Knowledge-enhanced Prompt Learning for Open-domain Commonsense Reasoning

ZHAO Xujiang, LIU Yanchi, CHENG Wei, OISHI Mika, OSAKI Takao, MATSUDA Katsushi, CHEN Haifeng

Abstract

Neural language models for commonsense reasoning often formulate the problem as a QA task and make predictions based on learned representations of language after fine-tuning. However, without providing any fine-tuning data and pre-defined answer candidates, can neural language models still answer commonsense reasoning questions only relying on external knowledge? In this work, we investigate a unique yet challenging problem-open-domain commonsense reasoning that aims to answer questions without providing any answer candidates and fine-tuning examples. A team comprising NECLA (NEC Laboratories America) and NEC Digital Business Platform Unit proposed method leverages neural language models to iteratively retrieve reasoning chains on the external knowledge base, which does not require task-specific supervision. The reasoning chains can help to identify the most precise answer to the commonsense question and its corresponding knowledge statements to justify the answer choice. This technology has proven its effectiveness in a diverse array of business domains.

Keywords
Commonsense Reasoning, External Knowledge, pre-trained language model, open-domain, reasoning chains

1. Introduction

In the realm of business, large-scale pretrained language models (PLMs)[1] have emerged as a dominant force in natural language processing (NLP). These PLMs acquire a foundational understanding of the world through extensive training on vast collections of general text data, followed by fine-tuning on specific datasets for various applications. While PLMs have demonstrated exceptional performance in numerous downstream tasks, they still grapple with two critical challenges in the context of reasoning-related endeavors:

(1) Incomplete Knowledge: PLMs often struggle when faced with tasks that require information not present in their training data or when dealing with test instances that do not conform to a question-and-answer format.

(2) Limited Reasoning Capabilities: PLMs make predictions based on implicitly encoded knowledge, which lacks the structured reasoning abilities necessary for complex tasks and fails to provide clear explanations for their chosen responses.

In the business world, addressing these challenges is crucial for leveraging PLMs effectively in applications such as customer support, data analysis, and content generation. Finding solutions to enhance their reasoning capabilities and handle diverse information sources is essential to unlock their full potential in various business contexts. As shown in **Fig. 1**, if we are presented with a question that its domain is different from examples seen during the training. For medical-domain questions like What are both Family Doctor and Surgeon refer to? we aim to generate an abstracted meaning for both entities without providing any answer candidates. However, without providing any fine-tuning instances, one of the state-of-the-art PLMs T5-3b[2] would generate an irreverent answer: *quizlet*. In addition, for commonsense questions: People aim to [MASK] at work, the paradigm of prompt learning with PLMs often formulates the problem to multiple-choice QA and calculate the likelihood of the whole sentence by filling in the blank with each answer candidate. However, both answers learning from others and completing the job are semantically correct. PLMs cannot provide justification for why a certain answer can be chosen. Both cases reveal that the prediction of commonsense reasoning requires robust and structured

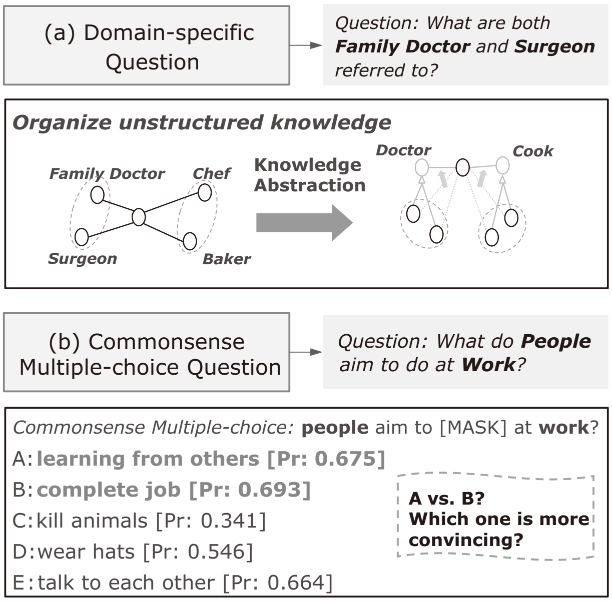


Fig. 1 Two cruxes of using PLMs in commonsense reasoning.

reasoning to integrate the explicit information offered by the question context and external knowledge.

