects examples iteratively to boost the performance of ICL.

2.2 Curriculum learning

Curriculum learning (CL) is one of the important data selection strategies, which is inspired by the way humans learn, and involves gradually increasing the complexity of training data, allowing models to build upon simpler concepts before tackling more complex ones (Wang et al., 2021). The key of CL is to define the complexity of training data and plan the training order (Soviany et al., 2022). This strategy not only facilitates more efficient learning but also potentially improves model generalization. CL is similar to the example combination part of our framework, but it is used for model training while we focus on the example selection for ICL.

3 Method

In this section, we introduce details of the framework, whose illustration is shown in Figure 2.

3.1 Quiz Phase

In the first phase, our goal is to select multiple example chains from the initial data pool, which contains a large number of question-answer pairs, as well as to construct a challenging data pool to be used in the next phase. Each example chain comprises k sequential good examples, with each example being a question-answer pair. We add one example per round until we complete k rounds, where k is a hyperparameter.

In the first round, we randomly select a number of questions from the data pool. We then prompt the LLM to take a quiz by answering the selected questions one by one with the following prompt:

P1: *[Please attempt to answer the question step by step:* ***<question>****]*

After reviewing the answers, questions answered correctly are paired with their answers to form good examples for the next round. Incorrectly answered questions are added to the challenging pool.

In subsequent rounds, we sample new questions from the remaining data and conduct a new quiz, leveraging the ICL from previously correctly answered examples. Prompt for these rounds is:

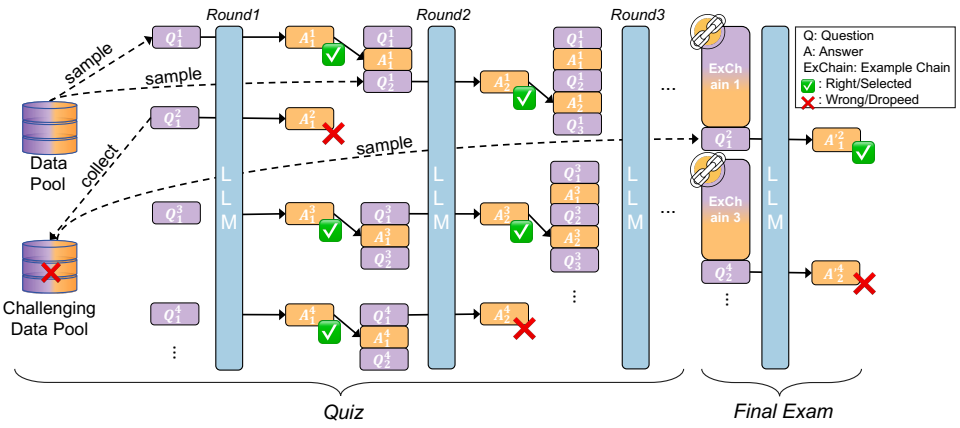

Figure 2: An illustration of the Chain-of-Quizzes framework, which consists of two phases: Quiz and Final Exam.

P2: *[Based on the following examples* ***<examples>****, please answer the question step by step:* ***<question>****]*

Drawing on Bruner's Spiral Learning theory, the complexity of new questions should be increased. Following the previous work (Lewis and Frank, 2016), we approximate the question's complexity by its length and ensure that each new question is longer than those in the examples of the previous rounds. If answered correctly, this indicates that the question-answer pair is a good example within the context of the provided examples. Incorrectly answered questions are again added to the challenging pool.

After k rounds, we obtain several example chains, each containing k sequentially arranged examples.

3.2 Final Exam Phase

Here, the objective is to find optimal example chains, which can help the LLMs answer the questions that they failed before. Recall that based on the quiz phase, we have curated multiple candidate example chains as well as accumulated a challenging data pool containing questions previously unanswered by the LLM. Guided by the principles of Mastery Learning theory, we now conduct a final exam to assess the LLM's proficiency over the example chains. The final exam process is straightforward: we randomly draw questions from the challenging data pool and require the LLM to answer them using the candidate example chain in the prompt. The prompt used is the same as **P2** in Section 3.1. A correct answer indicates that the LLM successfully "masters" the example chain; otherwise, the candidate chain is dropped.

