IBM Research AI

SMU SINGAPORE MANAGEMENT UNIVERSITY
School of Information Systems

A Co-Matching Model for Multi-choice Reading Comprehension

Shuohang Wang[1], Mo Yu[2], Shiyu Chang[2], Jing Jiang[1]

[1]Singapore Management University
[2]IBM Research

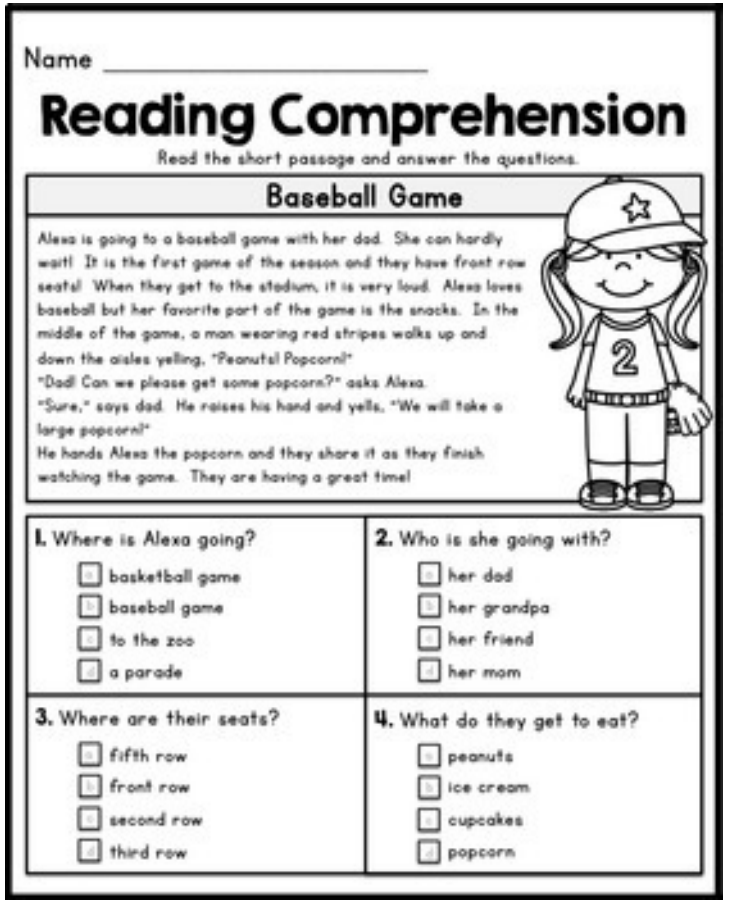
Reading Comprehension

- The task: to answer questions given a passage of text
- Datasets
  - CNN/Daily Mail [Hermann et al. 2015]
  - Children's Book Test [Hill et al. 2016]
  - SQuAD [Rajpurkar 2016]
  - MCTest [Richardson et al. 2013]
  - RACE [Lai et al. 2017]
  - NarrativeQA [Kocisky et al. 2018]

Cloze Style

Answer Span Extraction

Multi-choice Questions



Name _____

Reading Comprehension

Read the short passage and answer the questions.

Baseball Game

Alexa is going to a baseball game with her dad. She can hardly wait! It is the first game of the season and they have front row seats! When they get to the stadium, it is very loud. Alexa loves baseball but her favorite part of the game is the snacks. In the middle of the game, a man wearing red stripes walks up and down the aisles yelling, "Peanuts! Popcorn!"
"Dad! Can we please get some popcorn?" asks Alexa.
"Sure," says dad. He raises his hand and yells, "We will take a large popcorn!"
He hands Alexa the popcorn and they share it as they finish watching the game. They are having a great time!

