



# 基于深度学习的阅读理解

刘 康

中国科学院自动化研究所  
模式识别国家重点实验室  
2016年10月14日

# 任务描述

James the Turtle was always getting in trouble. Sometimes he'd reach into the freezer and empty out all the food. Other times he'd sled on the deck and get a splinter. His aunt Jane tried as hard as she could to keep him out of trouble, but he was sneaky and got into lots of trouble behind her back.

One day, James thought he would go into town and see what kind of trouble he could get into. He went to the grocery store and pulled all the pudding off the shelves and ate two jars. Then he walked to the fast food restaurant and ordered 15 bags of fries. He didn't pay, and instead headed home.

His aunt was waiting for him in his room. She told James that she loved him, but he would have to start acting like a well-behaved turtle. After about a month, and after getting into lots of trouble, James finally made up his mind to be a better turtle.

1) What is the name of the trouble making turtle?

- A) Fries
- B) Pudding
- C) James
- D) Jane

2) What did James pull off of the shelves in the grocery store?

- A) pudding
- B) fries
- C) food
- D) splinters

3) Where did James go after he went to the grocery store?

- A) his deck
- B) his freezer
- C) a fast food restaurant
- D) his room

# 挑战

- 不完全依赖于外部知识

- 推理

- 指代消解

- 需要常识知识

James the Turtle was always getting in trouble. Sometimes he'd reach into the freezer and empty out all the food. Other times he'd sled on the deck and get a splinter. His aunt Jane tried as hard as she could to keep him out of trouble, but he was sneaky and got into lots of trouble behind her back. One day, James thought he would go into town and see what kind of trouble he could get into. He went to the grocery store and pulled all the pudding off the shelves and ate two jars. Then he walked to the fast food restaurant and ordered 15 bags of fries. He didn't pay, and instead headed home. His aunt was waiting for him in his room. She told James that she loved him, but he would have to start acting like a well-behaved turtle. After about a month, and after getting into lots of trouble, James finally made up his mind to be a better turtle.

1) What is the name of the trouble making turtle?

- A) Fries
- B) Pudding
- C) James
- D) Jane

2) What did James pull off of the shelves in the grocery store?

- A) pudding
- B) fries
- C) food
- D) splinters

3) Where did James go after he went to the grocery store?

- A) his deck
- B) his freezer
- C) a fast food restaurant
- D) his room

# 数据集

时间	数据集	描述	题型	大小
2015	AI2 Science Exams (UW)	美国科学小学考试（真实数据）	选择题	很小（百级）
2015	bAbi (Facebook)	考察20个方面文本理解和推理能力，简单的文本表达（模拟数据）	问答	很小（1000个训练、1000个测试、每个任务150个词）
2013	MCTest (Microsoft)	儿童读物（真实数据，众包）	选择题	很小（MC160：160篇；MC500：500篇）
2015	CNN/Daily Mail (Deep Mind)	新闻文本（真实数据，自动标注）	完形填空	较大（9w篇）
2016	CBTest (Facebook)	儿童读物（真实数据，自动标注）	完型填空	大（68w）
2016	Chinese RCD (iFlytek and HIT)	新闻文本儿童读物（真实数据，自动标注）	完形填空	大（87w）
2016	SQuAD (Stanford)	Wikipedia文本（真实数据，众包）	问答	较大（10w篇）
2016	Rochester Story	Short Commonsense Story（真实数据，众包）	选择题（句子级）	中等（1w篇）

# bAbi: Facebook AI Research

- 1,000 Training, 1,000 Testing
- Answer: single words or a list of words
- 20 tasks in English
- Simulated world: actions
- 150 words, 4 actors, 6 locations and 3 objects per task

**Task 1: Single Supporting Fact**

Mary went to the bathroom.  
John moved to the hallway.  
Mary travelled to the office.  
Where is Mary? **A:office**

**Task 2: Two Supporting Facts**

John is in the playground.  
John picked up the football.  
Bob went to the kitchen.  
Where is the football? **A:playground**

**Task 3: Three Supporting Facts**

John picked up the apple.  
John went to the office.  
John went to the kitchen.  
John dropped the apple.  
Where was the apple before the kitchen? **A:office**

**Task 4: Two Argument Relations**

The office is north of the bedroom.  
The bedroom is north of the bathroom.  
The kitchen is west of the garden.  
What is north of the bedroom? **A: office**  
What is the bedroom north of? **A: bathroom**

**Task 5: Three Argument Relations**

Mary gave the cake to Fred.  
Fred gave the cake to Bill.  
Jeff was given the milk by Bill.  
Who gave the cake to Fred? **A: Mary**  
Who did Fred give the cake to? **A: Bill**

**Task 6: Yes/No Questions**

John moved to the playground.  
Daniel went to the bathroom.  
John went back to the hallway.  
Is John in the playground? **A:no**  
Is Daniel in the bathroom? **A:yes**

**Task 7: Counting**

Daniel picked up the football.  
Daniel dropped the football.  
Daniel got the milk.  
Daniel took the apple.  
How many objects is Daniel holding? **A: two**

**Task 8: Lists/Sets**

Daniel picks up the football.  
Daniel drops the newspaper.  
Daniel picks up the milk.  
John took the apple.  
What is Daniel holding? **milk, football**

**Task 9: Simple Negation**

Sandra travelled to the office.  
Fred is no longer in the office.  
Is Fred in the office? **A:no**  
Is Sandra in the office? **A:yes**

**Task 10: Indefinite Knowledge**

John is either in the classroom or the playground.  
Sandra is in the garden.  
Is John in the classroom? **A:maybe**  
Is John in the office? **A:no**



**Task 11: Basic Coreference**

Daniel was in the kitchen.  
Then he went to the studio.  
Sandra was in the office.  
Where is Daniel? **A:studio**

