

Question answering

Deep learning for Natural Processing, Deepmind

180528_홍수린

목차

* 괄호 안은 원본 자료에서의 페이지입니다

1. Intro (p.2)
2. Semantic parsing (p.12)
3. Answer sentence selection (p.32)
4. Reading comprehension (p.22)
5. Visual QA (p. 38)
6. Summary (p.45)

‘질문’과 ‘답변’의 예시

질문	답변
When were the first pyramids built?	2630 BC
Jean-Claude Juncker	Jean-Claude Juncker is a Luxembourgish politician. Since 2014, Juncker has been President of the European Commission.
How old is Keir Starmer?	54 years
What is the current price for AAPL?	136.50 USD
What's the weather like in London?	7 degrees Celsius. Clear with some clouds.
Whom did Juncker meet with?	The European Commission president was speaking after meeting with Irish Taoiseach Enda Kenny in Brussels.
When did you get to this lecture?	Five minutes after it started.

Visual Question Answering

Who is wearing glasses?

man



woman



Where is the child sitting?

fridge



arms



Is the umbrella upside down?

yes



no



How many children are in the bed?

2



1



질의응답에 왜 관심을 가져야 할까?

- ❖ 질의응답은 AI의 최후의 task로 인식됨: 질의응답을 완벽히 풀어낼 수 있다면 다른 모든 문제들도 해결할 수 있는 것
- ❖ 당장 직면한 많은 적용 분야가 존재함: 검색, 대화, 정보 추출, 요약...
- ❖ 이미 상당히 많은 사례가 존재: IBM Watson, Siri, Google Search...
- ❖ 하지만 아직도 매-우 갈 길이 멀다: 수많은 풀어나지 못한 문제들과 흥미로운 연구가 기다리고 있다!

‘답변’의 출처

질문	답변의 출처
최초의 피라미드는 언제 지어졌나?	백과사전
김연아 선수	백과사전 / 위키
문재인 대통령의 나이는?	아주아주 최근 백과사전
삼성전자 최근 주식 가격은 얼마지?	나스닥 주가 알림 시스템
서울 날씨 어때?	기상청
최근 문재인 대통령이 만난 사람은?	최근 뉴스기사, YTN뉴스 “2차 남북 정상회담 깜짝 방문....”
당신은 수업에 언제 도착하셨나요?	개인 진술

질의응답은 3가지 데이터에 의존

Question

Factual questions*

Complex/narrative
questions

Information Retrieval

Library Reference

Context/Source

Sets of documents
(corpus)

A single document

Knowledge Base

Non-linguistic types
of data (GPS, images,
sensors, ...)

Answer

A single fact

An explanation

A document

A sentence or
paragraph extracted
from somewhere

An image or other
type of object

Another question

질의응답 시스템을 제작하려면

- ❖ Question answering을 시작할 때 질문보다는 ‘응답’의 출처나 형태에 집중하는 것이 더 쉽다!
- ❖ QA 시스템을 제작할 때 고려할 세 가지 질문
 1. 응답의 형태가 어떠해야 하는가?
 2. 그 응답을 어디서부터 얻을 수 있는가?
 3. 나의 학습 데이터가 어떤 형태여야 하는가?

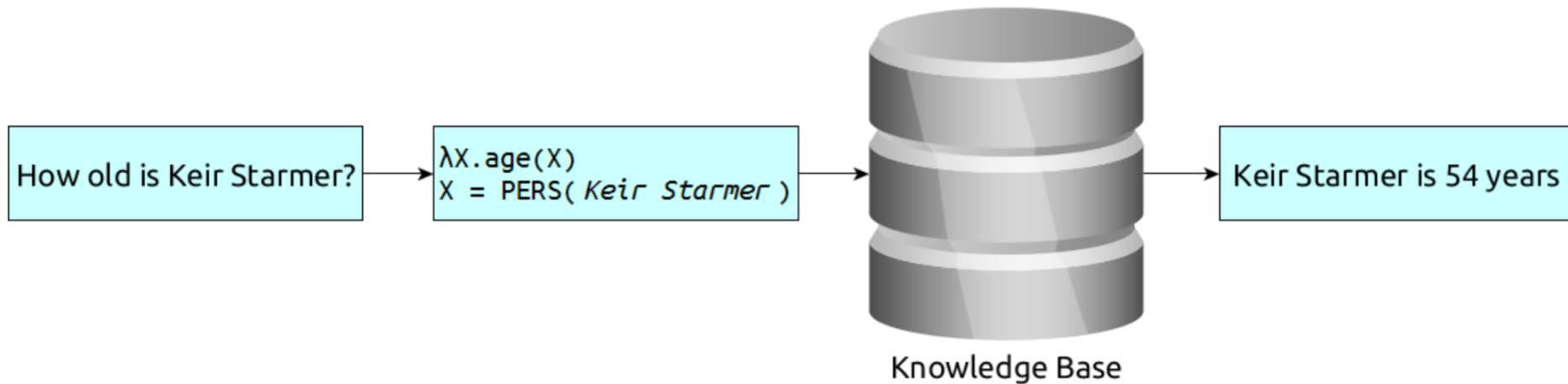
질의응답의 세부 분야

세부 분야	특징
Semantic parsing	<ul style="list-style-type: none">• 질문을 logical form으로 변환하는 데 집중• 'context'가 Knowledge base에 해당
Reading comprehension	<ul style="list-style-type: none">• 문서에 기반해 답변• 'context'는 특정한 문서에 해당
Sentence selection	<ul style="list-style-type: none">• 정답이 포함된 문장을 pick up해내는 task
Visual QA	<ul style="list-style-type: none">• 정답 형태가 단순하고 사실 기반• 'context'는 하나의 또는 복수의 이미지
Information retrieval	<ul style="list-style-type: none">• 정답은 문서, 단락, 문장 등• 'context'는 문서들의 집합
Library reference	<ul style="list-style-type: none">• 정답은 또다른 질문• 'context'는 도서관에서 또는 사서들이 접근 가능한 정형화된 지식

Semantic parsing

Semantic parsing

- 자연어를 정형화된 형태의 수식으로 구조화(mapping)하는 과정 ex.SQL
- Knowledge base에 query로 넣을 수 있도록 구조화
- ‘Semantic parsing’이란 질문을 Logical Form(LF)으로 바꾸는 과정이며, LF의 실제 정답을 찾는 것은 부수적인 과정으로 보는 경향 있음



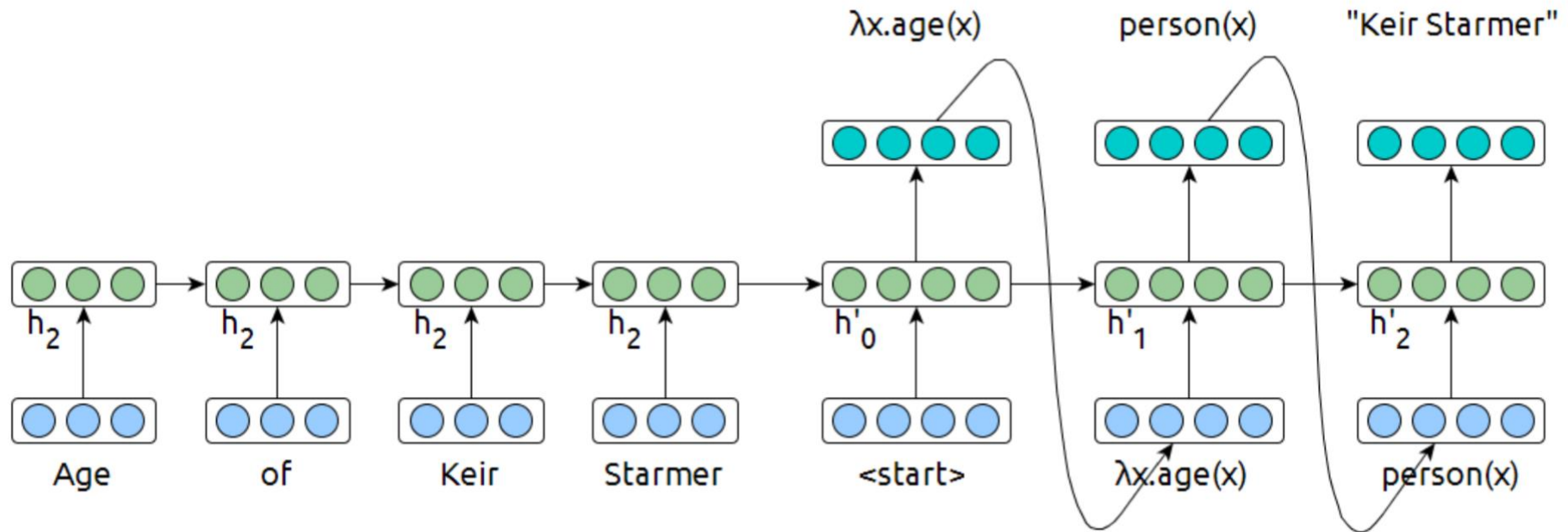
Question → Logical Form → KB Query → Answer