1. Where is Alexa going?
- a basketball game
- b baseball game
- c to the zoo
- d a parade

2. Who is she going with?
- a her dad
- b her grandpa
- c her friend
- d her mom

3. Where are their seats?
- a fifth row
- b front row
- c second row
- d third row

4. What do they get to eat?
- a peanuts
- b ice cream
- c cupcakes
- d popcorn

RACE

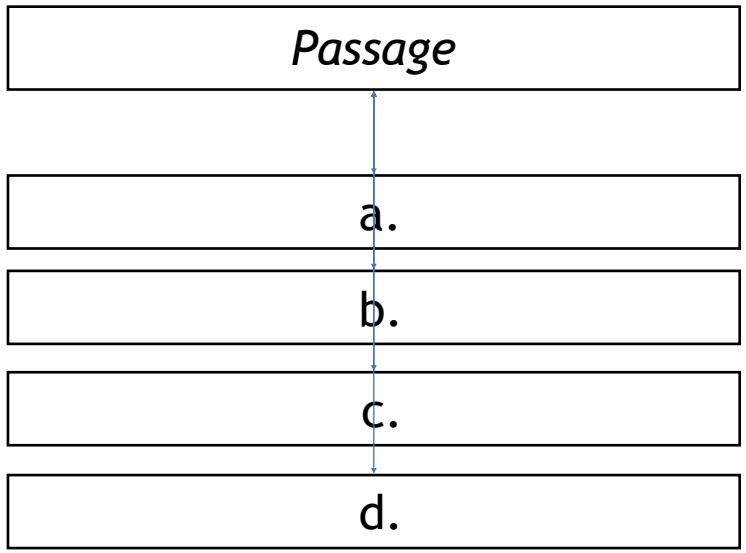
- Passage: *My father wasn't a king, he was a taxi driver, but … … I met Blandy at a party and he asked if I'd like to buy the island. Of course I said yes but I had no money-I was just an art teacher. I tried to find some business partners, who all thought I was crazy. So I sold some of my possessions, put my savings together and bought it … …*

- Question: *How did the author get the island?*
  - a. *It was a present from Blandy.*
  - b. *The king sold it to him.*
  - c. *He bought it from Blandy.*
  - d. *He inherited from his father.*

Challenge: to jointly model passage, question and candidate answers

Related Work

- Converted to sequence pair matching [Yin et al., 2016]
  - Each candidate answer is concatenated with the question
  - The concatenated sequences are matched against the passage