In this work, we focus on the Open-domain Commonsense Reasoning task, which requires machines to make human-like presumptions about the type and essence of ordinary situations without presenting any answer candidates and fine-tuning examples. In this work, we present the external KnowlEdge Enhanced Prompting method (KEEP) to achieve open-ended commonsense reasoning without pre-defining an answer candidate set and an answer scope. Firstly, to eliminate the requirement of answer candidates, KEEP leverages an external knowledge base (e.g., ConceptNet) as the answer searching space and iteratively extracts multi-hop reasoning paths relevant to the question. To avoid searching exhaustively over the whole knowledge base, we leverage PLMs to formulate the overall search criteria. The key insight is PLMs have certain reasoning abilities through their large-scale model parameters, which can be utilized to provide implicit knowledge in determining whether or not to keep expanding the reasoning paths or adopt the entity in the path as the final answer. Therefore, without restricting specific answer scopes and direct supervision of the reasoning process, KEEP can be applied in most real-world scenarios requiring commonsense reasoning. To further enhance the reasoning ability of the PLM, we propose to leverage task-agnostic reasoning paths extracted directly from the external knowledge base as training instances to finetune the PLM.

2. Knowledge Enhanced Prompting Method

In section 2, we first introduce the problem formulation, and then discuss the detailed framework of the proposed method, which can be divided into three components:

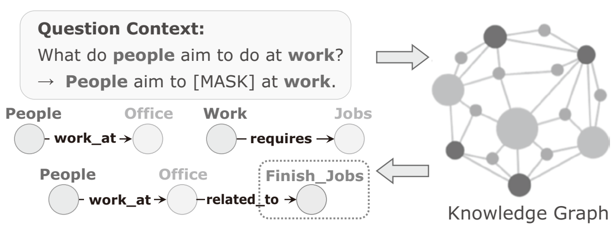
(1) Entity extraction and linking
(2) Local knowledge graph expansion
(3) Training strategy and answer prediction

2.1 Problem Formulation

We aim to solve open-ended commonsense reasoning questions by jointly using knowledge from a PLM and a structured knowledge graph. The knowledge graph (KG)G=(V, E) (e.g., ConceptNet) is a multi-relational heterogeneous graph. V is the set of entity nodes, $E \subseteq V \times R \times V$ is the set of edges that connect nodes in V, where R represents a set of relation types (e.g., locates_at or requires). Specifically, given an open-ended commonsense reasoning question q without providing answer candidates and regulating an answer scope, the target of this work is to determine 1) a local KG $Gq \in G$ contains relevant information of q; 2) a set of reasoning paths k = $\{k_1, k_2, \ldots, k_m\}$ extracted from Gq ; and 3) an entity $\hat{a}$ extracted from k that is precise to answer the question q. For example, in **Fig. 2**, to answer a commonsense question "what do people aim to do at work?", we aim at first extracting all relevant reasoning paths from the external KG that can provide us with logical information to answer the question. Among all the paths, we select the most precise one (i.e., people → office → finish_jobs) and extract the answer $\hat{a}$ = finish_jobs such that the following joint likelihood can be maximized.

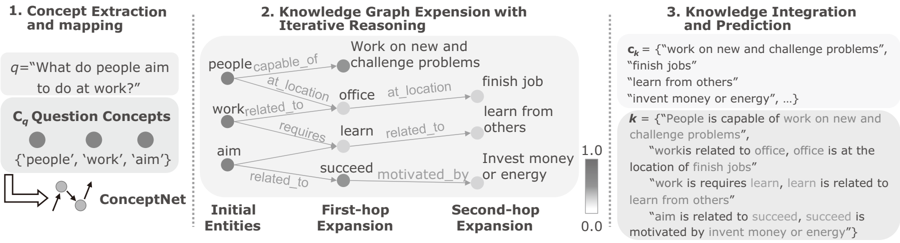
$$P\left(\hat{a},\mathbf{k}\middle|q,G_q\right) = P\left(\mathbf{k}\middle|q,G_q\right)\cdot P\left(\hat{a}\middle|\mathbf{k}\right)$$

Challenges: However, maximizing the joint likelihood is not a trivial task due to two critical obstacles. First,



The model takes the question as input and returns supporting reasoning paths (i.e., knowledge statements) with the best answer.

