Complex Question answering on Knowledge Bases and Text

Huang Xin

Complexity of ComplexQA dataset

- **Question with constraint**

1. What educational institution with men's basketball sports team named Temple Owls did Kevin Hart go to school? (and relation type)

```
ns:m.0ggp1n ns:people.person.education ?y .
?y ns:education.education.institution ?x .
?x ns:education.educational_institution.sports_teams ns:m.03yfq3l .
```

2. What country located in the Greenwich Mean Time Zone that is the main trading partner of China? (or relation type)

```
{
  ns:m.0d05w3 ns:location.statistical_region.places_exported_to ?y .
  ?y ns:location.imports_and_exports.exported_to ?x .
}
UNION
{
  ns:m.0d05w3 ns:location.statistical_region.places_imported_from ?y .
  ?y ns:location.imports_and_exports.imported_from ?x .
}
```

Complexity of ComplexQA dataset

- **Question with order by and filter operation**

  1. when is the last time the the team has a team moscot named Lou Seal won the world series?

```
?c ns:sports.sports_team.team_mascot ns:m.03_dwn .
?c ns:sports.sports_team.championships ?x .
?x ns:time.event.start_date ?sk0 .
}
ORDER BY DESC(xsd:datetime(?sk0))
```

  2. Who is the leader of the country that has national anthem Allhu Akbar?

```
FILTER(NOT EXISTS {?y ns:government.government_position_held.from ?sk0} ||
EXISTS {?y ns:government.government_position_held.from ?sk1 .
FILTER(xsd:datetime(?sk1) <= "2015-08-10"^^xsd:dateTime) })
FILTER(NOT EXISTS {?y ns:government.government_position_held.to ?sk2} ||
EXISTS {?y ns:government.government_position_held.to ?sk3 .
FILTER(xsd:datetime(?sk3) >= "2015-08-10"^^xsd:dateTime)
```

Complexity of ComplexQA dataset

- **Ambiguous Question**

  1. what movies has miley cyrus been in

  ns:m.0bdxs5 ns:film.actor.film ?y .
  ?y ns:film.performance.film ?x .
  ?x ns:film.film.starring ?c .
  ?c ns:film.performance.character ns:g.121cr_v1 .

  2. Which countries speak Portugsese?

  ?c ns:common.topic.image ns:m.0ccjyhs .
  ?c ns:language.human_language.countries_spoken_in ?x .

Complexity of ComplexQA dataset

- **Ambiguous Question**

  1. what movies has miley cyrus been in

  ns:m.0bdxs5 ns:film.actor.film ?y .
  ?y ns:film.performance.film ?x .
  ?x ns:film.film.starring ?c .
  ?c ns:film.performance.character ns:g.121cr_v1 .

  2. Which countries speak Portugsese?

  ?c ns:common.topic.image ns:m.0ccjyhs .
  ?c ns:language.human_language.countries_spoken_in ?x .