Semantic parsing의 KB

- ❖ 일반적으로 triple 형태로 구조화 됨: entity1, entity2, relation
 - *married-to, Michelle Obama, Barack Obama*
 - *member-of, United Kingdom, European Union*
- ❖ 무료 KB: Freebase, Wikidata, OpenStreetMap, GeoQuery,...
- ❖ 이러한 KB는 단순히 관계성만으로 이루어져 있기 때문에 제작이 용이하고 대량의 triple을 추출할 수 있음
- ❖ 하지만 supervised learning용 학습 데이터로 변환하기 위해서는 전문가의 annotation이 필요하므로 어려움

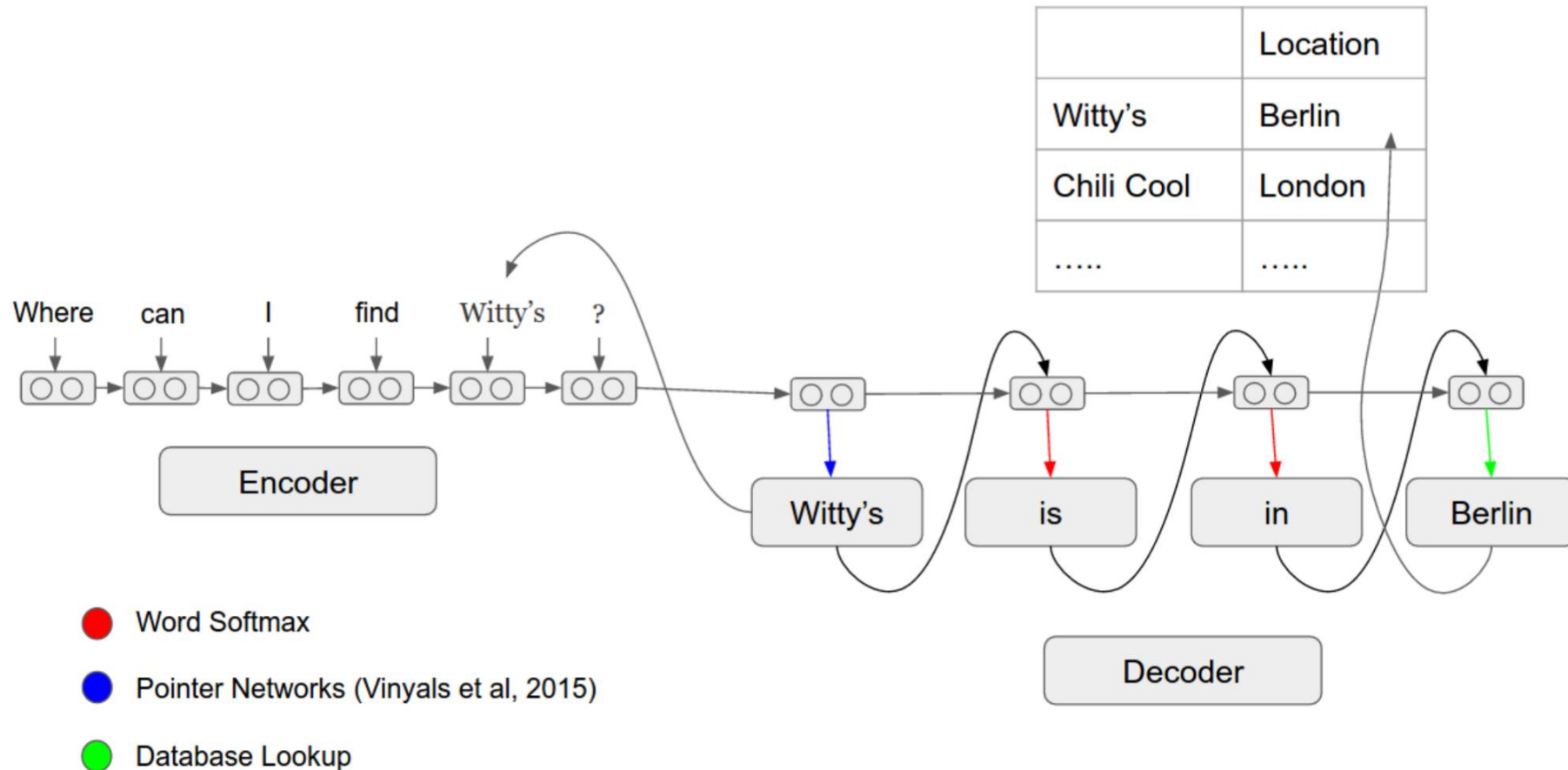
Deep learning approach

❖ 기계번역과 유사한 seq2seq model로 접근 가능



❖ 학습 데이터 확보가 어려움, target 언어가 LF이므로 성능이 좋지 않음

여러 source로부터의 생성



Recent studies

Ringgaard, Michael, Rahul Gupta, and Fernando CN Pereira. "SLING: A framework for frame semantic parsing." *arXiv preprint arXiv:1710.07032* (2017).

Goldman, O., Latacinnik, V., Naveh, U., Globerson, A., & Berant, J. (2017). Weakly-supervised Semantic Parsing with Abstract Examples. *arXiv preprint arXiv:1711.05240*.

Dong, Li, Chris Quirk, and Mirella Lapata. "Confidence Modeling for Neural Semantic Parsing." *arXiv preprint arXiv:1805.04604* (2018).

Answer sentence selection

Answer sentence selection

서울 인구 수



All

Images

News

Maps

Videos

More

Settings

Tools

About 17,500,000 results (0.52 seconds)

재외국민(1만472명)을 제외한 **서울** 내국인 **인구**는 이미 3월 기준으로 999만9116명까지 내려왔다. 통계청 관계자는 "이런 추세라면 곧 **서울** 주민등록 **인구** 1천만명 선이 붕괴될 것 같다"고 말했다. Apr 26, 2016

