No Need for Large-Scale Search: Exploring Large Language Models in Complex Knowledge Base Question Answering

Shouhui Wang, Biao Qin*

School of Information, Renmin University of China

Background & Motivation

➢ The existing gap between natural language questions and structured knowledge representations;

➢ Large search and reasoning spaces, and the challenge of ranking a massive number of candidate logical forms;

➢ Large language models (LLMs) offers an opportunity to address above challenges.
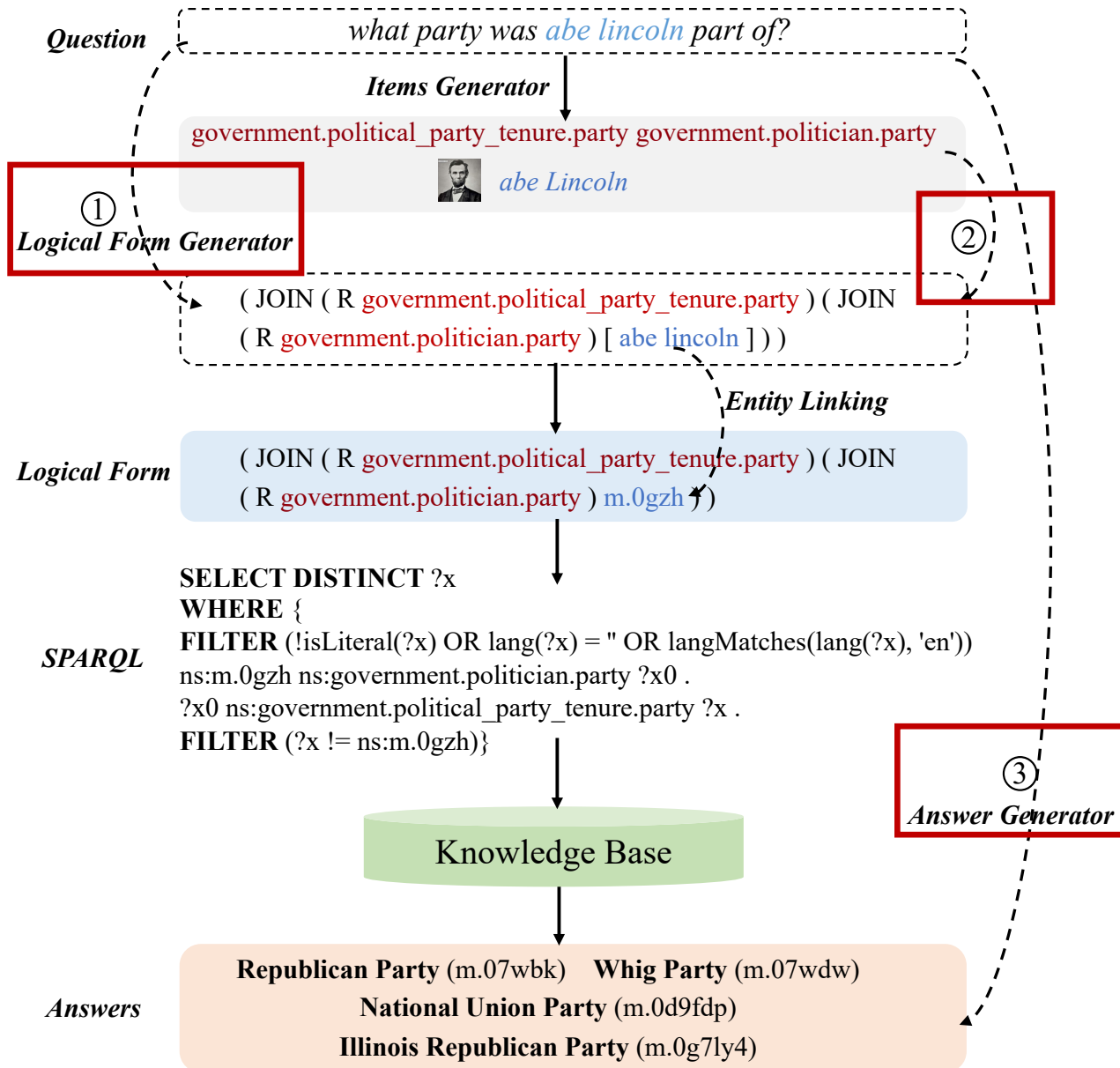
We adopted LLMs to propose a **T**hree-step **F**ine-tune **S**trategy based on a large language model to implement the KBQA system (**TFS-KBQA**).
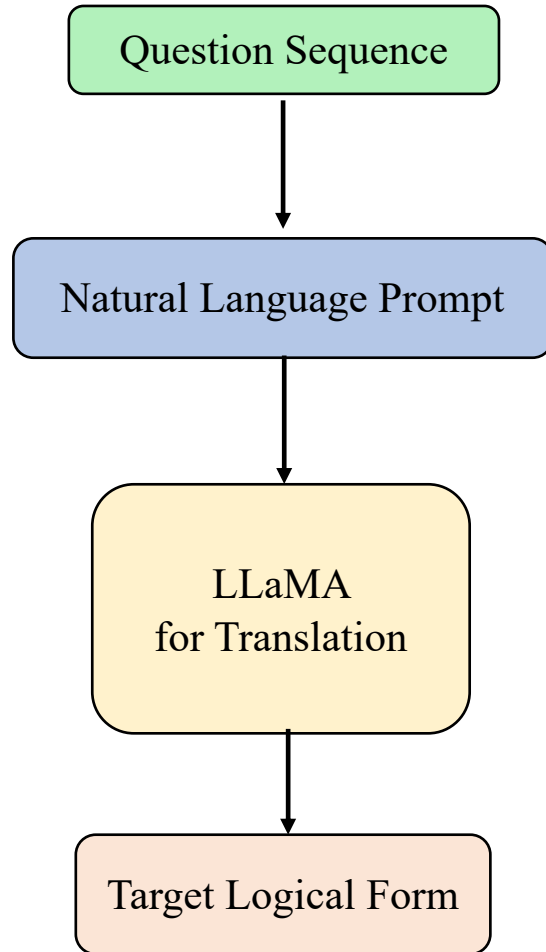
Strategy 1: Utilize an LLM to directly convert the question into its corresponding logical form.