After completing both phases, we have now several validated example chains. Note that one can repeat the phases several times until enough validated example chains are obtained.

3.3 Inference

During the inference, when presented with a new test question, we can thus obtain different answers using different example chains. A majority vote is then used to determine the final answer. This voting strategy also helps to mitigate the randomness inherent in our framework, ensuring more reliable and consistent results.

4 Experiments

4.1 Datasets

We conducted experiments across three distinct categories of QA datasets to assess the capability of our method. Specifically, for **Mathematics**, we utilized GSM8K (Cobbe et al., 2021), AddSub (Hosseini et al., 2014), AQuA (Ling et al., 2017), SingleEq (Koncel-Kedziorski et al., 2015), and SVAMP (Patel et al., 2021). For **Commonsense Reasoning**, we employed CSQA (Saha et al., 2018) and StrategyQA (Geva et al., 2021), and for **Symbolic Reasoning**, we utilized the Last Letter and Coin Flip datasets (Wang et al., 2022a). These diverse datasets enabled us to thoroughly validate the effectiveness of our method across various reasoning and computational tasks.

4.2 Experimental Settings

Our experiments were conducted using GPT-3.5-Turbo-0613 as the underlying LLM.

4.2.1 Baselines

Zero-shot approach directly employs the LLM to answer questions without presenting any examples. Chain-of-thought (**CoT**) (Kojima et al., 2022) uses

Methods	*Mathematics*					*Commonsense*		*Symbolic*		Avg
	GSM8K	AddSub	AQuA	SingleEq	SVAMP	CSQA	Strategy	Letter	Coin	
Zero-shot	69.4	89.6	53.9	93.8	82.0	59.7	75.5	22.0	55.0	66.8
Zero-shot-CoT	76.8	86.0	55.5	91.3	81.0	63.5	69.9	65.6	91.2	75.6
Zero-shot-PS	72.9	86.3	53.5	90.9	78.0	66.3	**77.2**	45.2	59.8	70.0
Few-shot	23.1	83.7	25.5	88.3	66.0	45.8	70.9	27.0	81.0	56.8
ToT	75.9	85.0	65.0	88.0	80.0	73.5	71.4	61.4	75.8	75.1
Auto-CoT	78.3	90.6	57.4	94.4	80.0	65.8	74.3	79.0	**95.0**	79.4
Self-Consistency	85.5	**91.3**	64.5	94.6	84.0	77.1	70.3	83.8	58.2	78.8
CoQ	**88.6**	**91.3**	**70.8**	**94.9**	**92.0**	**77.7**	70.6	**91.5**	84.9	**84.7**

Table 1: Results of experiments on different datasets.

Methods	GSM8K	AQuA	SVAMP
Zero-shot-PS	66.7	54.7	75.0
Auto-CoT	72.4	48.8	79.0
Self-Consistency	77.8	**69.7**	91.0
CoQ	**82.1**	67.8	**92.0**

Table 2: Results with Gemini as underlying LLM.

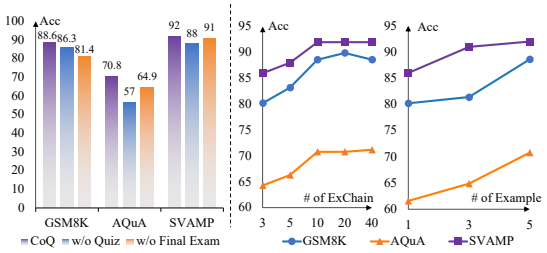

Figure 3: Results of ablation study.

the "Let's think step by step" prompt to encourage the LLM's reasoning ability. Plan-and-Solve Prompting (**PS**) (Wang et al., 2023) addresses the problem by first creating a plan to break down the task into smaller subtasks, and then executing these subtasks according to the plan. **Few-shot** approach provides the LLM with a fixed set of examples before it attempts to answer the questions; the examples are randomly chosen question-answer pairs. **Tree of Thought (ToT)** (Yao et al., 2023) is a framework allow LLM to explore multiple reasoning paths and make decisions through a tree-like structure of intermediate thoughts. **Auto-CoT** (Zhang et al., 2022c) aims to automatically construct diverse and effective examples by sampling questions and generating reasoning chains. **Self-Consistency** (Wang et al., 2022b) employs a strategy of first generating 40 reasoning paths and then identifying the most reliable answer by a majority voting.