Fig. 2 Example of the open-ended commonsense reasoning.



1) concept extraction and entity linking; 2) local knowledge graph expansion with iterative reasoning steps, and 3) knowledge integration and final answer prediction.

Fig. 3 The framework of the proposed method consists.

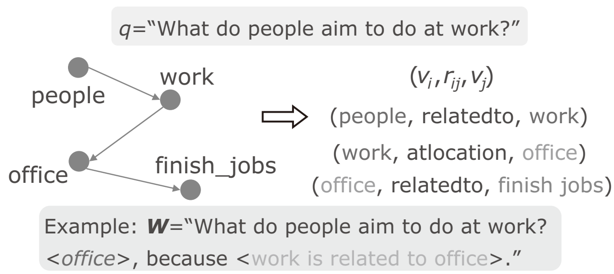
retrieving the question-relevant reasoning paths k (i.e., knowledge statements) is difficult since we cannot build a local KG between question entities and answer candidates under the open-ended setting as existing works[3)-5)] do. Moreover, without regulating a pre-defined answer scope as differentiable method[6)] does, the search space would be the whole knowledge graph. Next, to solve both challenges, we discuss how to initiate the local KG and iteratively reason over it to find all plausible knowledge statements and the most convincing answer. We demonstrate the overall framework in **Fig 3**.

2.2 Local Graph Construction and Expansion Knowledge Graph Entity Linking.

Conceptual knowledge graphs (e.g., ConceptNet) enable a variety of useful context-oriented reasoning tasks over real-world texts, which provides us with the most suitable structured knowledge in open-ended commonsense reasoning. To reason over a given commonsense context using knowledge from both PLM and G, the first step of the framework is to extract the set of critical entities $c_q = \{c_q^{(1)}, \cdots, c_q^{(i)}, \cdots\}$ from the question q that have the surjective mapping to a node set $V_q \in V$ in the KG. And we follow the prior work[7)] to map informative entities c_q from q to conjunct concept entities V_q in KG by leveraging the latent representation of the query context and relational information stored in G.

Reasoning Over Local Knowledge Graph: To imitate the human reasoning process, we aim to retrieve reasoning paths within L hops from G to form the local knowledge subgraph G_q that has the highest coverage to the question concepts c_q. Ideally, each path in G_q can be regarded as a reasoning chain that helps to locate the most precise answer and its explanation to the question q. However, expanding L-hop subgraph G_q from c_q is computationally prohibited.

Reasoning Path Pruning: In order to make the process of reasoning path expansion scalable, we incorporate the implicit knowledge in PLMs to prune irreverent paths. Specifically, we pair the question q with the text of node v along with the reasoning-path-transformed knowledge statement to form a cloze-based prompt $W = [q; v_j; (v_i, r_{ij}, v_j)]$ in order to turn the local graph expansion problem into an explicit reasoning procedure by directly answering the question with its derived reasoning path. For example, in **Fig. 4**, the prompt is formatted as What do people aim to do at work? <answer_node>, because <reasoning path>. We leverage a pre-defined template to transform the triplet (v_i, r_{ij}, v_j) into natural language. For instance, the triplet (work, antonym, unemployment) can be translated to work is the antonym of unemployment as illustrated in Fig. 4. To evaluate whether we keep the reasoning path, we propose leveraging the PLM to score the relevance of each reasoning path given the context of the question. Formally, suppose the prompt W consists of N tokens $W = \{\omega_1, \ldots, \omega_{n-1}, \omega_n, \omega_{n+1}, \ldots, \omega_N\}$, the commonsense score $\phi_l(W)$ of the logical sentence W composed at l-th hop expansion is defined as:



Relation Groups	Merged Relations	Relation Text
antonym/distinctfrom	antonym	is the antonym of
atlocation/locatednear	atlocation	is at location of
Relatedto/similarto/synonym	relatedto	is related to

Fig. 4 Knowledge statement transformation and cloze-based prompt construction.