서울 '인구 1000만 시대', 28년 만에 무너진다 - 허핑턴포스트코리아

https://www.huffingtonpost.kr/2016/04/26/story_n_9777352.html

? About this result

Feedback

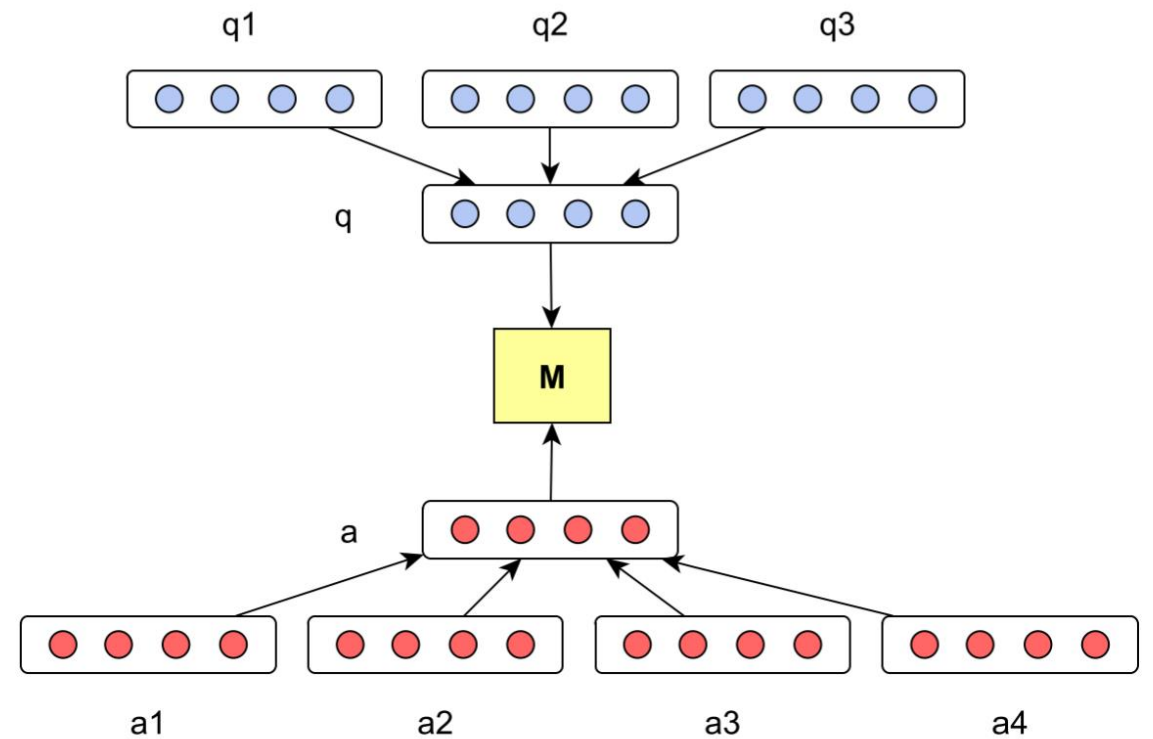
Answer sentence selection(cont.)

- ❖ Corpus에서 정답이 포함된 문장을 선별하는 것
- ❖ 사실 기반 질문(factual question)을 대상으로 함
- ❖ IR 시스템이나 웹 페이지를 source로 하여 주로 시도됨
- ❖ 결과물로는 하나 또는 복수 개의 정답 후보 문장 리턴됨
- ❖ RC의 정답이 부분 추출되거나 생성되는 데 반해 확실히 존재하는 문장에 통째로 추출됨
- ❖ Corpora: TREC QA track, MS MARCO

Neural model

- ❖ 정답 후보 a 와 질문 q 가 매칭될 확률을 계산해야 함
- ❖ 개별 후보들에 대한 값을 각기 따로 계산해준다는 점에서 이후 RC와 다름

$$p(y = 1|q, a) = \sigma(q^T M a + b)$$



Summary

- ❖ 정답 자체를 리턴하지 않고 해당하는 문장이나 context를 통째로 리턴한다는 점에서 다른 시스템들에 비해 robust하지만 동시에 한계 존재
- ❖ IR과 QA의 중간 파이프라인 역할(이후의 MRC에서 필수적)

Reading Comprehension

독해력(RC; Reading Comprehension)의 중요성과 어려움

일본에서 도쿄대 입시를 목표로 학습한 인공지능 로봇 'Torobo 군', 독해력이 향상되지 않아 언어 영역 점수 미달로 4수 실패

암기가 가능한 다른 지식들과 달리 언어는 '추론'을 통해 정답을 이끌어내는 과정이 필요함


ex.

Context: 아랍에미리트(UAE) 수도 아부다비 시청이 오랫동안 세차하지 않아 먼지를 뒤집어쓴 차의 주인 수백 명에게 과태료 3천디르함(약 91만원)씩을 부과했다고 현지 언론들이 25일(현지시간) 보도했다. 보도에 따르면 올해 4월1일부터 이달 22일까지 석 달여간 세차하지 않고 공용 주차장에 차를 주차해 과태료 처분을 받은 차주는 479명에 달했다. UAE에서 신호를 위반할 때 내야 하는 범칙금이 800디르함(약 25만원)인 점을 고려하면 꽤 높은 금액이다. 아부다비 시청은 심하게 더러운 차가 공용 주차장에 방치되면 도시의 미관이 나빠진다는 이유로 과태료를 부과했다. 시청은 세 번까지 경고장을 받은 뒤에도 세차하지 않은 차주에 대해 과태료 처분을 내리고 있다.

- Q1. 왜 아부다비에서는 더러운 차 주인에게 벌금을 내게 했어?
- Q2. 아랍에미리트에서 신호위반 범칙금 얼마야?
- Q3. 아부다비 시청은 세차하지 않은 차주에게 몇 번까지 경고해?

Boosting reading comprehension ability key to AI's future success

January 8, 2017 (Mainichi Japan)

 Japanese version



The AI system "Torobo-kun" uses robot arms to write out answers for a mock university entrance exam, in Tokyo's Chiyoda Ward on Nov. 14, 2016. (Mainichi)

Researchers on an artificial intelligence (AI) system called "Torobo-kun," which is being developed with an ultimate goal of passing the entrance exam for the University of Tokyo, Japan's top university, gave up on the dream for now after studies led by the National Institute of Informatics (NII) hit a snag.

While AI has been progressing rapidly, beating a former world Go champion and showing off other skills, the system is known to be poor in its reading comprehension ability that can lead to the right answer. A team of researchers from the NII and other entities have remodeled Torobo-kun by having it try out mock university entrance exams since 2013.

This fiscal year, Torobo-kun for the first time succeeded in earning mock exam scores that show

it has an 80 percent chance of passing the entrance exams for such prestigious private schools as Meiji University, Aoyama Gakuin University, Rikkyo University, Chuo

독해력(RC; Reading Comprehension)의 중요성과 어려움

일본에서 도쿄대 입시를 목표로 학습한 인공지능 로봇 'Torobo 군', 독해력이 향상되지 않아 언어 영역 점수 미달로 4수 실패

암기가 가능한 다른 지식들과 달리 언어는 '추론'을 통해 정답을 이끌어내는 과정이 필요함


ex.

Context: 아랍에미리트(UAE) 수도 아부다비 시청이 오랫동안 세차하지 않아 먼지를 뒤집어쓴 차의 주인 수백 명에게 과태료 3천디르함(약 91만원)씩을 부과했다고 현지 언론들이 25일(현지시간) 보도했다. 보도에 따르면 올해 4월1일부터 이달 22일까지 석 달여간 세차하지 않고 공용 주차장에 차를 주차해 과태료 처분을 받은 차주는 479명에 달했다. UAE에서 신호를 위반할 때 내야 하는 범칙금이 **800디르함(약 25만원)**인 점을 고려하면 꽤 높은 금액이다. 아부다비 시청은 **심하게 더러운 차가 공용 주차장에 방치되면 도시의 미관이 나빠진다는 이유**로 과태료를 부과했다. 시청은 **세 번**까지 경고장을 받은 뒤에도 세차하지 않은 차주에 대해 과태료 처분을 내리고 있다.

- Q1. 왜 아부다비에서는 더러운 차 주인에게 벌금을 내게 했어?
Q2. 아랍에미리트에서 신호위반 범칙금 얼마야?
Q3. 아부다비 시청은 세차하지 않은 차주에게 몇 번까지 경고해?

Boosting reading comprehension ability key to AI's future success

January 8, 2017 (Mainichi Japan)

 Japanese version



The AI system "Torobo-kun" uses robot arms to write out answers for a mock university entrance exam, in Tokyo's Chiyoda Ward on Nov. 14, 2016. (Mainichi)

Researchers on an artificial intelligence (AI) system called "Torobo-kun," which is being developed with an ultimate goal of passing the entrance exam for the University of Tokyo, Japan's top university, gave up on the dream for now after studies led by the National Institute of Informatics (NII) hit a snag.

While AI has been progressing rapidly, beating a former world Go champion and showing off other skills, the system is known to be poor in its reading comprehension ability that can lead to the right answer. A team of researchers from the NII and other entities have remodeled Torobo-kun by having it try out mock university entrance exams since 2013.

