MERIt: Meta-path Guided Contrastive Learning for Logical Reasoning

SHANDONG UNIVERSITY

Fangkai Jiao, Yangyang Guo, Xuemeng Song, Liqiang Nie

NUS National University of Singapore

Introduction

Logical reasoning requires to correctly infer the semantic relations with respect to the constituents among different sentences. As shown below, to find the correct answer for the given question, one needs to extract the logical structures residing in a pair of each option and the whole context, and justify its reasonableness.

Although the pretrained language models have achieved significant progress on several benchmarks, they still struggle on inferring the logical relations beneath natural language since the goal of general pre-training, i.e., masked language modeling, deviates largely from that of logical reasoning.

In this paper, we present MERIt, a MEta-path guided contrastive learning approach for logical ReasonIng of text. To the best of our knowledge, we are the first to explore self-supervised pre-training for logical reasoning. Our contribution is summarized as follows:

1. We successfully employ the meta-path strategy to mine the potential logical structure in raw text. It is able to automatically generate negative candidates for contrastive learning via logical relation editing.
2. We propose a simple yet effective counterfactual data augmentation method to eliminate the information shortcut during pre-training.
3. We evaluate our method on two logical reasoning tasks, LogiQA and ReClor. The experimental results show that our method achieves the new state-of-the-art performance on two benchmark datasets.

Context: Economist: **(1)** A country's rapid emergence from an economic recession (r_1) requires **(2)** substantial new investment in that country's economy. Since **(3)** people's confidence in the economic policies of their country (r_2) is a precondition for **(2)** any new investment, **(4)** countries that put collective goals before individuals' goals (r_3) cannot **(1)** emerge quickly from an economic recession.

Question:

Which one of the following, if assumed, enables the economist's conclusion to be properly drawn?

Options:

A. People in **(4)** countries that put collective goals before individuals' goals (r_4) lack **(3)** confidence in the economic policies of their countries.

B. A country's economic policies are the most significant factor determining whether that country's economy will experience a recession.

C. If the people in a country that puts individuals' goals first are willing to make new investments in their country's economy, their country will emerge quickly from an economic recession.

D. No new investment occurs in any country that does not emerge quickly from an economic recession.

Answer: A

Logic Structure: $(4) \xrightarrow{r_4} (3) \xrightarrow{r_2} (2) \xrightarrow{\bar{r}_1} (1) \Leftrightarrow (4) \xrightarrow{r_3} (1)$

Fig. 1: An example of logical reasoning from ReClor.

Method

Q1: How to discover the logical structure in a raw document?
A1: The logical reasoning process can be formulated as:

$$\langle v_i, r_{i,j}, v_j \rangle \leftarrow \left(v_i \xrightarrow{r_{i,i+1}} v_{i+1} \xrightarrow{r_{i+1,i+2}} \cdots \xrightarrow{r_{j-1,j}} v_j \right) (1)$$

Take entity e_i as the logical variable v_i:

$$\langle e_i, r_{i,j}, e_j \rangle \leftarrow \left(e_i \xrightarrow{r_{i,i+1}} e_{i+1} \xrightarrow{r_{i+1,i+2}} \cdots \xrightarrow{r_{j-1,j}} e_j \right) (2)$$