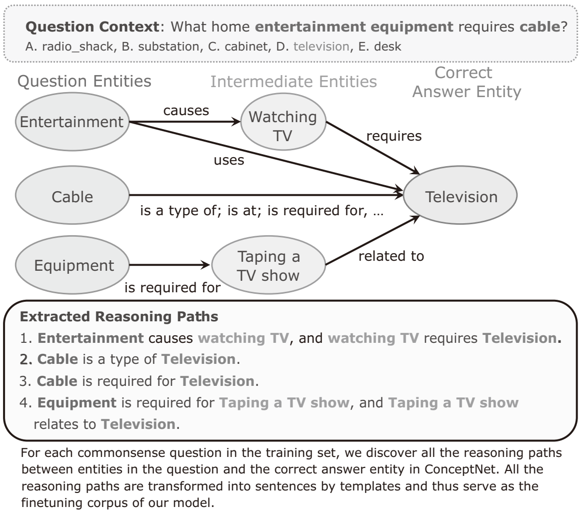


Fig. 5 Training Corpus Generation.

$$\phi_l(W) := \sum_{n=0}^{N} \log\left(\mathrm{p}_\theta\left(\omega_\mathrm{n} | \mathrm{W}_{\backslash n}\right)\right) / \mathrm{N}$$

where the $W_{\backslash n}$ indicates replacing the token ω_n to the [MASK], and the denominator N reduces the influence of the sentence length on the score prediction.

As we iteratively expand G_q , each ϕ_l (W) scores a unique reasoning path at a particular $l \in [1, L]$ depth in the graph. As marked in Fig. 3, a higher score ϕ_l (W) indicates the node v_j should be kept for the next $(l + 1)$ hop expansion.

2.3 Training Strategy and Answer Prediction.

In order to further enhance the PLM's reasoning capability, we propose to finetune PLMs on the knowledge examples constructed from ConceptNet. Specifically, we aim to enhance the p_θ's reasoning capability by correctly identifying the knowledge triplets on ConceptNet. As depicted in **Fig. 5**, given a commonsense question q = "What home entertainment equipment requires cable?" and its correct answer $\tilde{a}$ = "television", we identify reasoning paths [(v_1, r_1, v_2), ..., (v_{L-1}, r_{L-1}, v_L)] on G from each entity $c_q^{(i)}$ in c_q to $\tilde{a}$. Note that there may exist multiple paths $c_q^{(i)}$ to $\tilde{a}$.; e.g., "Cable is a type of Television" and "Cable is required for Television". Each reasoning path is then transformed as natural language sentences with templates as illustrated in the table of Fig. 4. We follow the standard masked language modeling task to finetune the model. By randomly masking a small portion (i.e., 15%) of tokens in each sentence, we aim to let the PLM comprehend the latent logic behind each retrieved reasoning path by learning to fill masks.

Answer Prediction: After we obtained the subgraph G_q consisting of all reasoning paths k within L-hop with a high commonsense score, each path $k_i \in$ k can be regarded as an individual supporting knowledge explanation to an answer a_i.

$$\log P_\theta(a_i | k_i) \propto \phi_\mathrm{L} = \sum_{l=1}^{L} \phi_\mathrm{l}$$

where the ϕ_L denotes the final score for each answer ai within L-hop and can be interpreted as approximating the likelihood of answer ai given a singular reasoning path $\{c \rightarrow v1 \rightarrow \cdots \rightarrow a\}$. To better improve efficiency, we utilize beam search to only keep high-confidence reasoning paths. We can thus pick the answer $\hat{a}$ and its reasoning path $\hat{k}$ with the highest score ϕ_L as the final answer and supporting knowledge.

3. Conclusion

A team comprised of members from NECLA and NEC Digital Business Platform Unit developed an off-the-shelf framework KEEP to predict answers for open-ended commonsense reasoning without requiring answer candidates and a pre-defined answer scope. By applying real-world tasks to address commonsense answering challenges, this technology has proven its effectiveness in a diverse array of business domains. We believe this work poses a new direction to automated commonsense reasoning under the zero-shot and open-ended setting in the Large Language Model era.