This fiscal year, Torobo-kun for the first time succeeded in earning mock exam scores that show

it has an 80 percent chance of passing the entrance exams for such prestigious private schools as Meiji University, Aoyama Gakuin University, Rikkyo University, Chuo

쉬운 문서, 어려운 문서

어려운 질문(?)	쉬운 질문(?)
<p>컴퓨터(영어: computer, 문화어: 콤퓨터) 또는 셈틀은 수식이나 논리적 언어로 표현된 계산을 수행하거나 작업을 통제하는 기계다. 사전에서 찾을 수 있는 위와 같은 정의는 정확하지만 가리키는 범위가 너무 넓어 이전의 컴퓨터와 현재의 컴퓨터 그리고 미래의 컴퓨터를 정확하게 설명하기는 어렵다. 더 의미 있는 질문은 '컴퓨터의 종류에는 어떤 것들이 있는가?' 아니면 '현재의 컴퓨터의 능력과 기능의 특별한 것은 무엇인가?'일 것이다. 파스칼은 세무사였던 아버지의 작업을 돕기 위해 덧셈과 뺄셈을 자동으로 수행하는 기계식 계산기를 고안했다. 1950년대 초에 '컴퓨팅 머신'(Computing Machine)이라는 말이 전산기를 지시하기 위해서 쓰였다. 마침내, 더 짧은 컴퓨터(computer)라는 말이 컴퓨팅 머신을 대체했다. 본래, 산술은 수학적인 문제와 밀접하게 연관되어 있지만, 현대 컴퓨터들은 값이 싸지고 용도가 다양해짐에 따라, 수학과는 관계 없는 많은 일에도 쓰인다.</p>	<p>영철이는 옆 반 미나를 좋아하게 되었다. 미나는 공부도 잘 하고 인기가 많았다. 영철이는 미나가 운동을 잘 하는 사람을 좋아한다는 사실을 알게 되었다. 영철이는 열심히 농구 연습을 하기 시작했다.</p>
Q. 기계식 계산기를 발명한 사람은?	Q. 영철이가 농구 연습을 하게 된 이유는?
A. 파스칼	A. 좋아하는 사람이 생겨서

첫 번째 '어려운 질문'은 위키피디아에서 본문을 가져온 것으로, 문서 자체는 어른이 작성하였고 복잡해 보이지만 본문 내에서 답을 그대로 찾아낼 수 있음 - 기계가 잘 답할 수 있는 유형

두 번째 '쉬운 질문'은 어린이들의 문체이므로 본문이 단순하고 단어가 쉽지만 질문에 답하기 위해서는 여러 문장에 걸친 추론이 필요 - 기계가 어려워하는 유형

Reading comprehension

- ❖ 주어진 문서 내에서 질문에 대한 답을 찾아 줌
- ❖ 정답이 본문에 단어나 단어 구 형태로 포함되어 있어야 함
- ❖ CNN/DailyMail, CBT, SQuAD, ...

<https://rajpurkar.github.io/SQuAD-explorer/>

The screenshot shows the SQuAD website. The header is red with 'SQuAD' in white and 'The Stanford Question Answering Dataset' below it. The main content area is divided into two columns. The left column, titled 'What is SQuAD?', describes the dataset and includes buttons for 'Explore SQuAD and model predictions' and 'Read the paper (Rajpurkar et al. '16)'. The right column, titled 'Leaderboard', shows a table of the best models on the test set of v1.1.

Rank	Model	EM	F1
	Human Performance Stanford University (Rajpurkar et al. '16)	82.304	91.221
1	QANet (ensemble) Google Brain & CMU Mar 19, 2018	83.877	89.737
2	MARS (ensemble) YUANFUDAO research NLP May 10, 2018	83.520	89.612
3	QANet (ensemble) Google Brain & CMU Mar 06, 2018	82.744	89.045
4	MARS (single model) YUANFUDAO research NLP May 09, 2018	82.587	88.880

CNN/Daily mail dataset example

CNN article

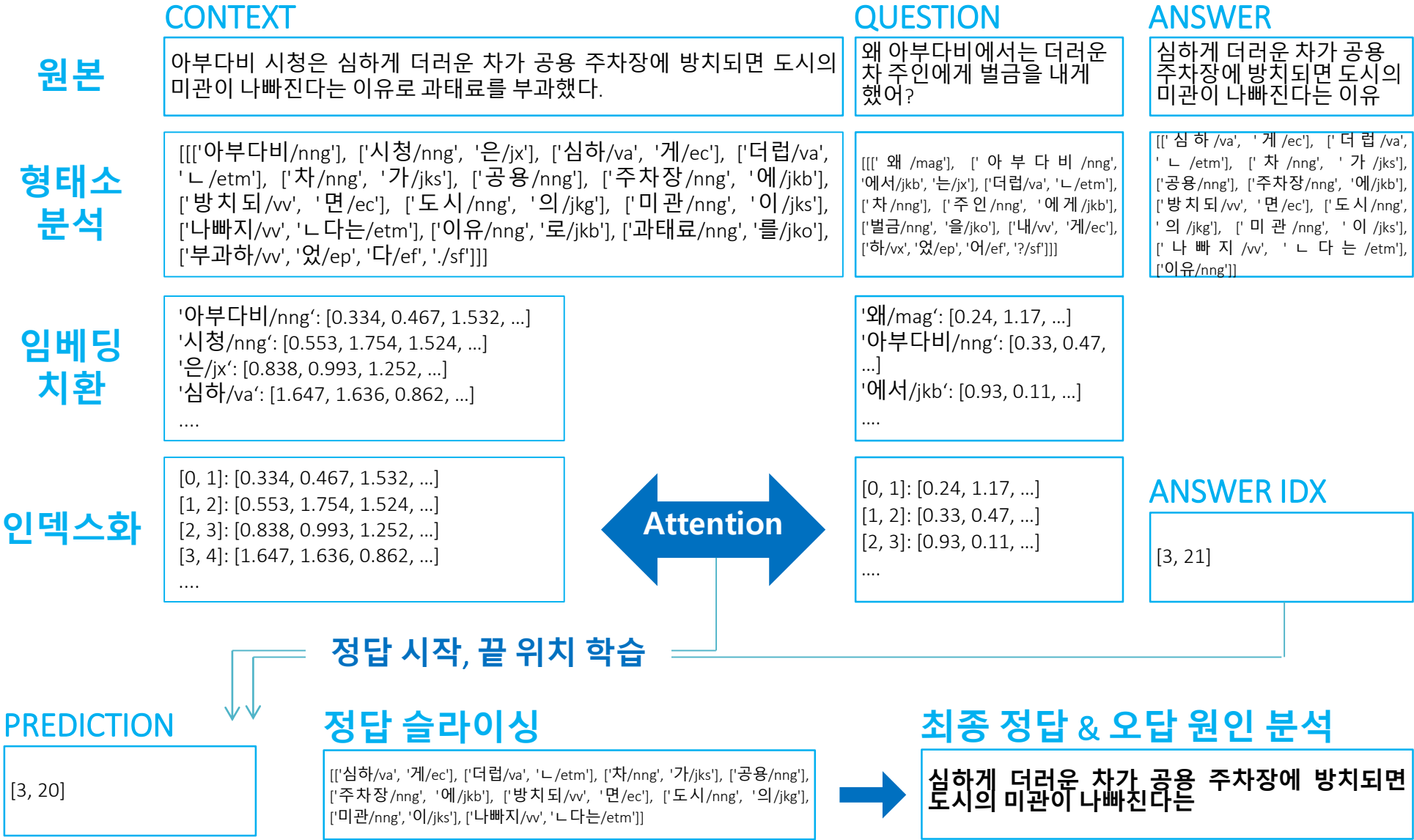
Document The BBC producer allegedly struck by Jeremy Clarkson will not press charges against the “Top Gear” host, his lawyer said Friday. Clarkson, who hosted one of the most-watched television shows in the world, was dropped by the BBC Wednesday after an internal investigation by the British broadcaster found he had subjected producer Oisin Tymon “to an unprovoked physical and verbal attack.” ...

Query Producer **X** will not press charges against Jeremy Clarkson, his lawyer says.

Answer Oisin Tymon

We formulate *Cloze* style queries from the story paraphrases.