**Task 12: Conjunction**

Mary and Jeff went to the kitchen.  
Then Jeff went to the park.  
Where is Mary? **A: kitchen**  
Where is Jeff? **A: park**

**Task 13: Compound Coreference**

Daniel and Sandra journeyed to the office.  
Then they went to the garden.  
Sandra and John travelled to the kitchen.  
After that they moved to the hallway.  
Where is Daniel? **A: garden**

**Task 14: Time Reasoning**

In the afternoon Julie went to the park.  
Yesterday Julie was at school.  
Julie went to the cinema this evening.  
Where did Julie go after the park? **A:cinema**  
Where was Julie before the park? **A:school**

**Task 15: Basic Deduction**

Sheep are afraid of wolves.  
Cats are afraid of dogs.  
Mice are afraid of cats.  
Gertrude is a sheep.  
What is Gertrude afraid of? **A:wolves**

**Task 16: Basic Induction**

Lily is a swan.  
Lily is white.  
Bernhard is green.  
Greg is a swan.  
What color is Greg? **A:white**

**Task 17: Positional Reasoning**

The triangle is to the right of the blue square.  
The red square is on top of the blue square.  
The red sphere is to the right of the blue square.  
Is the red sphere to the right of the blue square? **A:yes**  
Is the red square to the left of the triangle? **A:yes**

**Task 18: Size Reasoning**

The football fits in the suitcase.  
The suitcase fits in the cupboard.  
The box is smaller than the football.  
Will the box fit in the suitcase? **A:yes**  
Will the cupboard fit in the box? **A:no**

**Task 19: Path Finding**

The kitchen is north of the hallway.  
The bathroom is west of the bedroom.  
The den is east of the hallway.  
The office is south of the bedroom.  
How do you go from den to kitchen? **A: west, north**  
How do you go from office to bathroom? **A: north, west**

**Task 20: Agent's Motivations**

John is hungry.  
John goes to the kitchen.  
John grabbed the apple there.  
Daniel is hungry.  
Where does Daniel go? **A:kitchen**  
Why did John go to the kitchen? **A:hungry**

# MCTest (Microsoft Research)

James the Turtle was always getting in trouble. Sometimes he'd reach into the freezer and empty out all the food. Other times he'd sled on the deck and get a splinter. His aunt Jane tried as hard as she could to keep him out of trouble, but he was sneaky and got into lots of trouble behind her back.

One day, James thought he would go into town and see what kind of trouble he could get into. He went to the grocery store and pulled all the pudding off the shelves and ate two jars. Then he walked to the fast food restaurant and ordered 15 bags of fries. He didn't pay, and instead headed home.

His aunt was waiting for him in his room. She told James that she loved him, but he would have to start acting like a well-behaved turtle. After about a month, and after getting into lots of trouble, James finally made up his mind to be a better turtle.

1) What is the name of the trouble making turtle?

- A) Fries
- B) Pudding
- C) James
- D) Jane

2) What did James pull off of the shelves in the grocery store?

- A) pudding
- B) fries
- C) food
- D) splinters

3) Where did James go after he went to the grocery store?

- A) his deck
- B) his freezer
- C) a fast food restaurant
- D) his room



# MCTest (Microsoft Research)

- Short fictional story (150-300 words)
- Annotation: Crowd Sourcing (Amazon Mechanical Turk)
- MCT160, MCT500 (Grammar Test)

Corpus	Stories	Median writing time	Average Words Per:		
			Story	Question	Answer
MC160	160	26 min	204	8.0	3.4
MC500	500	20 min	212	7.7	3.4

MC160	Train and Dev: 400 Q's		Test: 240 Q's	
	SW	SW+D	SW	SW+D
Single	59.46	68.11	64.29	75.89
Multi	59.53	67.44	48.44	57.81
All	59.50	67.75	55.83	66.25

MC500	Train and Dev: 1400 Q's		Test: 600 Q's		All
	SW	SW+D	SW	SW+D	SW+D
Single	55.13	61.77	51.10	57.35	60.44
Multi	49.80	55.28	51.83	56.10	55.53
All	52.21	58.21	51.50	56.67	57.75

# CNN/Daily Mail (Deep Mind)

- 93k CNN articles, 220k Daily Mail articles
- bullet points
- 完形填空

Original Version	Anonymised Version
<b>Context</b>	
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.” ...	the <i>ent381</i> producer allegedly struck by <i>ent212</i> will not press charges against the “ <i>ent153</i> ” host , his lawyer said friday . <i>ent212</i> , who hosted one of the most - watched television shows in the world , was dropped by the <i>ent381</i> wednesday after an internal investigation by the <i>ent180</i> broadcaster found he had subjected producer <i>ent193</i> “ to an unprovoked physical and verbal attack . ” ...
<b>Query</b>	
Producer <b>X</b> will not press charges against Jeremy Clarkson, his lawyer says.	producer <b>X</b> will not press charges against <i>ent212</i> , his lawyer says .
<b>Answer</b>	
Oisin Tymon	<i>ent193</i>

# CNN/Daily Mail

	CNN			Daily Mail		
	train	valid	test	train	valid	test
# months	95	1	1	56	1	1
# documents	90,266	1,220	1,093	196,961	12,148	10,397
# queries	380,298	3,924	3,198	879,450	64,835	53,182
Max # entities	527	187	396	371	232	245
Avg # entities	26.4	26.5	24.5	26.5	25.5	26.0
Avg # tokens	762	763	716	813	774	780
Vocab size	118,497			208,045		

Top N	Cumulative %	
	CNN	Daily Mail
1	30.5	25.6
2	47.7	42.4
3	58.1	53.7
5	70.6	68.1
10	85.1	85.5