Complexity of ComplexQA dataset

- **Other complexity characteristic**

  a. Questions can be answered by different hop

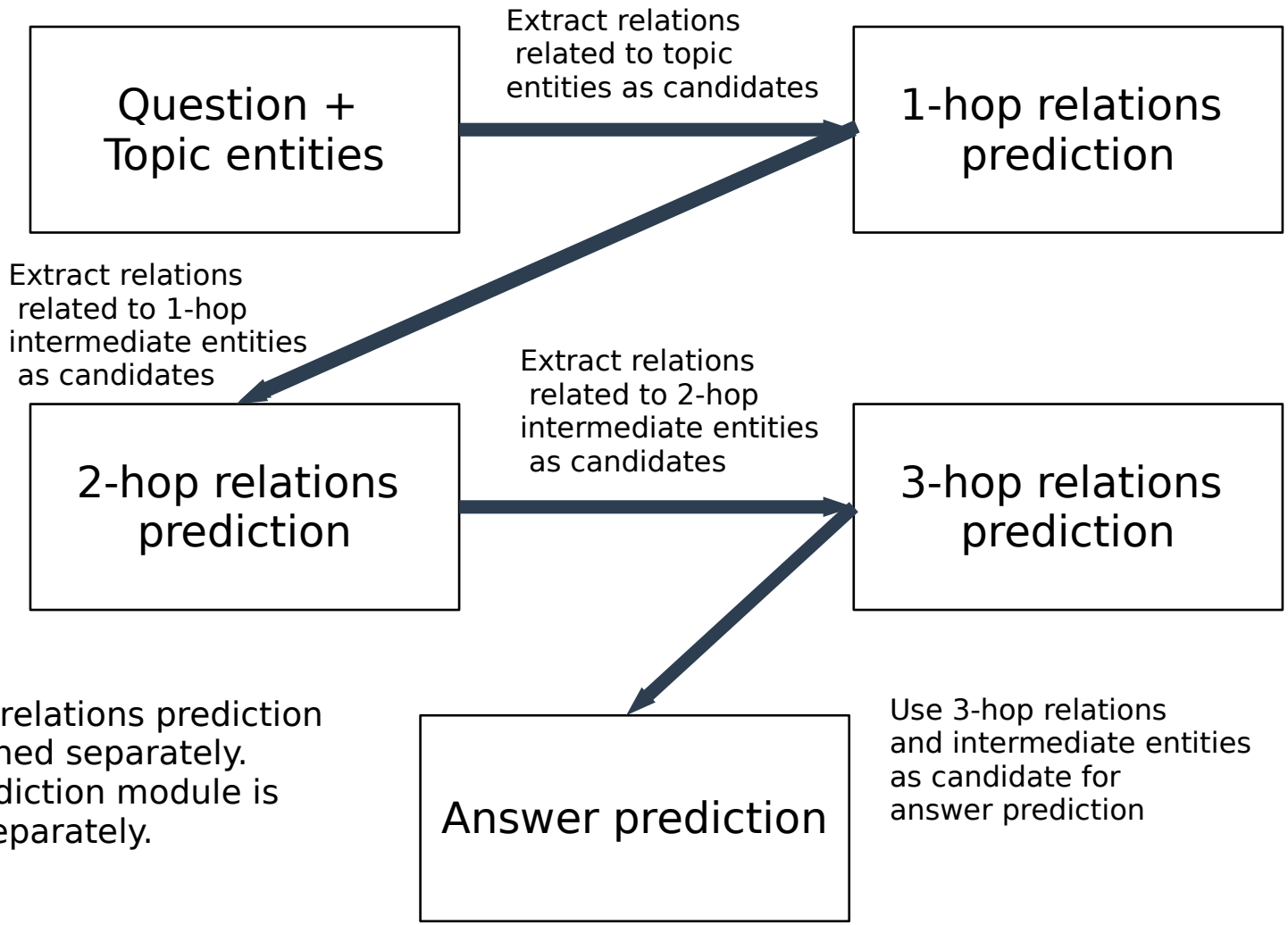
  b. Loop question which answers are topic entities.

  c. Questions with multiple topic entities and multiple answers.

  d. Questions with variant length of path leading to answers.

  e. Huge knowledge graph containing 8k+ relations could be hard for finding relations.

New Approach applied to complexQA



Note:
1. Each K-hop relations prediction module is trained separately.
2. Answer prediction module is also trained separately.

New Approach applied to complexQA

K-hop relation prediction module

For each K-hop relation prediction, we would use 1 to K-1 relations as constraint feature.

For example, for a 3-hop question, we would use relations including [r1, r2, *] as features for the prediction of r3.

$$q_{emb} = LSTM(q_{1,} q_{2,..}, q_n)$$
$$r_{emb} = LSTM(r_{1,} r_{2,} r_3)$$
$$r_{emb} = Dense(r_{emb})$$
$$S(q, r) = sigmoid(q * r)$$

The relation embedding is looked up in the relation embedding table.
The final relations are ranked by their similarity score between question and relation embedding.
A special relation “EOD” is added to specify this question doesn’t need further relation prediction.

Subgraph generation module performance

Performance of K-hop relation prediction

We always choose K=3 for all questions.
So the 3-hop relation prediction is the final subgraph generation performance.

Metrics	1-hop relation prediction	2-hop relation prediction	3-hop relation prediction
Hit@1	80.2%	65.3%	69.9%
Recall (Top-3)	83.5%	70.6%	71.9%

The metric is measured by the exact relation chain occurred in the ground truth path.

Example of K-hop relation prediction

Question

when is the last time the giants won the world series

Ground truth path

[['sports.sports_team.team_mascot']]
[['sports.sports_team.team_mascot', 'sports.sports_team.championships']]

Predicted path

['sports.sports_team.team_mascot',
'sports.sports_team.championships',
 'EOD']