학습 과정에서의 데이터 형태 변화



응용 알고리즘: Reasonet(2016-09)

Reasonet: Reasoning Network (Microsoft 발표-소스코드는 없음)

일정한 횟수만 document 내에서 반복적으로 추론을 하던 기존 기법들과 달리 terminate state를 도입하여 자동으로 추론을 중지할 시점을 찾음

Terminate state를 찾는 과정에서는 강화학습 사용

Table 2: The performance of Reasoning Network on CNN and Daily Mail dataset.

	CNN		Daily Mail	
	valid	test	valid	test
Deep LSTM Reader [6]	55.0	57.0	63.3	62.2
Attentive Reader [6]	61.6	63.0	70.5	69.0
MemNets [7]	63.4	66.8	-	-
AS Reader [8]	68.6	69.5	75.0	73.9
Stanford AR [3]	72.2	72.4	76.9	75.8
DER Network [11]	71.3	72.9	-	-
Iterative Attention Reader [18]	72.6	73.3	-	-
EpiReader [22]	73.4	74.0	-	-
GA Reader [5]	73.0	73.8	76.7	75.7
AoA Reader [4]	73.1	74.4	-	-
Reasonet	72.9	74.7	77.6	76.6

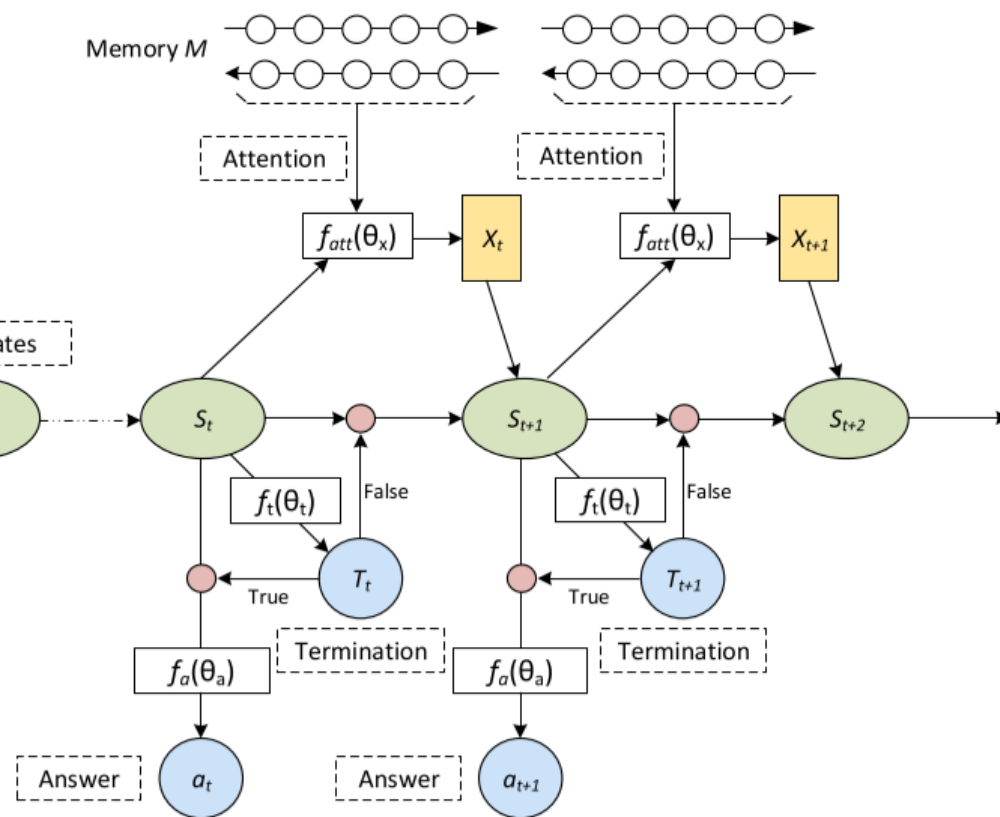


Figure 1: Reasonet Architecture.

응용 알고리즘: Bi-DAF(2016-11)

정답을 실제로 찾아내기 위해 character-level word embedding, 의미적 전달을 위해 word-level word embedding 사용

Context2Query, Query2Context가 발생하는 Attention Flow Layer가 핵심

모든 소스코드가 github에 공개되어 있고 매뉴얼도 잘 설명되어 있으므로 입문용으로 학습해보기 좋음

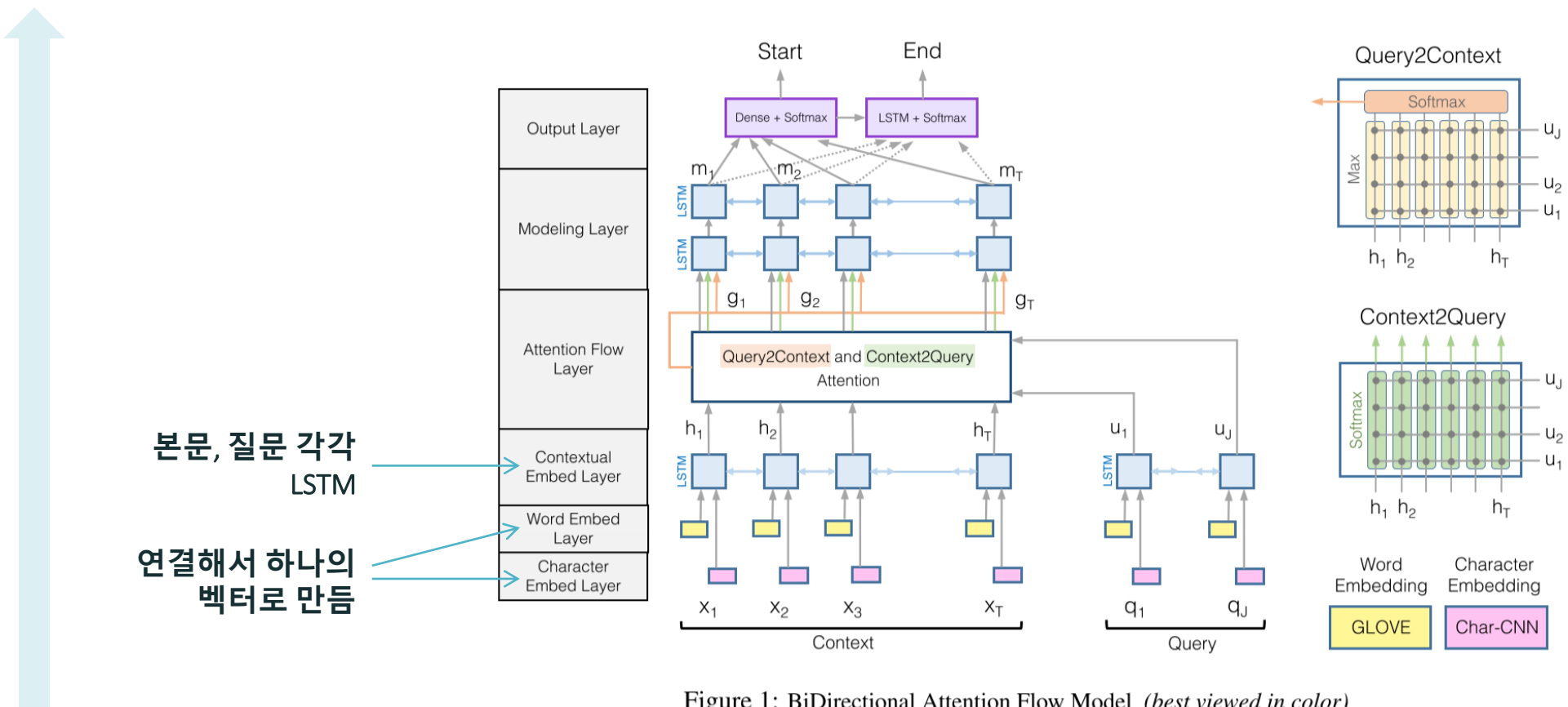


Figure 1: BiDirectional Attention Flow Model (best viewed in color)

응용 알고리즘: Bi-DAF cont.