References

1) Mihir Kale and Abhinav Rastogi: Text-to-Text Pre-Training for Data-to-Text Tasks, The 13th International Conference on Natural Language Generation, pp. 97-102, 2020
https://aclanthology.org/2020.inlg-1.14/
2) Iz Beltagy, Kyle Lo and Arman Cohan: SciBERT: A Pretrained Language Model for Scientific Text, 2019 Conference on Empirical Methods in Natural Language Processing (EMNLP), 2019
https://arxiv.org/abs/1903.10676
3) Michihiro Yasunaga, Hongyu Ren, Antoine Bosselut, Percy Liang and Jure Leskovec: QA-GNN: Reasoning with Language Models and Knowledge Graphs for Question Answering, The 2021 Conference of the North American Chapter of the Association for Computational Linguistics (NAACL), 2021
https://arxiv.org/abs/2104.06378
4) Xikun Zhang, Antoine Bosselut, Michihiro Yasunaga et al.: GreaseLM: Graph REASoning Enhanced Language Models, 9th International Conference on Learning Representations (ICLR) , 2022
https://arxiv.org/abs/2201.08860
5) Bill Yuchen Lin, Xinyue Chen, Jamin Chen and Xiang Ren: KagNet: Knowledge-Aware Graph Networks for Commonsense Reasoning, 2019 Conference on Empirical Methods in Natural Language Processing (EMNLP), 2019
https://arxiv.org/abs/1909.02151
6) Bill Yuchen Lin, Haitian Sun et al.,:Differentiable Open-Ended Commonsense Reasoning, the 2021 Conference of the North American Chapter of the Association for Computational Linguistics (NAACL), 2021
https://arxiv.org/abs/2010.14439
7) Maria Becker, Katharina Korfhage and Anette Frank: COCO-EX: A Tool for Linking Concepts from Texts to ConceptNet, The 16th Conference of the European Chapter of the Association for Computational Linguistics (EACL), 2021
https://aclanthology.org/2021.eacl-demos.15/

Authors' Profiles

ZHAO Xujiang
Researcher
NEC Laboratories America

LIU Yanchi
Researcher
NEC Laboratories America

CHENG Wei
Senior Researcher
NEC Laboratories America

OISHI Mika
Software and System Engineering Department

OSAKI Takao
Director
Software and System Engineering Department

MATSUDA Katsushi
Professional
Software and System Engineering Department

CHEN Haifeng
Department Head
NEC Laboratories America

Information about the NEC Technical Journal

Thank you for reading the paper.
If you are interested in the NEC Technical Journal, you can also read other papers on our website.

Link to NEC Technical Journal website

Japanese | English

Vol.17 No.2 Special Issue on Revolutionizing Business Practices with Generative AI
– Advancing the Societal Adoption of AI with the Support of Generative AI Technologies

Remarks for Special Issue on Revolutionizing Business Practices with Generative AI
Approaches to Generative AI Technology: From Foundational Technologies to Application Development and Guideline Creation

Papers for Special Issue

Market Application of Rapidly Spreading Generative AI
NEC Innovation Day 2023: NEC’s Generative AI Initiatives
Streamlining Doctors’ Work by Assisting with Medical Recording and Documentation
Using Video Recognition AI x LLM to Automate the Creation of Reports
Understanding of Behaviors in Real World through Video Analysis and Generative AI
Automated Generation of Cyber Threat Intelligence
NEC Generative AI Service (NGS) Promoting Internal Use of Generative AI
Utilization of Generative AI for Software and System Development
LLMs and MI Bring Innovation to Material Development Platforms
Disaster Damage Assessment Using LLMs and Image Analysis

Fundamental Technologies that Enhance the Potential of Generative AI
NEC’s LLM with Superior Japanese Language Proficiency
NEC’s AI Supercomputer: One of the Largest in Japan to Support Generative AI
Towards Safer Large Language Models (LLMs)
Federated Learning Technology that Enables Collaboration While Keeping Data Confidential and its Applicability to LLMs
Large Language Models (LLMs) Enable Few-Shot Clustering
Knowledge-enhanced Prompt Learning for Open-domain Commonsense Reasoning
Foundational Vision-LLM for AI Linkage and Orchestration
Optimizing LLM API usage costs with novel query-aware reduction of relevant enterprise data

For AI Technology to Penetrate Society
Movements in AI Standardization and Rule Making and NEC Initiatives
NEC’s Initiatives on AI Governance toward Respecting Human Rights
Case Study of Human Resources Development for AI Risk Management Using RCModel

NEC Information

2023 C&C Prize Ceremony

Vol.17 No.2
June 2024

Special Issue TOP