In our approach, each example chain contains 5 examples, and we have 10 example chains for majority voting. We conduct the following ablation experiments on three main datasets: 1) **w/o Quiz**: replace the quiz by randomly selecting the same number of examples. 2) **w/o Final Exam**: remove the final exam phase. 3) Vary the number of examples in each set and the number of example chains. 4) Replace the underlying LLM with Gemini-Pro.

4.3 Experiment Results

We report the accuracy of various methods, with all results averaged over three runs.

From Table 1, we observe the following: 1) Compared to Zero-shot, the performance of the Few-shot drops significantly when directly adding examples to the prompt (e.g., accuracy on GSM8K decreases from 69.4% to 23.1%). This may be because the examples in the Few-shot are question-answer pairs, which forces the LLM to generate the answer directly without any reasoning. 2) Our method surpasses baselines on most datasets, particularly in Mathematics. This proves the effectiveness of our example selection framework. Compared to Self-Consistency, we use fewer reasoning paths and achieve better performance. 3) Our method did not yield optimal results on the StrategyQA and Coin Flip datasets. Upon analyzing these datasets, we find that there exists a unified problem-solving approach. For instance, in StrategyQA, the specific problem-solving approach is encompassed by the PS method, leading to its exceptional performance on this dataset.

From Table 2, it is also evident that our framework maintains its superiority even when the underlying Large Language Model (LLM) is changed (Gemini), demonstrating its generality.

From the ablation study presented in Figure 3, we can draw several conclusions: 1) The removal of the quiz phase results in a significant decline in performance, as evidenced by the decrease in AQuA's accuracy from 70.8% to 57.0%. This suggests the efficacy of our iterative example selection methodology. 2) Eliminating the final exam phase also causes a minor reduction in performance, suggesting that the final exam contributes positively to

the refinement of example chains. 3) An increase in the number of example chains correlates with improved accuracy. However, to achieve an optimal balance between speed and accuracy, selecting 10 example chains is found to be the best choice. 4) The larger the number of examples in an example chain, the better the performance. This proves the effective interaction among the examples.

5 Conclusion

In this work, inspired by pedagogy theories, we propose a novel framework Chain-of-Quizzes (CoQ) to select examples for the in-context learning of LLM. Extensive results demonstrate the effectiveness of our approach. In the future, we will: 1) experiment with other question complexity measurements and 2) expand the CoQ to more tasks.

6 Limitations

In this section, we discuss the limitations of our work:

• The current implementation of our framework can be regarded as preliminary. It however presents an opportunity for further exploration, particularly in diversifying the settings. For example, the definition of "complex" can be extended in Section 3.1.

• Our approach utilizes examples in the form of question-answer pairs. While this format has its merits and is quite commonly used nowadays, there exists potential for more effective usage. Incorporating reasoning steps within these examples could provide a deeper context for the LLMs, potentially leading to improved performance.

Acknowledgments

This work was supported by the National Natural Science Foundation of China (62441605, 62376245, 62376243, 62037001, U20A20387), the Starry Night Science Fund of Zhejiang University Shanghai Institute for Advanced Study (SN-ZJU-SIAS-0010), the Key Research and Development Program of Zhejiang Province, China (No. 2024C03255), the Fundamental Research Funds for the Central Universities(No. 226-2022-00051).

Finally, we would like to thank the anonymous reviewers for their helpful feedback and suggestions.

References

Ibrahim Adeshola and Adeola Praise Adepoju. 2023. The opportunities and challenges of chatgpt in education. *Interactive Learning Environments*, pages 1–14.

James H Block, Peter W Airasian, John Bissell Carroll, and Benjamin Samuel Bloom. 1971. *Mastery learning: Theory and practice*. Holt, Rinehart and Winston,.

