

Wizard of Wikipedia

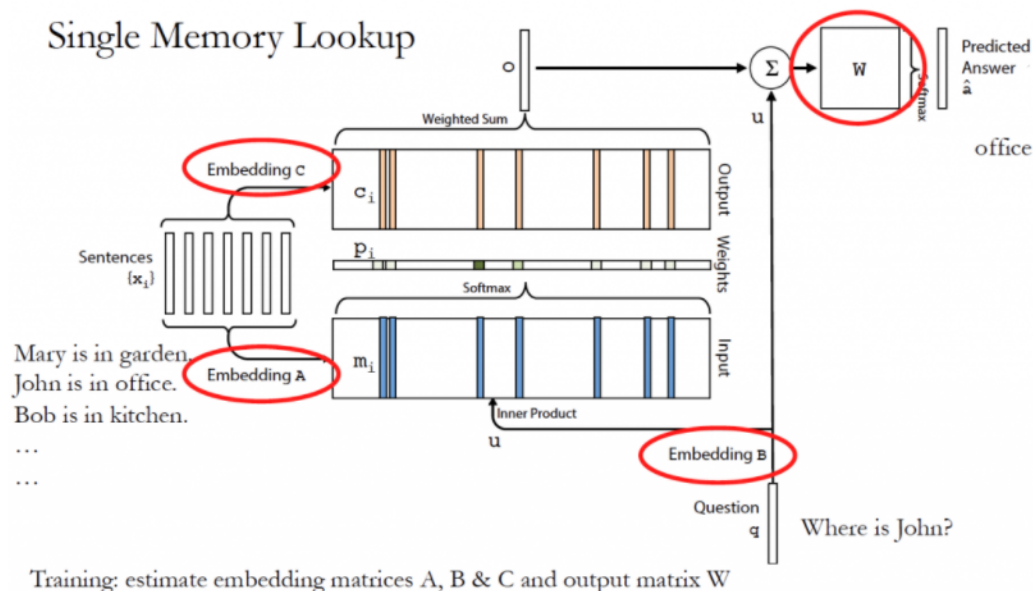
🕒 작성일시	@2021년 10월 5일 오후 2:45
🕒 최종 편집일시	@2021년 12월 23일 오전 7:56
☰ 문서 유형	간략 논문
📌 상태	완료
👤 작성자	
👤 최종 편집자	
☰ 학회	
☰ 키워드	

목표 : 데이터 셋 구성, 평가 요소, (option 모델)

참고 : 모델로는 End to End memory networks + Transformer 를 활용함

End to End memory networks

1. Single layer networks



- Inner memory - Key 느낌 (linear projection o)
- Outer memory - Value 느낌 (linear projection o)
- Question - Query 느낌 (linear projection o)

그래서 결국 Softmax(QKT)V 진행하는 것 과 유사함.

다만 메모리 형식으로 Sentences가 들어온다.

Sam walks into the kitchen.
Sam picks up an apple.
Sam walks into the bedroom.
Sam drops the apple.
Q: Where is the apple?
A. Bedroom

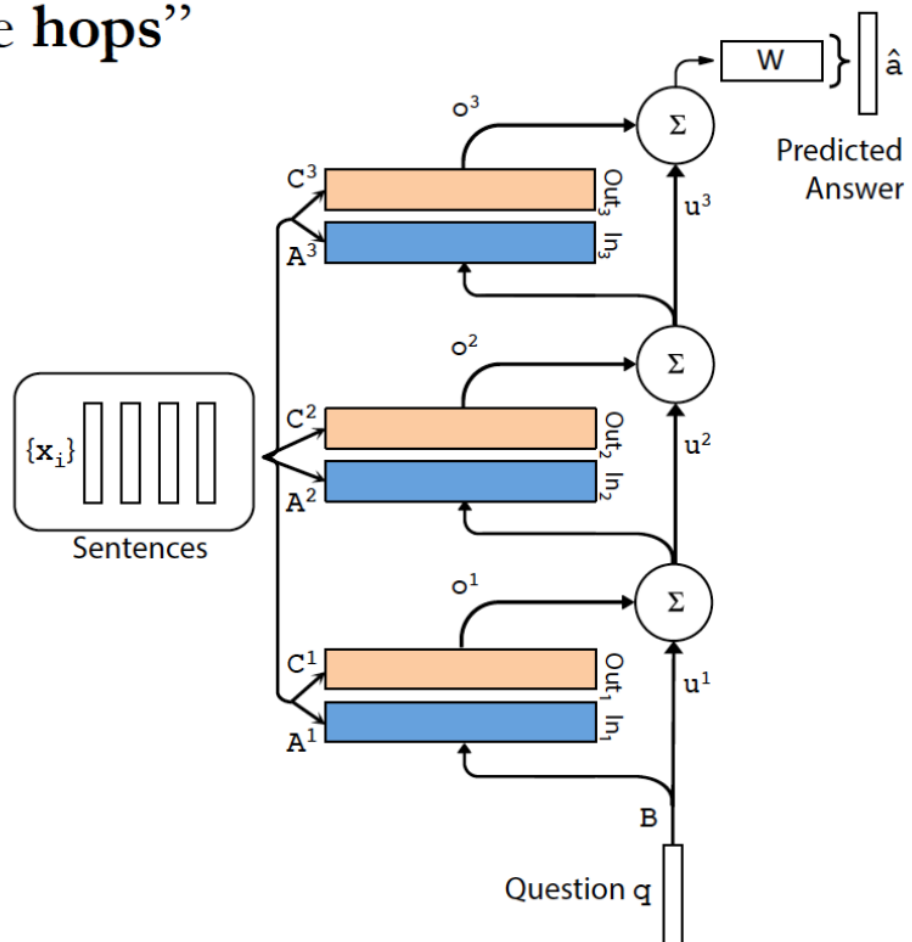
Brian is a lion.
Julius is a lion.
Julius is white.
Bernhard is green.
Q: What color is Brian?
A. White