# CBT (Children's Book Test)

## Facebook

- Children's Book
- Contexts: 20 sentences
- Answer: entity in the 21st sentence

	TRAINING	VALIDATION	TEST
NUMBER OF BOOKS	98	5	5
NUMBER OF QUESTIONS (CONTEXT+QUERY)	669,343	8,000	10,000
AVERAGE WORDS IN CONTEXTS	465	435	445
AVERAGE WORDS IN QUERIES	31	27	29
DISTINCT CANDIDATES	37,242	5,485	7,108
VOCABULARY SIZE	53,628		



# CBT (Children's Book Test)

## Facebook

"Well, Miss Maxwell, I think it only fair to tell you that you may have trouble with those boys when they do come. Forewarned is forearmed, you know. Mr. Cropper was opposed to our hiring you. Not, of course, that he had any personal objection to you, but he is set against female teachers, and when a Cropper is set there is nothing on earth can change him. He says female teachers can't keep order. He 's started in with a spite at you on general principles, and the boys know it. They know he'll back them up in secret, no matter what they do, just to prove his opinions. Cropper is sly and slippery, and it is hard to corner him."

"Are the boys big ?" queried Esther anxiously.

"Yes. Thirteen and fourteen and big for their age. You can't whip 'em -- that is the trouble. A man might, but they'd twist you around their fingers. You'll have your hands full, I'm afraid. But maybe they'll behave all right after all."

Mr. Baxter privately had no hope that they would, but Esther hoped for the best. She could not believe that Mr. Cropper would carry his prejudices into a personal application. This conviction was strengthened when he overtook her walking from school the next day and drove her home. He was a big, handsome man with a very suave, polite manner. He asked interestedly about her school and her work, hoped she was getting on well, and said he had two young rascals of his own to send soon. Esther felt relieved. She thought that Mr. Baxter had exaggerated matters a little.

S: 1 Mr. Cropper was opposed to our hiring you .  
2 Not , of course , that he had any personal objection to you , but he is set against female teachers , and when a Cropper is set there is nothing on earth can change him .  
3 He says female teachers ca n't keep order .  
4 He 's started in with a spite at you on general principles , and the boys know it .  
5 They know he 'll back them up in secret , no matter what they do , just to prove his opinions .  
6 Cropper is sly and slippery , and it is hard to corner him . ''  
7 `` Are the boys big ? ''  
8 queried Esther anxiously .  
9 `` Yes .  
10 Thirteen and fourteen and big for their age .  
11 You ca n't whip 'em -- that is the trouble .  
12 A man might , but they 'd twist you around their fingers .  
13 You 'll have your hands full , I 'm afraid .  
14 But maybe they 'll behave all right after all . ''  
15 Mr. Baxter privately had no hope that they would , but Esther hoped for the best.  
16 She could not believe that Mr. Cropper would carry his prejudices into a personal application .  
17 This conviction was strengthened when he overtook her walking from school the next day and drove her home .  
18 He was a big , handsome man with a very suave , polite manner .  
19 He asked interestedly about her school and her work , hoped she was getting on well , and said he had two young rascals of his own to send soon .  
20 Esther felt relieved .

Q: She thought that Mr. \_\_\_\_\_ had exaggerated matters a little .

C: Baxter, Cropper, Esther, course, fingers, manner, objection, opinion, right, spite.

a: Baxter

# 任务

contexts/documents

question/query

answers/candidates

Document

One night I was at my friend's house where he threw a party. We were enjoying our dinner at night when all of a sudden we heard a knock on the door. I opened the door and saw this guy who had scar on his face. (.....)As soon as I saw him I ran inside the house and called the cops. The cops came and the guy ran away as soon as he heard the cop car coming. We never found out what happened to that guy after that day.

Question

1: What was the strange guy doing with the friend?

Candidate answers

- A) enjoying a meal
- B) talking about his job
- C) talking to him
- \*D) trying to beat him

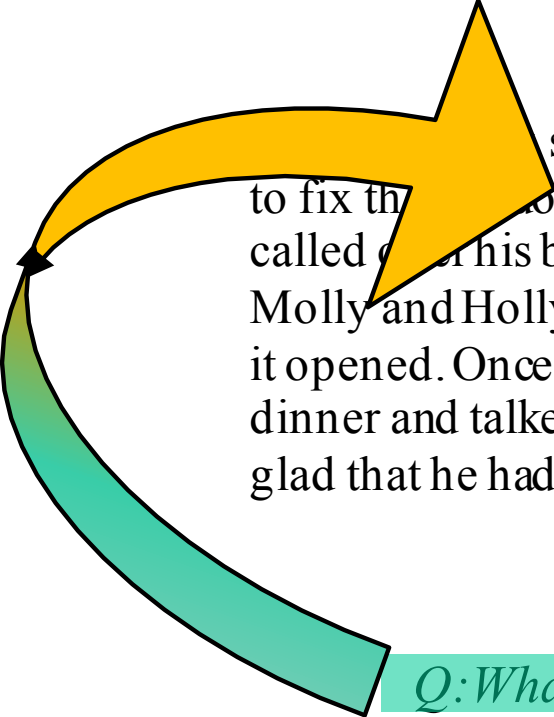
2: Why did the strange guy run away?

- \*A) because he heard the cop car
- B) because he saw his friend
- C) because he didn't like the dinner
- D) because it was getting late

# 传统方法

- Two-Steps
  - Facts/Statements Retrieval
  - Inferencing Answers

# Facts Retrieval



some things around the house. He had to fix the door. He had to fix the window. But before he did anything he had to fix the toilet. Tom called on his best friend Jim to help him. Jim brought with him his friends Molly and Holly.[...]. They all pushed on the window really hard until finally it opened. Once the window was fixed the four of them made a delicious dinner and talked about all of the good work that they had done. Tom was glad that he had such good friends to help him with his work.