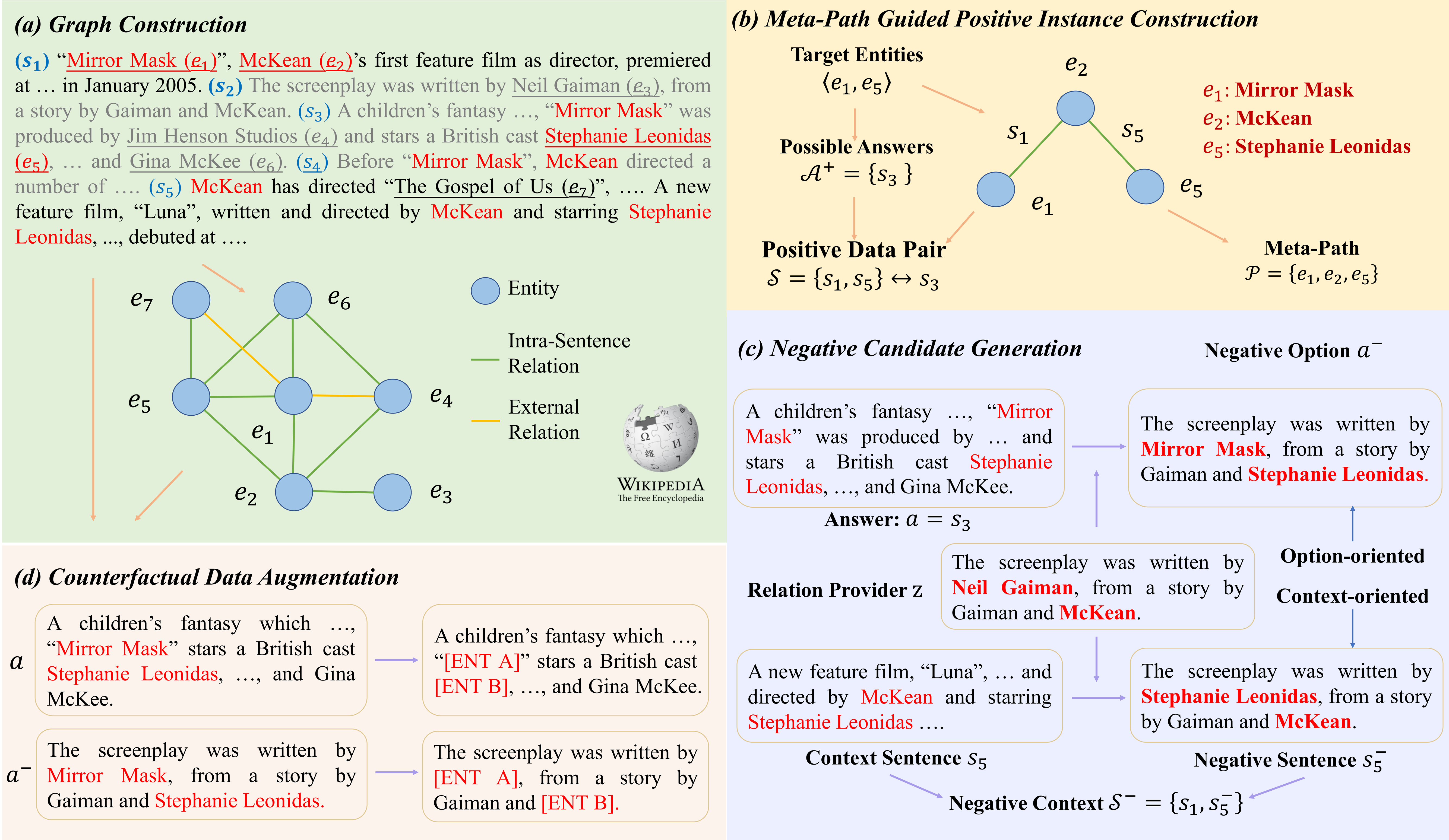
The right part is a meta-path connecting $\langle e_i, e_j \rangle$ through indirect relations, while the left part is a direct relation triplet. We then employ Eqn.2 to discover the logical structure in a document.

Q2: How to construct logically consistent instance pair?
A2: We assume the pair defined in Eqn. 2 is logically consistent under the same context, following which we construct the positive context-option pair. An example is shown in Fig. 2 (a) and (b). The external relation are annotated labels from the Wikipedia knowledge base while the intra-sentence relation indicates that a pair of entity has been mentioned in the same sentence. We employ the Depth-first Search to find the meta-path connecting an entity pair and treat the sentences directly mentioning them as the answer candidates.

Q3: How to generate negative instances?
A3: We generate the negative instances through the modification of the relations in the positive instance pair, which is further implemented via entity replacement as shown in Fig. 2 (c).

Q4: Are there any possible trivial solutions during the contrastive learning process?
A4: Yes. The commonsense knowledge entailed in the pre-trained language models may lead to trivial solutions by simply checking the factuality of each option.

Q5: How to eliminate the shortcut?
A5: We devise a counterfactual data augmentation strategy by replacing the original entities with those from different documents (Fig. 2 (d)). The operation aims at making both the positive and negative instance pairs contradictng to the world knowledge.



Optimization

The objective of logical reasoning can be formulated as:

$$\mathcal{L} = L(x, x^+, \mathcal{X}^-) = -\log \frac{\exp f(x, x^+)}{\sum_{x' \in \mathcal{X}^- \cup \{x^+\}} \exp f(x, x')} . (1)$$

In our method, the objective of the stage can thus be defined as:

$$\mathcal{L}_{OCL} = L(\mathcal{S}, a, \mathcal{A}^-), (2)$$
$$\mathcal{L}_{CCL} = L(a, \mathcal{S}, \mathcal{C}^-), (3)$$

where $\mathcal{S}$ is the set of context sentences, a is the answer candidate, $\mathcal{A}^-$ and $\mathcal{C}^-$ are the sets of the constructed negative options and contexts, respectively.

To avoid the catastrophic forgetting problem, we also add the MLM objective during pre-training.