Mary journeyed to the den.
Mary went back to the kitchen.
John journeyed to the bedroom.
Mary discarded the milk.
Q: Where was the milk before the den?
A. Hallway

→ 검은색 글씨들이 들어오고, 이것이 memory에 저장됨(이 때
로 구성

2. Multi layer networks

“Multiple hops”



여러 layer를 통과시키는 것임.

Wizard Of Wikipedia

Open Domain, chit chat, 대화 주제는 변경 가능, wizard(wikipedia 접근 가능), apprentice

Conversation flow

1. wizard 혹은 apprentice → topic pick
2. apprentice가 wizard에게 질문할 때, wizard는 관련 knowledge를 보고, 관련된 sentence를 선택해서 response를 생성한다.(선택 안 할 수도 있음)
3. 대화가 끝날 때까지 진행함(4~5 turn)

Topic

1. 1,365 → 각각이 Wikipedia와 연결됨.

Knowledge Retrieval

검색 방식은 fix

1. Inverted index lookup 을 활용함.
2. Articles와 query는 TF-IDF weighted bag of word와 n-gram vectors(hashing trick)활용.
 - 7가지 articles(첫문단만)~대화의 마지막 두 턴에 관련된 retrieve
 - 1가지 article (첫 10 문장만) ~ 대화 주제에 관련된 retrieve이를 제목과 함께해서, Wizard에게 제공한다.

Knowledge Selection and Response Generation

UI에서, Wizard가 만드려고 하는 response에 관련된 문서를 선택하고, 문장을 선택한다(간편하게 1개씩만 선택함) 또는 선택하지 않아도 됨.(no sentence used)

Table 1: Dataset statistics of the Wizard of Wikipedia task.

Wizard of Wikipedia Task	Train	Valid	Test Seen	Test Unseen
Number of Utterances	166,787	17,715	8,715	8,782
Number of Dialogues	18,430	1,948	965	968
Number of Topics	1,247	599	533	58
Average Turns per Dialogue	9.0	9.1	9.0	9.1
Knowledge Database	5.4M articles		93M sentences	

Final Dialogue Dataset The final dialogue dataset we collect consists of 22,311 dialogues with 201,999 turns, which we divide into 166,787 for train, 17,715 for validation, and 17,497 for test. The test set is split into two subsets, Test Seen and Test Unseen. Test Seen contains 533 overlapping topics with the training set with new dialogues about those topics. Test Unseen consists of 58 topics never seen before in train or validation. Overall data statistics can be found in Table 1, and further statistics and examples of collected conversations in Appendix A.2. We observe wizards and apprentices both asking and answering questions, and providing each other with a mixture of facts and personal feelings during their general discussion.