*Q: What did Tom need to fix first?*

- A) Door
- B) House
- C) Window
- \*D) Toilet



# Inferencing Answers

Tom had to fix some things around the house. He had to fix the door. He had to fix the window. **But before he did anything he had to fix the toilet.** Tom called over his best friend Jim to help him. Jim brought with him his friends Molly and Holly. [...]. They all pushed on the window really hard until finally it opened. Once the window was fixed the four of them made a delicious dinner and talked about all of the good work that they had done. Tom was glad that he had such good friends to help him with his work.

**RTE**

*Q: What did Tom need to fix first?*

- A) Door
- B) House
- C) Window
- \*D) Toilet

Tom need to fix toilet first.

# Formulation

Document:  $D$

Document sentences:  $\{s_1, \dots, s_n\}$

$S$ : support sentences to answer the question

Question:  $q$

Candidate Answers:  $\{a_1, \dots, a_4\}$

$$p(a|q, D) = \underbrace{p(S|q, D)}_{AS} \underbrace{p(a|q, S)}_{RTE}$$

# Feature Engineering for RTE

- Alignments among words (Sachan, et al. ACL 2015)
- Discourse Relation: Causal, Temporal, Explanation (Narasimhan et al. ACL 2015)
- Linguistic Features: POS, Dependency, Word Embeddings, Frames (Wang et al. ACL 2015)

Relation	Word list
Causal	because, why, due, so
Temporal	when, between, soon, before, after, during, then, finally, now, nowadays, first
Explanation	how, by, using

# Difficulties

- 上述方法仍然基于符号表示，语义鸿沟问题
- 对于简单问题比较有效，但是对于需要深层次语义推理的问题，往往无法处理
- 多级处理带来的噪声积累和传递问题

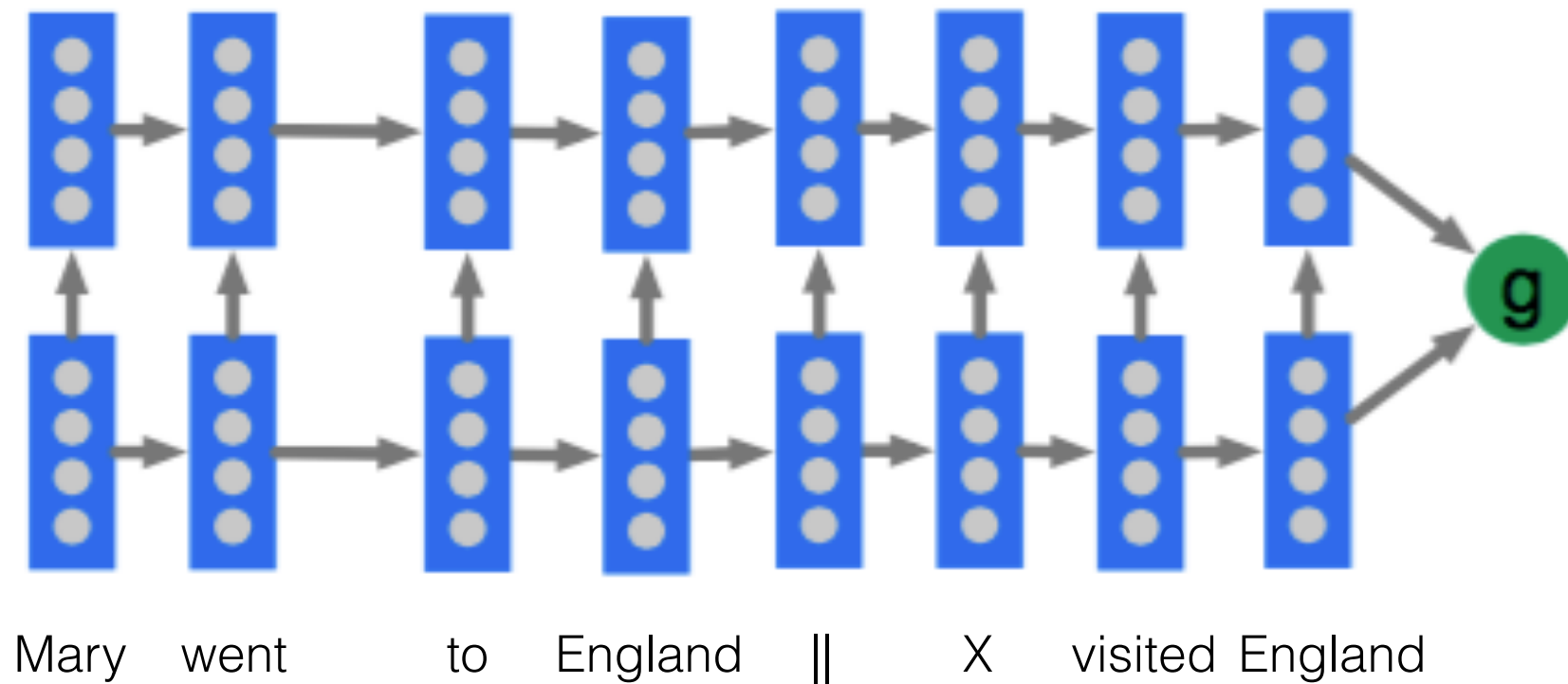


# 基于深度学习的机器阅读理解

- LSTM (Hermann, et al. 2015)
- Attention based Model (Chen, et al. 2016, Hermann, et al. 2015)
- Memory Network (Weston, et al. 2015)
- Hierarchical CNN (Yin, et al. 2015)
- Hierarchical Attention (Fang, et al. 2016)
- Neural Turing Machine (Graves, et al. 2014)