Tom Brown, Benjamin Mann, Nick Ryder, Melanie Subbiah, Jared D Kaplan, Prafulla Dhariwal, Arvind Neelakantan, Pranav Shyam, Girish Sastry, Amanda Askell, et al. 2020. Language models are few-shot learners. *Advances in neural information processing systems*, 33:1877–1901.

Ting-Yun Chang and Robin Jia. 2023. Data curation alone can stabilize in-context learning.

Mingda Chen, Jingfei Du, Ramakanth Pasunuru, Todor Mihaylov, Srini Iyer, Veselin Stoyanov, and Zornitsa Kozareva. 2022. Improving in-context few-shot learning via self-supervised training. *arXiv preprint arXiv:2205.01703*.

Szu-Wei Cheng, Chung-Wen Chang, Wan-Jung Chang, Hao-Wei Wang, Chih-Sung Liang, Taishiro Kishimoto, Jane Pei-Chen Chang, John S Kuo, and Kuan-Pin Su. 2023. The now and future of chatgpt and gpt in psychiatry. *Psychiatry and clinical neurosciences*, 77(11):592–596.

Aakanksha Chowdhery, Sharan Narang, Jacob Devlin, Maarten Bosma, Gaurav Mishra, Adam Roberts, Paul Barham, Hyung Won Chung, Charles Sutton, Sebastian Gehrmann, et al. 2023. Palm: Scaling language modeling with pathways. *Journal of Machine Learning Research*, 24(240):1–113.

Zheng Chu, Jingchang Chen, Qianglong Chen, Weijiang Yu, Tao He, Haotian Wang, Weihua Peng, Ming Liu, Bing Qin, and Ting Liu. 2023. A survey of chain of thought reasoning: Advances, frontiers and future. *arXiv preprint arXiv:2309.15402*.

Karl Cobbe, Vineet Kosaraju, Mohammad Bavarian, Mark Chen, Heewoo Jun, Lukasz Kaiser, Matthias Plappert, Jerry Tworek, Jacob Hilton, Reiichiro Nakano, et al. 2021. Training verifiers to solve math word problems. *arXiv preprint arXiv:2110.14168*.

Jacob Devlin, Ming-Wei Chang, Kenton Lee, and Kristina Toutanova. 2018. Bert: Pre-training of deep bidirectional transformers for language understanding. *arXiv preprint arXiv:1810.04805*.

Qingxiu Dong, Lei Li, Damai Dai, Ce Zheng, Zhiyong Wu, Baobao Chang, Xu Sun, Jingjing Xu, and Zhifang Sui. 2022. A survey for in-context learning. *arXiv preprint arXiv:2301.00234*.

Yao Fu, Hao Peng, Ashish Sabharwal, Peter Clark, and Tushar Khot. 2022. Complexity-based prompting for multi-step reasoning. *arXiv preprint arXiv:2210.00720*.

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.

Mohammad Javad Hosseini, Hannaneh Hajishirzi, Oren Etzioni, and Nate Kushman. 2014. Learning to solve arithmetic word problems with verb categorization. In *Proceedings of the 2014 Conference on Empirical Methods in Natural Language Processing (EMNLP)*, pages 523–533, Doha, Qatar. Association for Computational Linguistics.

Takeshi Kojima, Shixiang Shane Gu, Machel Reid, Yutaka Matsuo, and Yusuke Iwasawa. 2022. Large language models are zero-shot reasoners. *Advances in neural information processing systems*, 35:22199–22213.

Rik Koncel-Kedziorski, Hannaneh Hajishirzi, Ashish Sabharwal, Oren Etzioni, and Siena Dumas Ang. 2015. Parsing Algebraic Word Problems into Equations. *Transactions of the Association for Computational Linguistics*, 3:585–597.

Molly L Lewis and Michael C Frank. 2016. The length of words reflects their conceptual complexity. *Cognition*, 153:182–195.

Xiaonan Li and Xipeng Qiu. 2023. Finding support examples for in-context learning.