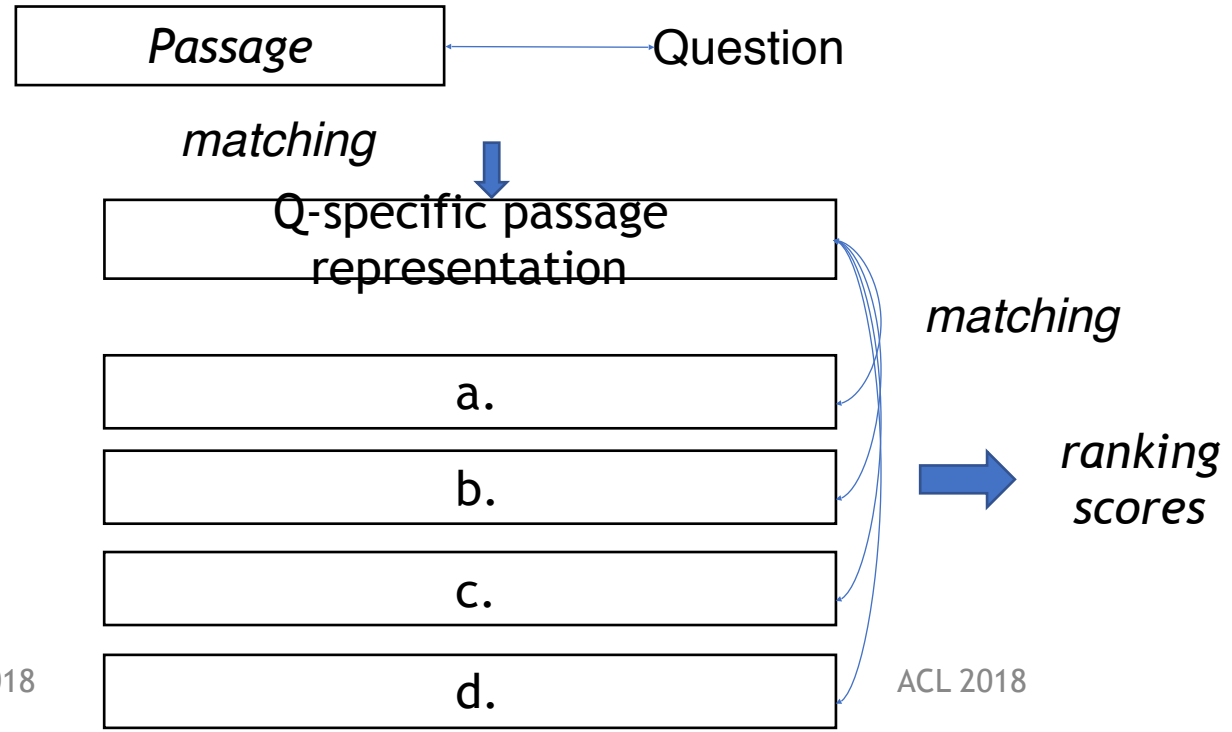
matching

ranking scores

Limitation:
Question and answers are not clearly separated. Interaction information between a question and an answer is lost.

Related Work

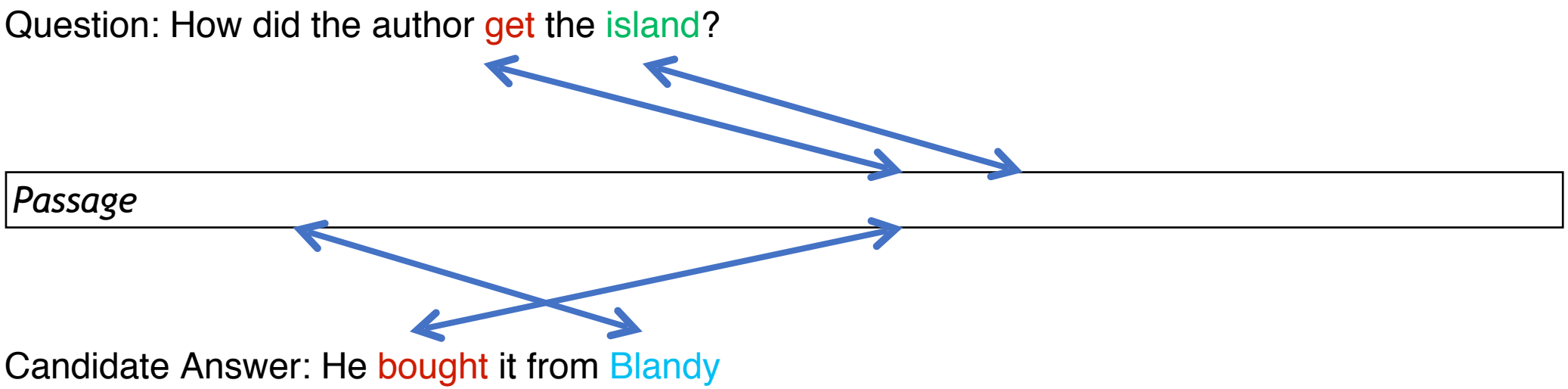
- Matching sequences pair by pair [Lai et al., 2017]
  - Match passage and question first
  - Then this representation is used to match candidate answers



Limitation:
Matching *P* & *Q* may not give meaningful representations for questions like "Which statement of the following is true?"