$$p(a|d, q) \propto \exp(W(a)g(d, q))$$

# LSTM



$$x'(t, k) = x(t) || y'(t, k-1), \quad y(t) = y'(t, 1) || \dots || y'(t, K)$$

$$i(t, k) = \sigma(W_{kxi}x'(t, k) + W_{khi}h(t-1, k) + W_{kci}c(t-1, k) + b_{ki})$$

$$f(t, k) = \sigma(W_{kxf}x(t) + W_{khf}h(t-1, k) + W_{kcf}c(t-1, k) + b_{kf})$$

$$c(t, k) = f(t, k)c(t-1, k) + i(t, k) \tanh(W_{kxc}x'(t, k) + W_{khc}h(t-1, k) + b_{kc})$$

$$o(t, k) = \sigma(W_{kxo}x'(t, k) + W_{kho}h(t-1, k) + W_{kco}c(t, k) + b_{ko})$$

$$h(t, k) = o(t, k) \tanh(c(t, k))$$

$$y'(t, k) = W_{ky}h(t, k) + b_{ky}$$

$$g^{\text{LSTM}}(d, q) = y(|d| + |q|)$$

# Attention Model

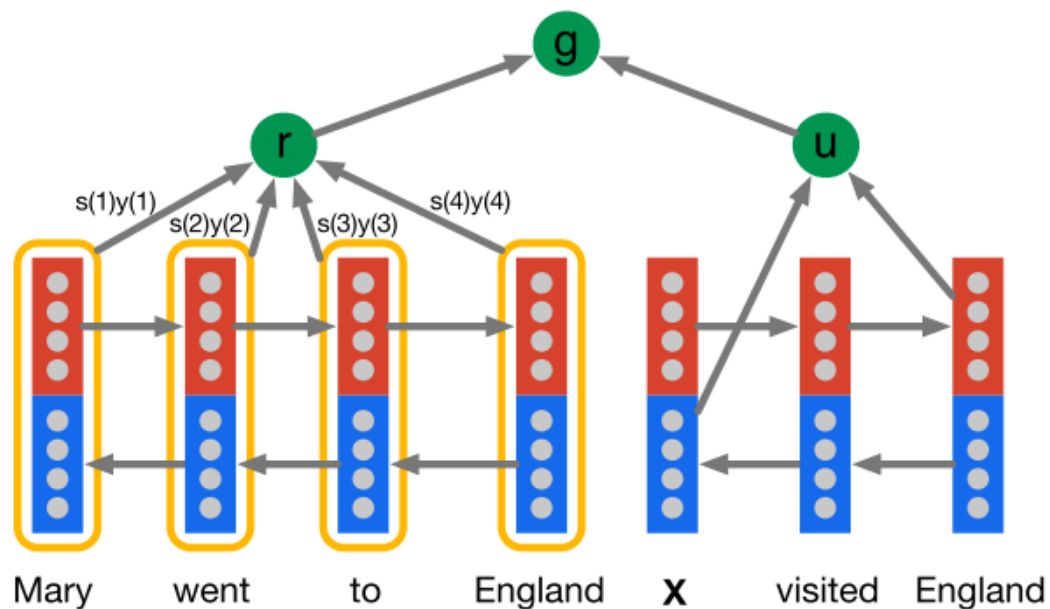
Context : Michael Jordan abruptly retired from Chicago Bulls before the beginning of the 1993-94 NBA season to pursue a career in baseball.

Question1: When did Michael Jordan retired from NBA?

Question2: Which sports does Michael Jordan participate after his retirement from NBA?

# Attention Model

## Attention Reader



$$y_d(t) = \vec{y}_d(t) \parallel \overleftarrow{y}_d(t) \quad u = \vec{y}_q(|q|) \parallel \overleftarrow{y}_q(1)$$