Wang Ling, Dani Yogatama, Chris Dyer, and Phil Blunsom. 2017. Program induction by rationale generation: Learning to solve and explain algebraic word problems. In *Proceedings of the 55th Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*, pages 158–167, Vancouver, Canada. Association for Computational Linguistics.

Costas Mavromatis, Balasubramaniam Srinivasan, Zhengyuan Shen, Jiani Zhang, Huzefa Rangwala, Christos Faloutsos, and George Karypis. 2023. Which examples to annotate for in-context learning? towards effective and efficient selection. *arXiv preprint arXiv:2310.20046*.

Sewon Min, Mike Lewis, Luke Zettlemoyer, and Hannaneh Hajishirzi. 2021. Metaicl: Learning to learn in context. *arXiv preprint arXiv:2110.15943*.

Arkil Patel, Satwik Bhattamishra, and Navin Goyal. 2021. Are NLP models really able to solve simple math word problems? In *Proceedings of the 2021 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies*, pages 2080–2094, Online. Association for Computational Linguistics.

Chengwei Qin, Aston Zhang, Anirudh Dagar, and Wenming Ye. 2023. In-context learning with iterative demonstration selection. *arXiv preprint arXiv:2310.09881*.

Alec Radford, Jeffrey Wu, Rewon Child, David Luan, Dario Amodei, Ilya Sutskever, et al. 2019. Language models are unsupervised multitask learners. *OpenAI blog*, 1(8):9.

Ohad Rubin, Jonathan Herzig, and Jonathan Berant. 2022. Learning to retrieve prompts for in-context learning.

Amrita Saha, Vardaan Pahuja, Mitesh Khapra, Karthik Sankaranarayanan, and Sarath Chandar. 2018. Complex sequential question answering: Towards learning to converse over linked question answer pairs with a knowledge graph. In *Proceedings of the AAAI conference on artificial intelligence*, volume 32.

Petru Soviany, Radu Tudor Ionescu, Paolo Rota, and Nicu Sebe. 2022. Curriculum learning: A survey. *International Journal of Computer Vision*, 130(6):1526–1565.

Keiichi Takaya. 2008. Jerome bruner's theory of education: From early bruner to later bruner. *Interchange*, 39:1–19.

Lei Wang, Wanyu Xu, Yihuai Lan, Zhiqiang Hu, Yunshi Lan, Roy Ka-Wei Lee, and Ee-Peng Lim. 2023. Plan-and-solve prompting: Improving zero-shot chain-of-thought reasoning by large language models. *arXiv preprint arXiv:2305.04091*.

Peifeng Wang, Aaron Chan, Filip Ilievski, Muhao Chen, and Xiang Ren. 2022a. Pinto: Faithful language reasoning using prompt-generated rationales. *arXiv preprint arXiv:2211.01562*.

Xin Wang, Yudong Chen, and Wenwu Zhu. 2021. A survey on curriculum learning. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 44(9):4555–4576.

Xuezhi Wang, Jason Wei, Dale Schuurmans, Quoc Le, Ed Chi, Sharan Narang, Aakanksha Chowdhery, and Denny Zhou. 2022b. Self-consistency improves chain of thought reasoning in language models. *arXiv preprint arXiv:2203.11171*.

Jason Wei, Xuezhi Wang, Dale Schuurmans, Maarten Bosma, Fei Xia, Ed Chi, Quoc V Le, Denny Zhou, et al. 2022. Chain-of-thought prompting elicits reasoning in large language models. *Advances in Neural Information Processing Systems*, 35:24824–24837.

Jerry Wei, Le Hou, Andrew Lampinen, Xiangning Chen, Da Huang, Yi Tay, Xinyun Chen, Yifeng Lu, Denny Zhou, Tengyu Ma, et al. 2023. Symbol tuning improves in-context learning in language models. *arXiv preprint arXiv:2305.08298*.

Yiquan Wu, Kun Kuang, Yating Zhang, Xiaozhong Liu, Changlong Sun, Jun Xiao, Yueting Zhuang, Luo Si, and Fei Wu. 2020. De-biased court's view generation with causality. In *Proceedings of the 2020 Conference on Empirical Methods in Natural Language Processing (EMNLP)*, pages 763–780.