Our Solution

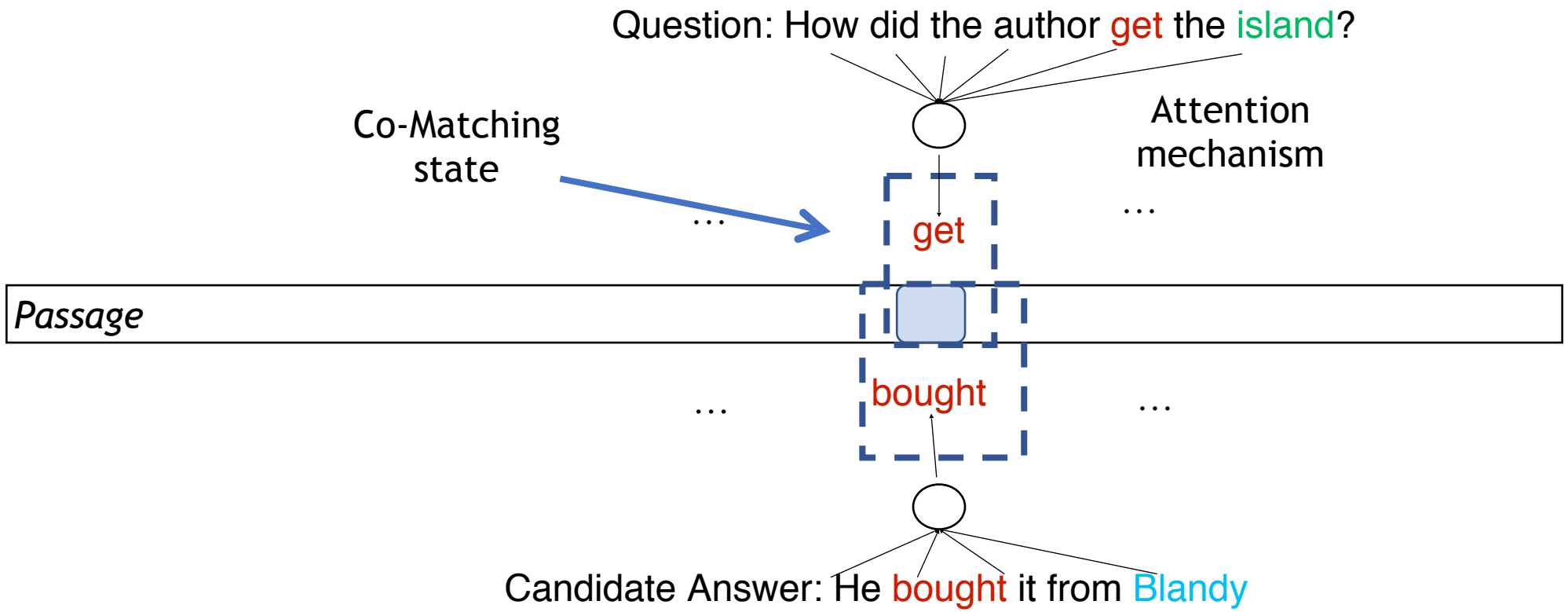
- Co-match each sentence in the passage with the question and the candidates answers separately.
- Make use of the alignments between sequences as follows:



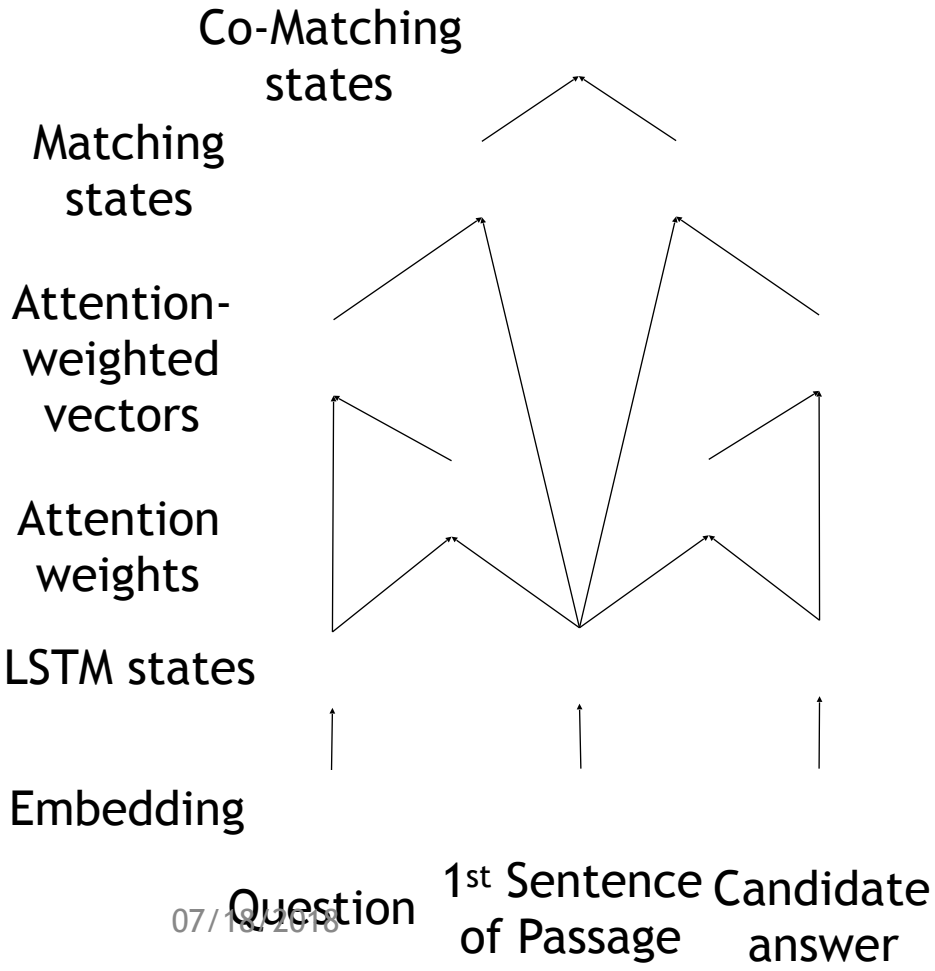
- Hierarchically aggregate the co-matching representations of (sentence, question, answer) triplets for final scoring.