Strategy 2: Leverage an LLM to transform the question into a logical form in two steps.

Strategy 3: Use an LLM to directly answer the natural language questions.

Question Sequence

Natural Language Prompt
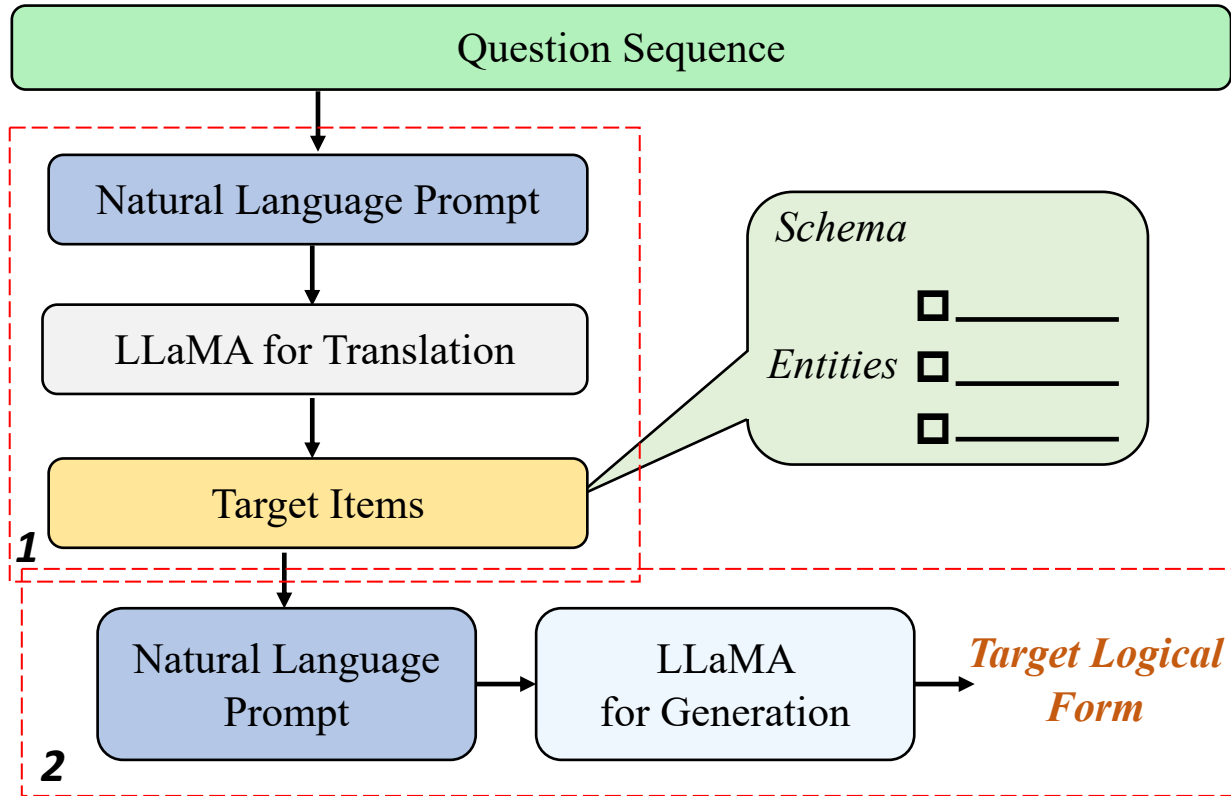
LLaMA
for Translation