Experiment

Model / Dataset	ReClor Dev	ReClor Test	ReClor Test-E	ReClor Test-H	LogiQA Dev	LogiQA Test
RoBERTa	62.6	55.6	75.5	40.0	35.0	35.3
DAGN	65.2	58.2	76.1	44.1	35.5	38.7
DAGN (Aug)	65.8	58.3	75.9	44.5	36.9	39.3
LReasoner (RoBERTa)‡	64.7	58.3	77.6	43.1	—	—
Focal Reasoner	66.8	58.9	77.1	44.6	**41.0**	40.3
MERIt	66.8	59.6	78.1	45.2	40.0	38.9
MERIt + LReasoner	67.4	60.4	78.5	46.2	—	—
MERIt + Prompt	**69.4**	**61.6**	79.3	**47.8**	39.9	**40.7**
MERIt + Prompt + LReasoner	67.3	61.4	**79.8**	46.9	—	—
ALBERT	69.1	66.5	76.7	58.4	38.9	37.6
MERIt (ALBERT)	74.2	70.1	81.6	61.0	43.7	**42.5**
MERIt (ALBERT) + Prompt	**74.7**	**70.5**	**82.5**	**61.1**	**46.1**	41.7
max						
LReasoner (RoBERTa)	66.2	62.4	81.4	47.5	38.1	40.6
MERIt	67.8	60.7	79.6	45.9	**42.4**	41.5
MERIt + Prompt	**70.2**	**62.6**	80.5	**48.5**	39.5	**42.4**
LReasoner (ALBERT)	73.2	70.7	81.1	62.5	41.6	41.2
MERIt (ALBERT)	73.2	71.1	83.6	61.3	43.9	**45.3**
MERIt (ALBERT) + Prompt	**75.0**	**72.2**	**82.5**	**64.1**	**45.8**	43.8

Tab. 1: The accuracy of different models on ReClor and LogiQA. We use RoBERTa-large and ALBERT-xxlarge as the backbone.

Model	Dev	Test	Test-E	Test-H
DeBERTa-v2-xlarge	76.7	71.0	83.8	60.9
MERIt (DeBERTa-v2-xlarge)	**78.0**	**73.1**	**86.2**	**64.4**
DeBERTa-v2-xxlarge	78.3	75.3	84.0	68.4
MERIt (DeBERTa-v2-xxlarge)	**80.6**	**78.1**	**84.6**	**72.9**

Tab. 3: The results on ReClor with DeBERTa as the backbone.

Model	Dev	Dev (P.)	Test	Test (P.)
MERIt	66.8	69.4	59.6	61.6
- DA	63.0	64.5	57.9	59.8
+ DA^2	65.3	67.8	**60.2**	61.3
+ DA^3	66.2	68.0	59.3	**61.9**
- Option-oriented CL	63.8	65.4	58.9	61.5
- Context-oriented CL	64.0	66.5	58.8	60.2
- Meta-Path	64.8	65.1	58.0	60.8

Tab. 2: Performance comparisons on ReClor between different variants of MERIt. *DA* means data augmentation and DA^N refers to 1:N ratio of the original data to the augmented data. P. is short for Prompt Tuning.

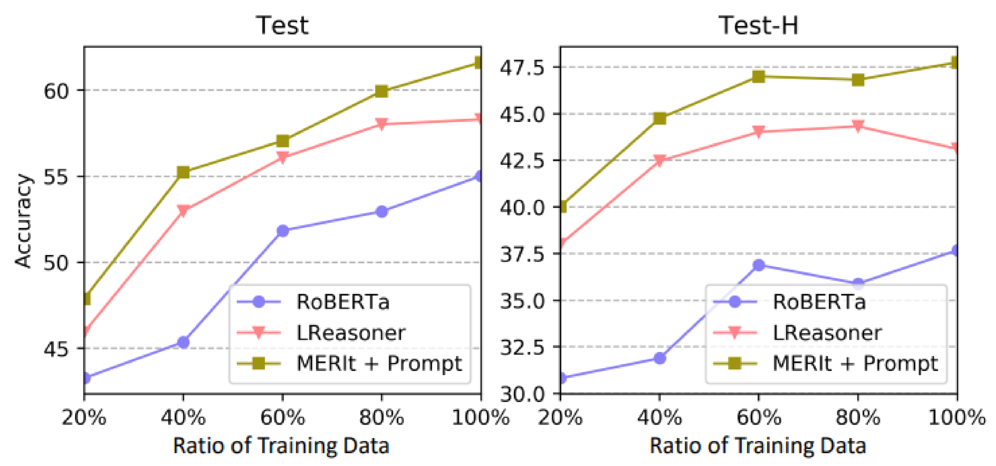


Fig. 2: The results on the test set (left) and the test-H set (right) of ReClor with different ratio of training data.