Co-Matching

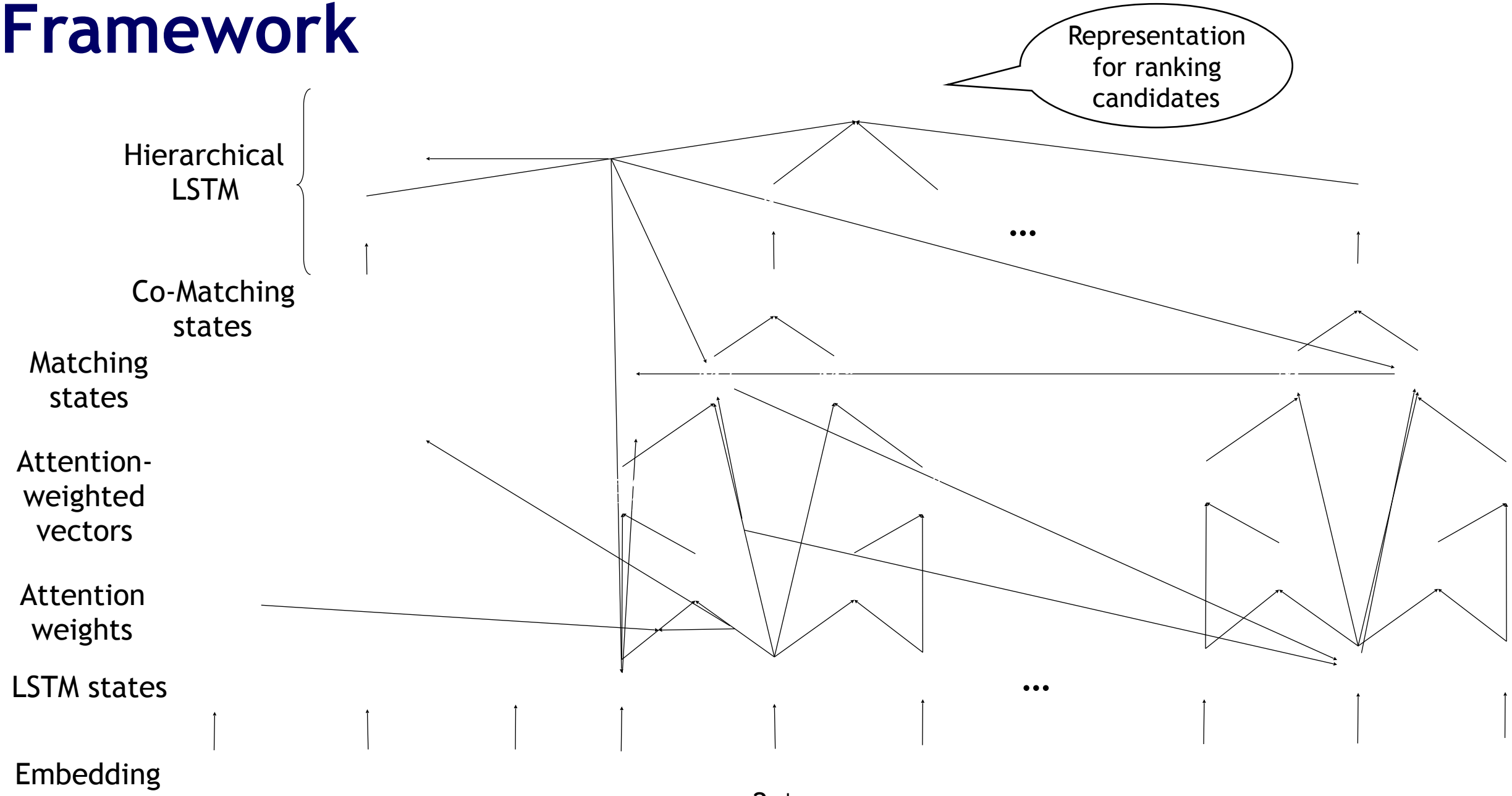
- For every word in sentence, we match it with the attention-weighted vectors computed based on the question and the candidate answer, respectively.