Target Logical Form

T1

Instruction (I_1): *Please translate the following questions into their corresponding logical forms.*

Input: *Who was vice president after kennedy died?*

Output: *(join (r government.us_president.vice_president) [john f. kennedy])*

Methodology-Strategy 2

Question Sequence

1
Natural Language Prompt
↓
LLaMA for Translation
↓
Target Items

Schema

Entities ☐_____
☐_____
☐_____

2
Natural Language Prompt → LLaMA for Generation → *Target Logical Form*

T2

t1

Instruction (I_{t1}): *Please translate the following questions into their corresponding relations and extract entities mentioned in the questions.*
Input: *Who was vice president after kennedy died?*
Output: *government.us_president.vice_president ; john f. kennedy*
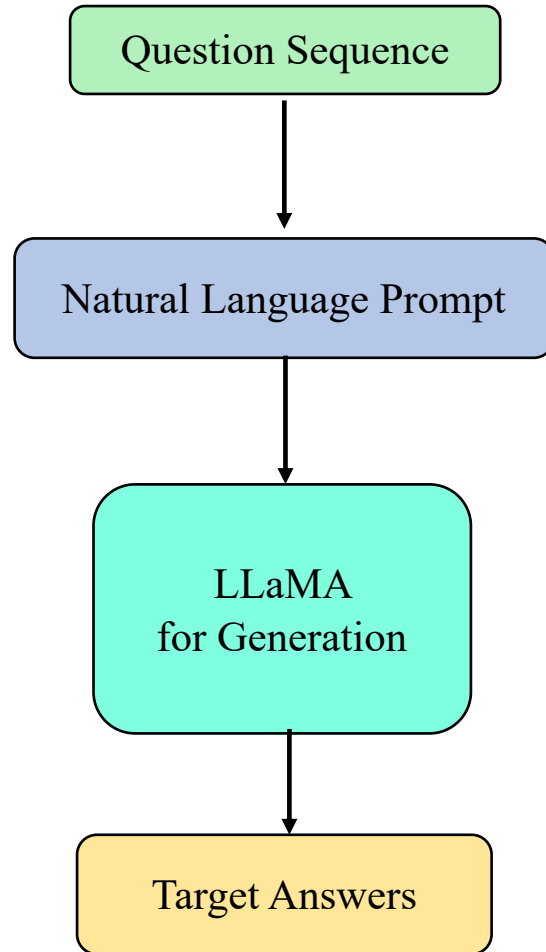
t2

Instruction (I_{t2}): *Please generate the target logical form according to the following sentence.*
Input: *government.us_president.vice_president ; john f. kennedy*
Output: *(join (r government.us_president.vice_president) [john f. kennedy])*

1

T3

Instruction (I_3): *Please provide answers to the following questions based on your understanding.*