$$m(t) = \tanh(W_{ym}y_d(t) + W_{um}u)$$

$$s(t) \propto \exp(w_{ms}^\top m(t)),$$

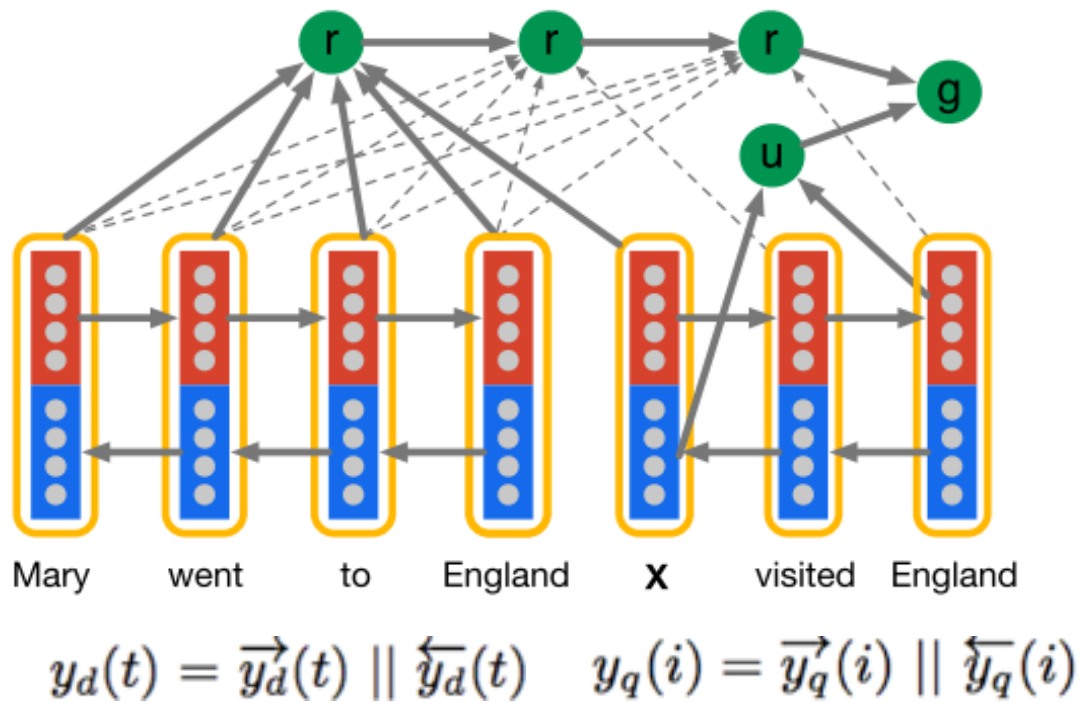
$$r = y_d s,$$

$$g^{\text{AR}}(d, q) = \tanh(W_{rg}r + W_{ug}u)$$



# Attention Model

## Impatient Reader



$$m(i, t) = \tanh(W_{dm}y_d(t) + W_{rm}r(i-1) + W_{qm}y_q(i))$$

$$s(i, t) \propto \exp(w_{ms}^\top m(i, t)),$$

$$r(0) = \mathbf{r}_0, \quad r(i) = y_d^\top s(i) + \tanh(W_{rr}r(i-1))$$

$$g^{\text{IR}}(d, q) = \tanh(W_{rg}r(|q|) + W_{qg}u)$$

# Memory Network

- 4 Components:
  - I (Input feature map): 将输入转换为内部的特征表示
  - G (Generalization): 根据输入更新Memory
  - O (Output feature map): 根据输入和当前Memory状态, 生成输出 (在特征空间中)
  - R (Response): 根据输出的特征表示, 生成答案

# Memory Network

Joe went to the kitchen. Fred went to the kitchen. Joe picked up the milk.

Joe travelled to the office. Joe left the milk. Joe went to the bathroom.

Where is the milk now? **A: office**

Where is Joe? **A: bathroom**

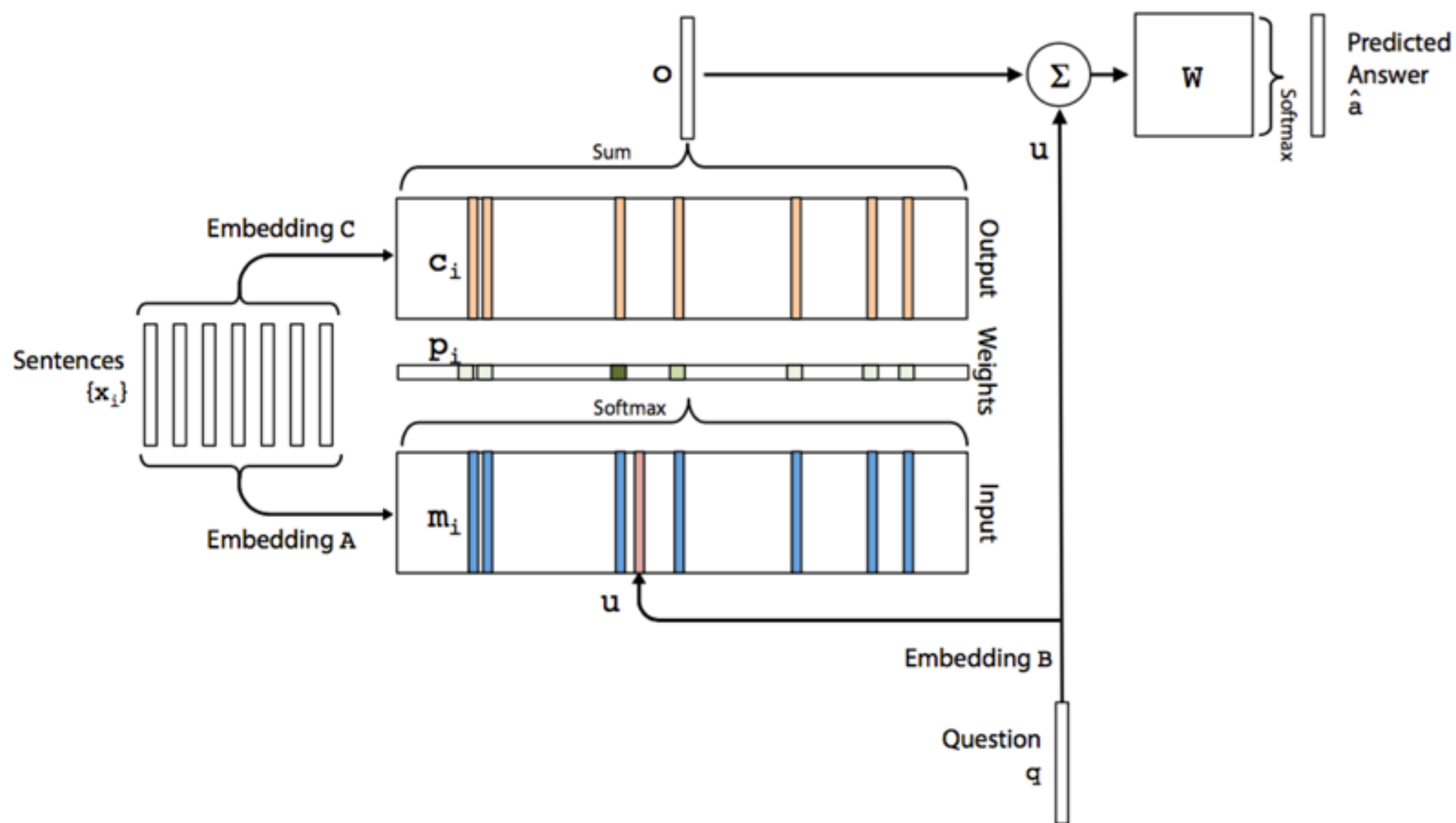
Where was Joe before the office? **A: kitchen**