→ 실제 데이터셋을 확인해보니

A.2 WIZARD OF WIKIPEDIA EXAMPLES	
Topic:	Lifeguard
Apprentice:	So I am a lifeguard. Know anything about saving lives in water?
Wizard:	I'm impressed! It's a big responsibility to supervise other people's safety in the water! Tell me more.
Apprentice:	Well, I help make sure people do not drown or get injured while in or near the water!
Knowledge:	A lifeguard is a rescuer who supervises the safety and rescue of swimmers, surfers, ... Lifeguards are strong swimmers and trained in CPR/AED first aid, certified in water In some areas, the lifeguard service also carries out mountain rescues, or may function as the primary EMS provider.
Wizard:	I've heard that in some places, lifeguards also help with other sorts of emergencies, like mountain rescues! Is that part of your job too?
Apprentice:	I have! I feel like you know much about this! What brings you to know so much?
Wizard:	Oh, that's about the extent of my knowledge. I've just been around beaches and I've always admired lifeguards. I'm not a super strong swimmer myself.
Topic:	Armadillo
Wizard:	I love animals and think armadillos are awesome with their leathery shell.
Apprentice:	I don't think I've ever seen an armadillo in real life!
Wizard:	I've seen them at the zoo. Armadillo means little armored one in Spanish.
Apprentice:	Are they native to a Spanish-speaking part of the world?
Knowledge:	Armadillos are New World placental mammals in the order Cingulata ... The word "armadillo" means "little armoured one" in Spanish. ... The nine-banded armadillo ("Dasypus novemcinctus"), or the nine-banded, long-nosed armadillo, is a medium-sized mammal found in North, Central, and South America.
Wizard:	Yes, they are most commonly found in North, Central, and South America
Topic:	Ice cream
Wizard:	I just love ice cream. I love the types with fruits and flavours. Do you like ice cream?
Apprentice:	I love Ice cream as much as any one. I especially like Gelato, foreign ice cream!
Knowledge:	Ice cream is a sweetened frozen food typically eaten as a snack or dessert... It is usually made from dairy products, such as milk and cream, and Bacon ice cream (or bacon-and-egg ice cream) is an ice cream generally created by adding bacon to egg custard and freezing the mixture.
Wizard:	Me too. There are some strange combinations though, have you heard of bacon ice cream? where they add bacon and even egg custard to the freezing mixture!
Apprentice:	Surprisingly bacon ice cream doesn't surprise me. That doesn't sound appealing to me, but perhaps it could be delicious...

Figure 3: **The Wizard of Wikipedia dataset.** Examples of collected conversations from the dataset, where both wizard and apprentice are humans. The wizard has access to an information retrieval system over Wikipedia, so that they can ask and answer questions, and make statements relevant to the discussion. For each utterance, knowledge retrieval is performed based on dialogue history, giving ~61 knowledge candidates per turn, with wizards clicking *no sentence used* 6.2% of the time. Assuming that a question contains a question mark or begins with 'how', 'why', 'who', 'where', 'what' or 'when', in the dataset Apprentices ask questions in 13.9% of training set utterances, and answer questions (i.e., the Wizard has asked a question) 39.5% of the time, while saying new or follow-on statements (neither asking nor answering a question) 49.3% of the time. Hence, the wizard and apprentice conduct conversations with a variety of dialogue acts.

1. chosen topic : Science fiction
2. persona <- 이것을 바탕으로 topic을 결정함
3. dialog -> wizard/apprentice (순서는 관계없음)
'speaker', 'text', 'checked_sentence', 'checked_passage', 'retrieved_passages', 'retrieved_topics'

dict_keys(['chosen_topic', 'persona', 'wizard_eval', 'dialog', 'chosen_topic_passage'])

```
In [38]: data[0]['chosen_topic']
Out[38]: 'Science fiction'

In [39]: data[0]['persona']
Out[39]: 'i enjoy movies about aliens invading the earth.'

In [40]: data[0]['wizard_eval']
Out[40]: 5

In [41]: data[0]['dialog'][0].keys()
Out[41]: dict_keys(['speaker', 'text', 'checked_sentence', 'checked_passage', 'retrieved_passages', 'retrieved_topics'])
```

checked_sentence → 실제 활용한 sentence

checked_passage → 실제 활용한 passage

retrieved passages → 검색한 article + 첫 paragraph (7가지)

retrieved topics → 검색한 topic(wikipedia에선 제목이라고 생각하면 됨) (7가지)

```
In [38]: data[0]['dialog'][0]['retrieved_passages'][0]
Out[38]:
{'Hyperspace (science fiction)': ['Hyperspace is a faster-than-light (FTL) method of traveling used in science fiction.',
'It is typically described as an alternative "sub-region" of space co-existing with our own universe which may be entered using an energy field or other device.',
'As seen in most fiction hyperspace is most succinctly described as a "somewhere else" within which the laws of general and special relativity decidedly do "not" apply - especially with respect to the speed of light being the cosmic speed limit.',
'Entering and exiting said "elsewhere" thus directly enables travel near or faster than the speed of light - almost universally with the aid of extremely advanced technology.']}
```