Framework


Co-Matching states
Matching states
Attention-weighted vectors
Attention weights
LSTM states
Embedding
Question
1st Sentence of Passage
Candidate answer

Framework


Representation for ranking candidates
Hierarchical LSTM
Co-Matching states
Matching states
Attention-weighted vectors
Attention weights
LSTM states
Embedding
Question
1st Sentence of Passage
Candidate answer
2nd Sentence of Passage
Nth Sentence of Passage

Experiments

	RACE-M	RACE-H	RACE
Random	24.6	25.0	24.9
Sliding Window	37.3	30.4	32.2
Stanford AR	44.2	43.0	43.3
GA	43.7	44.2	44.1
ElimiNet	-	-	44.7
HAF	45.3	47.9	47.2
MUSIC	51.5	45.7	47.4
Hier-Co-Matching	**55.8***	**48.2***	**50.4***
- Hier-Aggregation	54.2	46.2	48.5
- Co-Matching	50.7	45.6	46.4
Turkers	85.1	69.4	73.3
Ceiling	95.4	94.2	94.5

Our Hier-Co-Matching achieved the best performance compared with previous work.

We studied two key factors:
(1) the co-matching module
(2) the hierarchical aggregation approach

Conclusions

- We proposed a hierarchical co-matching model for answering multi-choice reading comprehension questions.
- We showed that our model could achieve state-of-the-art performance on the RACE dataset.
- There is still much room for improvement on RACE given the low absolute performance.
  - Latest results by OpenAI: 59%

Paper: 

Code: 

References

- Guokun Lai, Qizhe Xie, Hanxiao Liu, Yiming Yang, and Eduard Hovy. 2017. “RACE: Large-scale reading comprehension dataset from examinations.” In Proceedings of EMNLP.
- Wenpeng Yin, Sebastian Ebert, and Hinrich Schutze. 2016. “Attention-based convolutional neural network for machine comprehension.” In Proceedings of NAACL.
- Soham Parikh, Ananya B. Sai, Preksha Nema, Mitesh M. Khapra. 2018. “ElimiNet: A Model for Eliminating Options for Reading Comprehension with Multiple Choice Questions.” In Proceedings of IJCAI.
- Haichao Zhou, Wei Furu, Qin Bing, and Liu Ting. 2018. “Hierarchical attention flow for multiple-choice reading comprehension.” In Proceedings of AAAI.
- Yichong Xu, Jingjing Liu, Jianfeng Gao, Yelong Shen, and Xiaodong Liu. 2017. “Towards human-level machine reading comprehension: Reasoning and inference with multiple strategies.” 2018. arXiv.