Yiquan Wu, Yifei Liu, Weiming Lu, Yating Zhang, Jun Feng, Changlong Sun, Fei Wu, and Kun Kuang. 2022. Towards interactivity and interpretability: A rationale-based legal judgment prediction framework. In *Proceedings of the 2022 Conference on Empirical Methods in Natural Language Processing*, pages 4787–4799.

Yiquan Wu, Yifei Liu, Ziyu Zhao, Weiming Lu, Yating Zhang, Changlong Sun, Fei Wu, and Kun Kuang. 2024. De-biased attention supervision for text classification with causality. In *Proceedings of the AAAI Conference on Artificial Intelligence*, volume 38, pages 19279–19287.

Yiquan Wu, Weiming Lu, Yating Zhang, Adam Jatowt, Jun Feng, Changlong Sun, Fei Wu, and Kun Kuang. 2023a. Focus-aware response generation in inquiry conversation. In *Findings of the Association for Computational Linguistics: ACL 2023*, pages 12585–12599.

Yiquan Wu, Siying Zhou, Yifei Liu, Weiming Lu, Xiaozhong Liu, Yating Zhang, Changlong Sun, Fei Wu, and Kun Kuang. 2023b. Precedent-enhanced legal judgment prediction with llm and domain-model collaboration. *arXiv preprint arXiv:2310.09241*.

Shunyu Yao, Dian Yu, Jeffrey Zhao, Izhak Shafran, Thomas L. Griffiths, Yuan Cao, and Karthik Narasimhan. 2023. Tree of thoughts: Deliberate problem solving with large language models.

Jiacheng Ye, Zhiyong Wu, Jiangtao Feng, Tao Yu, and Lingpeng Kong. 2023. Compositional exemplars for in-context learning.

Rongzhi Zhang, Jiaming Shen, Tianqi Liu, Jialu Liu, Michael Bendersky, Marc Najork, and Chao Zhang. 2023a. Do not blindly imitate the teacher: Using perturbed loss for knowledge distillation. *arXiv preprint arXiv:2305.05010*.

Rongzhi Zhang, Jiaming Shen, Tianqi Liu, Haorui Wang, Zhen Qin, Feng Han, Jialu Liu, Simon Baumgartner, Michael Bendersky, and Chao Zhang. 2024a. Plad: Preference-based large language model distillation with pseudo-preference pairs.

Rongzhi Zhang, Yue Yu, Jiaming Shen, Xiquan Cui, and Chao Zhang. 2023b. Local boosting for weakly-supervised learning. In *Proceedings of the 29th ACM SIGKDD Conference on Knowledge Discovery and Data Mining*, pages 3364–3375.

Rongzhi Zhang, Yue Yu, Pranav Shetty, Le Song, and Chao Zhang. 2022a. Prboost: Prompt-based rule discovery and boosting for interactive weakly-supervised learning. *arXiv preprint arXiv:2203.09735*.

Wenqi Zhang, Yongliang Shen, Weiming Lu, and Yueting Zhuang. 2023c. Data-copilot: Bridging billions of data and humans with autonomous workflow. *arXiv preprint arXiv:2306.07209*.

Wenqi Zhang, Ke Tang, Hai Wu, Mengna Wang, Yongliang Shen, Guiyang Hou, Zeqi Tan, Peng Li, Yueting Zhuang, and Weiming Lu. 2024b. Agent-pro: Learning to evolve via policy-level reflection and optimization. *arXiv preprint arXiv:2402.17574*.

Yiming Zhang, Shi Feng, and Chenhao Tan. 2022b. Active example selection for in-context learning.

Zhuosheng Zhang, Aston Zhang, Mu Li, and Alex Smola. 2022c. Automatic chain of thought prompting in large language models. *arXiv preprint arXiv:2210.03493*.

Zhuosheng Zhang, Aston Zhang, Mu Li, Hai Zhao, George Karypis, and Alex Smola. 2023d. Multimodal chain-of-thought reasoning in language models. *arXiv preprint arXiv:2302.00923*.