$$o_1 = O_1(x, \mathbf{m}) = \arg \max_{i=1, \dots, N} s_O(x, \mathbf{m}_i)$$

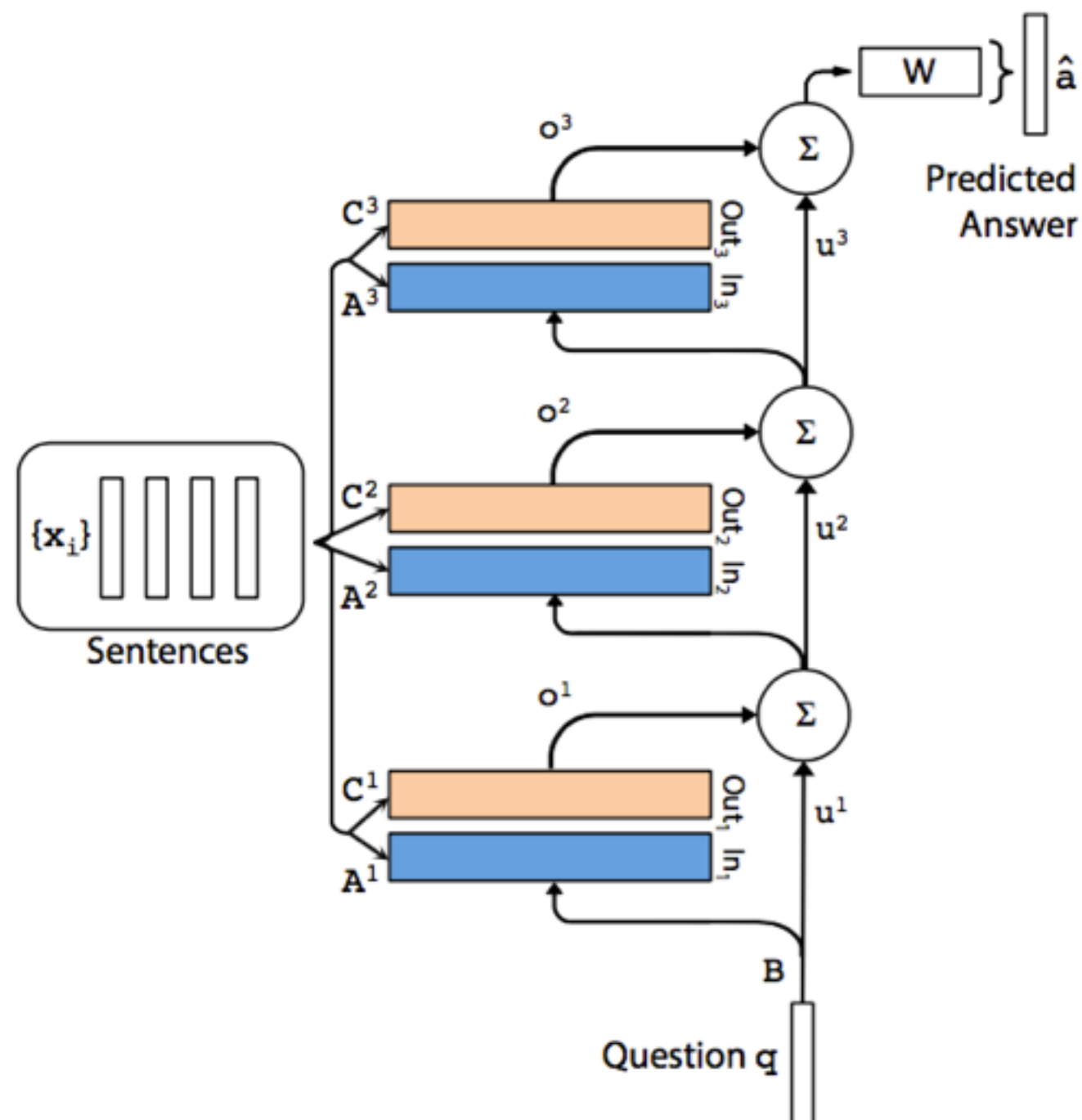
$$o_2 = O_2(x, \mathbf{m}) = \arg \max_{i=1, \dots, N} s_O([x, \mathbf{m}_{o_1}], \mathbf{m}_i)$$

$$r = \operatorname{argmax}_{w \in W} s_R([x, \mathbf{m}_{o_1}, \mathbf{m}_{o_2}], w)$$

# Memory Network: Single Lookup



# Memory Network: Multiple Lookup



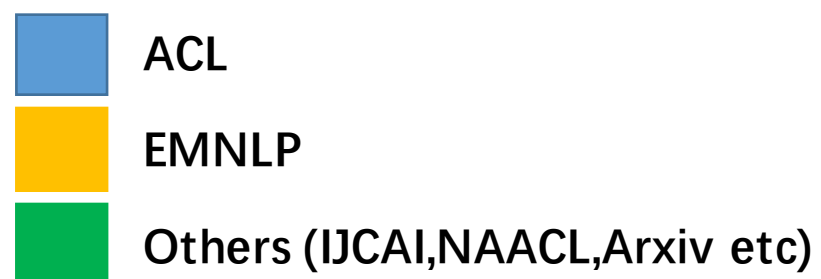
# 实验比较 (CNN/Daily Mail)

Model	CNN		Daily Mail	
	Dev	Test	Dev	Test
Frame-semantic model <sup>†</sup>	36.3	40.2	35.5	35.5
Word distance model <sup>†</sup>	50.5	50.9	56.4	55.5
Deep LSTM Reader <sup>†</sup>	55.0	57.0	63.3	62.2
Attentive Reader <sup>†</sup>	61.6	63.0	70.5	69.0
Impatient Reader <sup>†</sup>	61.8	63.8	69.0	68.0
MemNNs (window memory) <sup>‡</sup>	58.0	60.6	N/A	N/A
MemNNs (window memory + self-sup.) <sup>‡</sup>	63.4	66.8	N/A	N/A
MemNNs (ensemble) <sup>‡</sup>	66.2*	69.4*	N/A	N/A
Ours: Classifier	67.1	67.9	69.1	68.3
Ours: Neural net	<b>72.4</b>	<b>72.4</b>	<b>76.9</b>	<b>75.8</b>