	Single Model		Ensemble	
	EM	F1	EM	F1
Logistic Regression Baseline ^a	40.4	51.0	-	-
Dynamic Chunk Reader ^b	62.5	71.0	-	-
Fine-Grained Gating ^c	62.5	73.3	-	-
Match-LSTM ^d	64.7	73.7	67.9	77.0
Multi-Perspective Matching ^e	65.5	75.1	68.2	77.2
Dynamic Coattention Networks ^f	66.2	75.9	71.6	80.4
R-Net ^g	68.4	77.5	72.1	79.7
BiDAF (Ours)	68.0	77.3	73.3	81.1

(a) Results on the SQuAD test set

	EM	F1
No char embedding	65.0	75.4
No word embedding	55.5	66.8
No C2Q attention	57.2	67.7
No Q2C attention	63.6	73.7
Dynamic attention	63.5	73.6
BiDAF (single)	67.7	77.3
BiDAF (ensemble)	72.6	80.7

(b) Ablations on the SQuAD dev set

Bi-directional Attention Flow Demo

for [Stanford Question Answering Dataset \(SQuAD\)](#)

Direction : Select a paragraph and write your own question. The answer is always a subphrase of the paragraph - remember it when you ask a question!

Select Paragraph

[01] Warsaw ▼

Paragraph

One of the most famous people born in Warsaw was Maria Skłodowska-Curie, who achieved international recognition for her research on radioactivity and was the first female recipient of the Nobel Prize. Famous musicians include Władysław Szpilman and Frédéric Chopin. Though Chopin was born in the village of Żelazowa Wola, about 60 km (37 mi) from Warsaw, he moved to the city with his family when he was seven months old. Casimir Pulaski, a Polish general and hero of the American Revolutionary War, was born here in 1745.

Question

What was Maria Curie the first female recipient of?

new question!

Answer

Nobel Prize

응용 알고리즘: Match-LSTM and Answer pointer (2016-08)

마지막 answer pointer layer에서
(a) sequence model은 정답
시퀀스 자체를 예측, (b) boundary
model은 시작과 끝 위치만 예측

Boundary model 의 성능이 더
높음

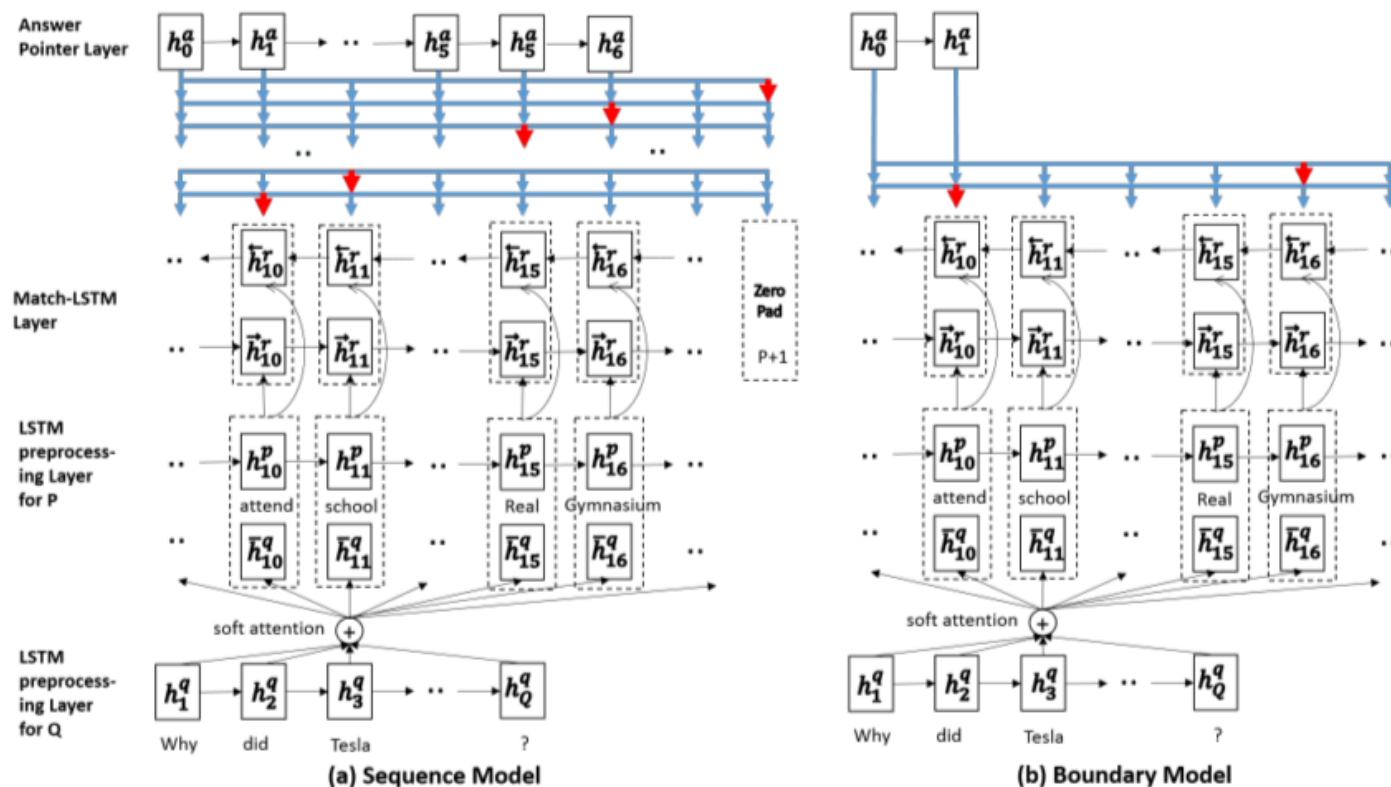
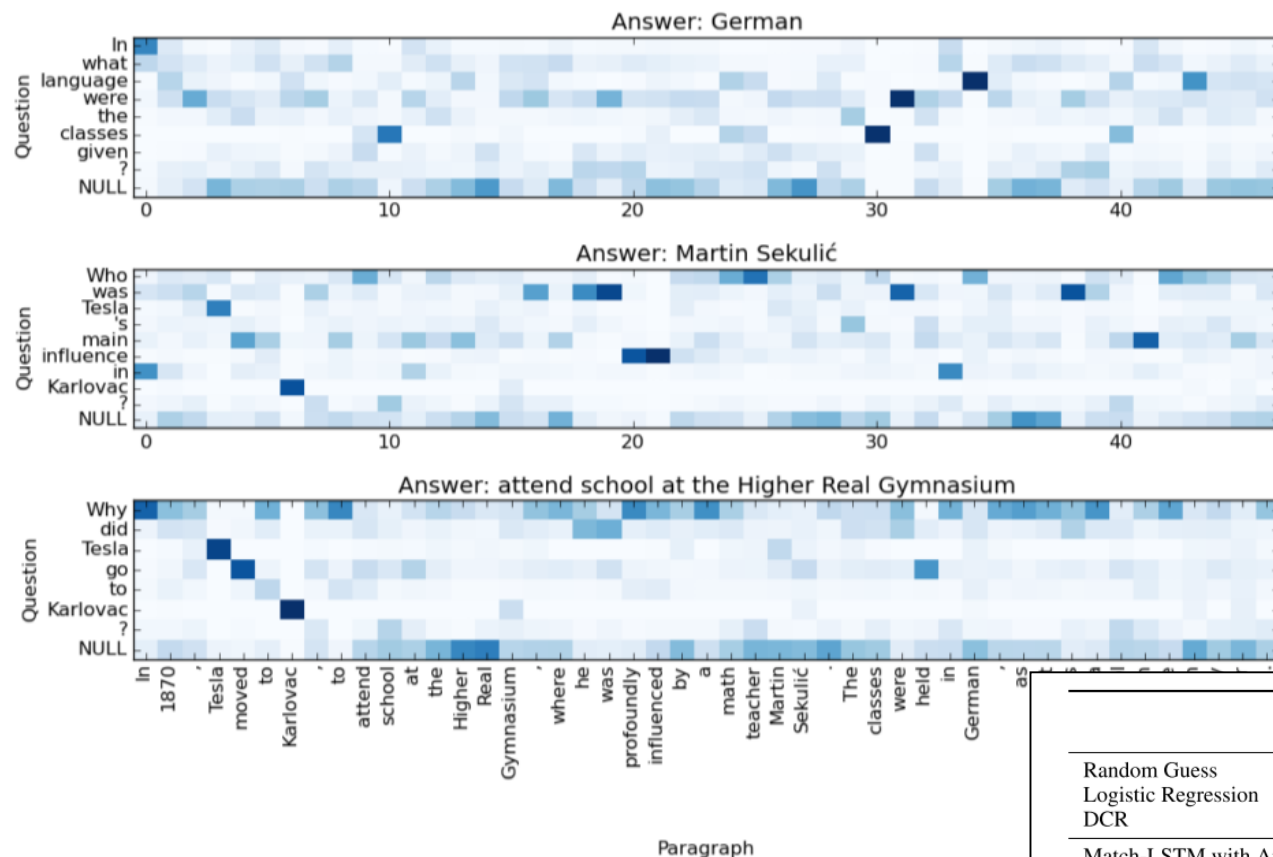