Input: *Who was vice president after kennedy died?*

Output: *Lyndon B. Johnson*

Experiments

Method	F1	Hits@1
IR-based Methods		
TransferNet (Shi et al., 2021)	–	71.4
NSM (He et al., 2021)	67.4	74.3
SR (Zhang et al., 2022)	74.5	**83.2**
SP-based Methods		
QGG (Lan and Jiang, 2020)	74.0	–
ReTraCk (Chen et al., 2021)	71.0	71.6
RnG-KBQA (Ye et al., 2022)	75.6	–
ArcaneQA (Gu and Su, 2022)	72.8	–
Program Transfer (Cao et al., 2022)	76.5	74.6
TIARA (Shu et al., 2022)	76.7	73.9
Pangu (Gu et al., 2023)	**79.6**	–
FC-KBQA (Zhang et al., 2023)	76.9	–
TFS-KBQA (Ours)	**79.9**	79.8
-Only use LLaMA-2-13b	79.5	79.5
-Only use LLaMA-2-7b	78.5	78.4
-w/o ReFinED entity linking	78.5	78.6
-w/o ELQ entity linking △	79.2	79.2
-w/o Google API entity linking	78.2	78.3
-w/o Strategy 1	78.5	78.6
-w/o Strategy 2	79.7	79.6
-w/o Strategy 3	77.4	76.9

WebQSP

Method	F1
PullNet (Sun et al., 2019)	47.2
QGG (Lan and Jiang, 2020)	40.4
NSM (He et al., 2021)	48.8
TransferNet (Shi et al., 2021)	48.6
RnG-KBQA (Ye et al., 2022)	42.3
Program Transfer (Cao et al., 2022)	**58.7**
FC-KBQA (Zhang et al., 2023)	53.1
TFS-KBQA (Ours)	**63.6**
-Only use LLaMA-2-13b	62.9
-Only use LLaMA-2-7b	62.2
-w/o ReFinED entity linking	61.1
-w/o Google API entity linking	61.4
-w/o Strategy 1	60.0
-w/o Strategy 2	62.7
-w/o Strategy 3	49.1

ComplexWebQuestion

Method	F1
Constraint-Based (Bao et al., 2016)	42.3
Embedding-Based (Luo et al., 2018)	42.8
QGG (Lan and Jiang, 2020)	**43.3**
TFS-KBQA (Ours)	**44.0**
-Only use LLaMA-2-13b	43.7
-Only use LLaMA-2-7b	41.1
-w/o ReFinED entity linking	42.8
-w/o Google API entity linking	42.7
-w/o Strategy 1	43.4
-w/o Strategy 2	43.3
-w/o Strategy 3	30.4

ComplexQuestions

Experiments

Example
[*Strategy 1*]
Question: *What currency is used in the jurisdiction where the Cabinet of Peru is located?*
Predict: (join (r location.country.currency_used) (join government.governmental_jurisdiction.government [cabinet of peru]))
Golden: (join (r location.country.currency_used) (join government.governmental_jurisdiction.governing_officials (join government.government_position_held.governmental_body [cabinet of peru])))
[*Strategy 2*]
Question: *What languages are spoken where Haitian Creole is spoken?*
Predict (t1): location.country.languages_spoken location.country.languages_spoken ; haitian creole
Golden: location.country.languages_spoken location.country.official_language ; haitian creole
Predict (t2): (join (r location.country.languages_spoken) (join location.country.languages_spoken [haitian creole]))
Golden: (join (r location.country.languages_spoken) (join location.country.official_language [haitian creole]))
[*Strategy 3*]
Question: *What artist recorded Elf's Lament?*
Predict: singer, actor, film producer, songwriter, record producer, radio personality, music artist, artist, songwriter
Golden: actor, singer, songwriter
[*Entity Linking*]
Question: *What is the type of government practiced in the country where the Israeli Lira is used?*
Logical Form: (join (r location.country.form_of_government) (join location.country.currency_formerly_used [israeli lira]))
Target Logical Form: (join (r location.country.form_of_government) (join location.country.currency_formerly_ used m.04r5x5 m.036g0k))

Conclusion

✓ Fully leverage the advantages of LLMs in the NLP field;

✓ Further improve KBQA performance on complex datasets, notably achieving an F1 score of 79.9% on WebQSP;

✓ Overcome the challenges of a large search and reasoning space;

✓ Avoided the ranking of massive candidate logical forms common in previous KBQA methods.

Q & A