Figure 1: An overview of our two models. Both models consist of an LSTM preprocessing layer, a match-LSTM layer and an Answer Pointer layer. For each match-LSTM in a particular direction, \vec{h}_i^q , which is defined as $\mathbf{H}^q \alpha_i^T$, is computed using the α in the corresponding direction, as described in either Eqn. (2) or Eqn. (5).

응용 알고리즘: Match-LSTM and Answer pointer cont.



질문과 본문의 단어들이 attention되는 모습을 시각화한 것

진한 부분이 질문의 단어에 따라 본문에서 연관성이 높은 부분

Figure 2: Visualization of the attention weights α for three questions associated with the passage.

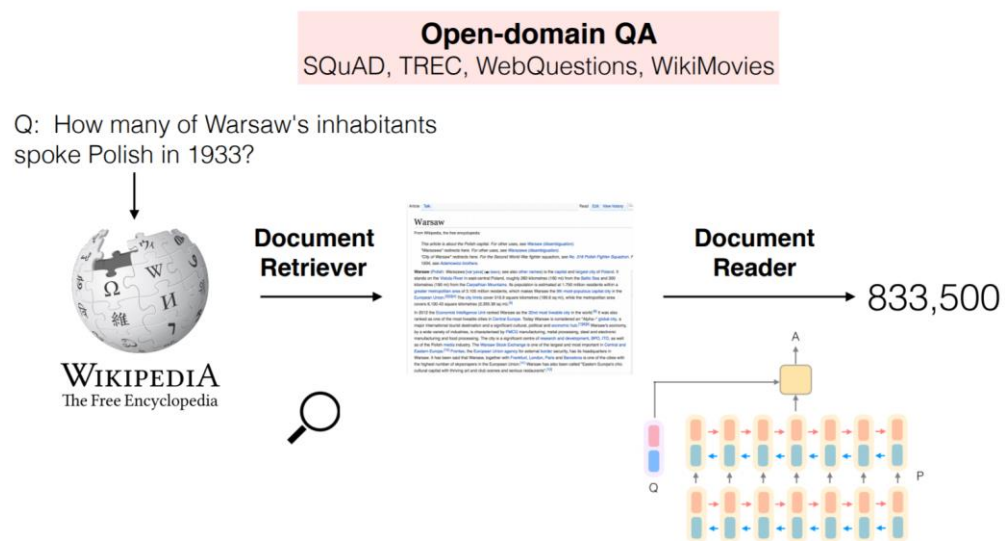
	l	$ \theta $	Exact Match		F1	
			Dev	Test	Dev	Test
Random Guess	-	0	1.1	1.3	4.1	4.3
Logistic Regression	-	-	40.0	40.4	51.0	51.0
DCR	-	-	62.5	62.5	71.2	71.0
Match-LSTM with Ans-Ptr (Sequence)	150	882K	54.4	-	68.2	-
Match-LSTM with Ans-Ptr (Boundary)	150	882K	61.1	-	71.2	-
Match-LSTM with Ans-Ptr (Boundary+Search)	150	882K	63.0	-	72.7	-
Match-LSTM with Ans-Ptr (Boundary+Search)	300	3.2M	63.1	-	72.7	-
Match-LSTM with Ans-Ptr (Boundary+Search+b)	150	1.1M	63.4	-	73.0	-
Match-LSTM with Bi-Ans-Ptr (Boundary+Search+b)	150	1.4M	64.1	64.7	73.9	73.7
Match-LSTM with Ans-Ptr (Boundary+Search+en)	150	882K	67.6	67.9	76.8	77.0

Table 2: Experiment Results. Here “Search” refers to globally searching the spans with no more than 15 tokens, “b” refers to using bi-directional pre-processing LSTM, and “en” refers to ensemble method.

응용 알고리즘: DrQA (2017-04)

machine reading at scale(MRS): 수백만의 document source에서 정답이 존재할 만한 유사한 문서를 찾고 그 안에서 정답을 추출해야 함

위키피디아만을 document source로 사용하여 외부 지식베이스에 의존성을 줄이고 경량화 함



Method	Dev		Test	
	EM	F1	EM	F1
Dynamic Coattention Networks (Xiong et al., 2016)	65.4	75.6	66.2	75.9
Multi-Perspective Matching (Wang et al., 2016) [†]	66.1	75.8	65.5	75.1
BiDAF (Seo et al., 2016)	67.7	77.3	68.0	77.3
R-net [†]	n/a	n/a	71.3	79.7
DrQA (Our model, Document Reader Only)	69.5	78.8	70.0	79.0

Figure 1: An overview of our question answering system DrQA.

학습 데이터 예시

```
{
  "paragraphs": [
    {
      "context": "[['KIA/sl'], ['타이거즈/nnp'], ['제공/nng', '한승혁/nnp', '(/ss', '24/sn', '·/nnb', 'KIA/sl', ')/ss', '은/jx'], ['지', '스프링/nng', '캠프/nng', '에서/jkb'], ['"/nng', '올해/nng', '는/jx'], ['절대로/mag'], ['오버페이스/nng'], ['하/vv', '지/ec'], ['않/vx', '·/sf'], ['거의/mag'], ['매년/mag'], ['스프링/nng', '캠프/nng', '에서/jkb'], ['힐/nng', '을/jko'], ['쏟/vv', '다가/ec'], ['개막/nng'], '을/jko'], ['당하/vv', '아/ec'], ['시즌/nng', '을/jko'], ['절대로/mag'], ['뛰/vv', '≒/etm'], ['수/nnb'], ['없/va', '있/ep', '면/etr', '올해/nng', '는/jx'], ['반복/nng', '하/xsv', '지/ec'], ['않/vx', '기/etn'], ['위하/vv', '아서/ec', '이/vcp', '있/ep', '다/ef', '·/s', '처럼/jkb'], ['100/sn', '₩/sw'], ['모드/nng', '로/jkb'], ['가/vv', '려/ec'], ['하/vx', '≒/etm'], ['때/nng', '마다/jx'], ['머릿속/r', '떠올리/vv', '있/ep', '다/ef', '·/sf'], ['그/mm'], ['결과/nng'], ['투수/nng', '와/jc'], ['타자/nng', '를/jko'], ['통틀/vv', '어/ec', '받/vv', '으며/ec'], ['김기태/nnp'], ['감독/nng', '이/jks'], ['직접/mag'], ['지정/nng', '하/xsv', 'ㄴ/etm'], ['"/ss', '스프링/nng', '있/ep', '다/ef', '·/sf'], ['시범/nng', '경기/nng', '를/jko'], ['치르/vv', '고/ec'], ['있/vx', '는/etm'], ['지금/mag', '·/sp'], ['마다/jx'], ['화제/nng', '예/jkb'], ['오르/vv', 'ㄴ다/ef', '·/sf'], ['2/sn', '월/nnb', '예/jkb'], ['연습/nng', '경기/nng', '부터/jx', '면/etm'], ['한승혁/nnp', '의/jkg'], ['구속/nng', '은/jx'], ['개막/nng', '을/jko'], ['논란/nng', '예/jkb'], ['두/vv', 'ㄴ/etm'], ['가', '다/ef', '·/sf'], ['시속/nng'], ['157/sn', 'km/nnb', '는/jx'], ['전/nng', '예/jkb', '도/jx'], ['한승혁/nnp', '의/jkg'], ['등/nng', '찍히/vv', 'ㄴ/etm'], ['구속/nng', '이/vcp', '다/ef', '·/sf'], ['그러나/maj'], ['지금/nng', '의/jkg'], ['157/sn', 'km/nnb', '는/j', '하/xsv', '고/ec'], ['있/vx', '다/ef', '·/sf'], ['갈/va', '은/etm'], ['구속/nng', '예/jkb', '도/jx'], ['모두/nng', '가/jks'], ['설', 'qas": [
      {
        "id": "c8_sports_25-1",
        "question": "[['김기태/nnp'], ['감독/nng', '이/jks'], ['올해/nng'], ['스프링/nng', '캠프/nng'], ['MVP/sl', '로/jkb'], ['지정/nng', '아/ef', '?/sf']]",
        "answers": [
          {
            "text": "[['한승혁/nnp']]",
            "answer_end": 4,
            "answer_start": 3
          }
        ]
      },
      {
        "id": "c8_sports_25-2",
        "question": "[['김기태/nnp'], ['감독/nng', '은/jx'], ['올해/nng'], ['스프링/nng', '캠프/nng'], ['MVP/sl', '로/jkb'], ['누구/np', '의/jkg'], ['한승혁/nnp', '의/jkg'], ['구속/nng', '은/jx'], ['개막/nng', '을/jko'], ['논란/nng', '예/jkb'], ['두/vv', 'ㄴ/etm'], ['가', '다/ef', '·/sf'], ['시속/nng'], ['157/sn', 'km/nnb', '는/jx'], ['전/nng', '예/jkb', '도/jx'], ['한승혁/nnp', '의/jkg'], ['등/nng', '찍히/vv', 'ㄴ/etm'], ['구속/nng', '이/vcp', '다/ef', '·/sf'], ['그러나/maj'], ['지금/nng', '의/jkg'], ['157/sn', 'km/nnb', '는/j', '하/xsv', '고/ec'], ['있/vx', '다/ef', '·/sf'], ['갈/va', '은/etm'], ['구속/nng', '예/jkb', '도/jx'], ['모두/nng', '가/jks'], ['설', '설']]"
      }
    ]
  ]
}
```

본문

질문

정답

정답 인덱스 (시작, 끝)

학습 주기 및 유지보수 방법

- ❖ 새로운 데이터가 얼마나 자주 축적되는지에 따라 다름
- ❖ 모델이 학습 데이터에 없는 질문도 찾을 수 있기 때문에 매일 학습할 필요는 없음
- ❖ 하지만 아래와 같은 사항을 반영하기 위해 일정 기간마다 재학습이 권장됨
- ❖ 동일한 유형의 문서라도(예. 뉴스) 시간이 지나면 문체나 형식이 약간씩 바뀌므로 이를 반영하기 위해 지속적인 학습이 필요
- ❖ 학습 데이터에 전혀 없는 새로운 유형의 질문이 발생할 수도 있음
- ❖ 사람이 직접 학습 데이터를 제작해야 하므로 아무것도 없는 상태에서 질문-답을 제작하기보다는 운영 로그를 참고하여 대답을 잘 하지 못했던 질문 위주로 추가하는 것이 권장됨

MRC의 취약점

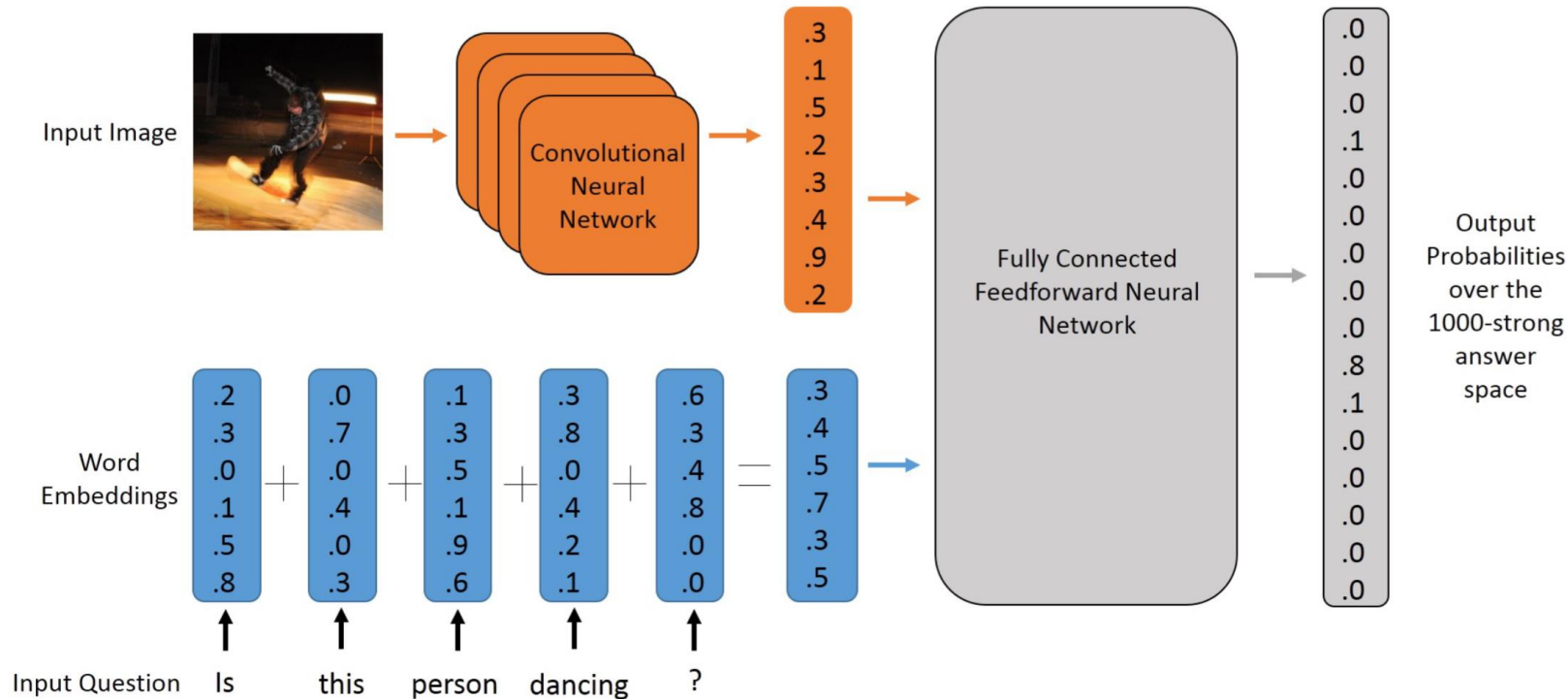
- ❖ *본문이 존재한 상태에서만 정답을 찾을 수 있음 == 실제 시스템의 체감 성능은 본문 검색 성능*RC의 성능*
- ❖ 데이터를 제작하는 데 많은 리소스가 투입됨. 자동으로 구축할 경우 양질의 데이터셋이 제작되지 못함
- ❖ 문종이나 본문의 형태가 변경될 경우 그에 맞춰 일정 부분 또 데이터를 구축해야 함. 음성인식의 적응학습 같은 개념으로 생각

Visual QA

General visual QA

- Question is language \rightarrow some encoder
- Context is a single picture \rightarrow convolutional network
- Answer is a single word \rightarrow classifier function

We have covered all the components already:



Entire Summary: 7 questions

1. 목표-task-가 무엇인가?
2. 질문, 맥락, 답변의 형태가 어떠한가?
3. 데이터는 어디에서부터 오는가?
4. 데이터를 증가시킬 수 있는가?
5. 질문과 맥락(context)을 어떻게 encode해야 하는가?
6. 질문과 맥락(context)을 어떻게 융합시켜야 하는가?
7. 답변을 어떻게 예측 또는 생성할 것인가?

답변에 더 집중하라고 하지만...(질문도)

이순신 장군 언제 돌아가셨어?

이순신 장군 사망일은 언제야?

이순신 장군 사망일은?

이순신 장군 사망일

이순신 사망일

이순신 죽은 날

Thank you