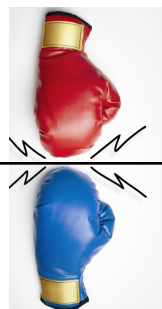
Chen, et al. 2016



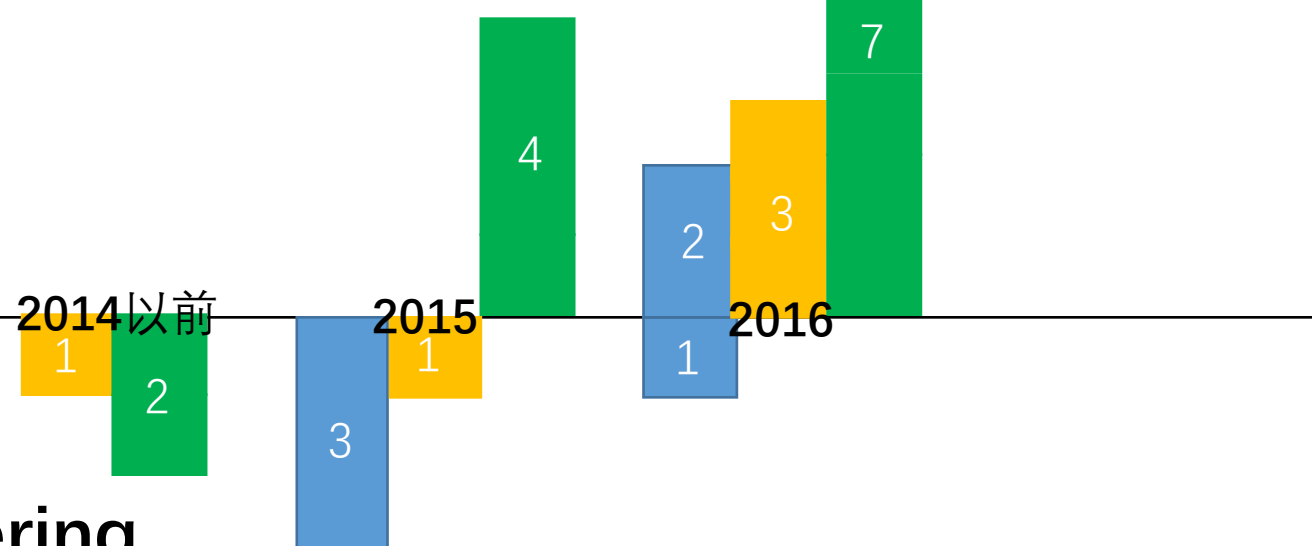
# 两类方法



Deep Learning



Feature Engineering



# 数据分析

Category	Question	Passage
Exact Match	<i>it 's clear @entity0 is leaning toward @placeholder</i> , says an expert who monitors @entity0	... @entity116 , who follows @entity0 's operations and propaganda closely , recently told @entity3 , <i>it 's clear @entity0 is leaning toward @entity60</i> in terms of doctrine , ideology and an emphasis on holding territory after operations . . . .
Paraphrase	@placeholder says he understands why @entity0 wo n't play at his tournament	... @entity0 called me personally to let me know that he would n't be playing here at @entity23 , " @entity3 said on his @entity21 event 's website . . . .
Partial clue	a tv movie based on @entity2 's book @placeholder casts a @entity76 actor as @entity5	... to @entity12 @entity2 professed that his @entity11 is not a religious book . . . .
Multiple sent.	he 's doing a his - and - her duet all by himself , @entity6 said of @placeholder	... we got some groundbreaking performances , here too , tonight , @entity6 said . we got @entity17 , who will be doing some musical performances . he 's doing a his - and - her duet all by himself . . . .
Coref. Error	rapper @placeholder " disgusted , " cancels upcoming show for @entity280	... with hip - hop star @entity246 saying on @entity247 that he was canceling an upcoming show for the @entity249 . . . . (but @entity249 = @entity280 = SAEs)
Hard	pilot error and snow were reasons stated for @placeholder plane crash	... a small aircraft carrying @entity5 , @entity6 and @entity7 the @entity12 @entity3 crashed a few miles from @entity9 , near @entity10 , @entity11 . . . .

# 数据分析

No.	Category	(%)
1	Exact match	13
2	Paraphrasing	41
3	Partial clue	19
4	Multiple sentences	2
5	Coreference errors	8
6	Ambiguous / hard	17

数据

Category	Classifier	Neural net
Exact match	13 (100.0%)	13 (100.0%)
Paraphrasing	32 (78.1%)	39 (95.1%)
Partial clue	14 (73.7%)	17 (89.5%)
Multiple sentences	1 (50.0%)	1 (50.0%)
Coreference errors	4 (50.0%)	3 (37.5%)
Ambiguous / hard	2 (11.8%)	1 (5.9%)
All	66 (66.0%)	74 (74.0%)

结果

简单的Attention Model已经可以达到“天花板”  
但是对于需要深度推理的问题，效果仍然不好

# 数据规模的影响

"Well, Miss Maxwell, I think it only fair to tell you that you may have trouble with those boys when they do come. Forewarned is forearmed, you know. Mr. Cropper was opposed to our hiring you. Not, of course, that he had any personal objection to you, but he is set against female teachers, and when a Cropper is set there is nothing on earth can change him. He 's started in with a spite at you on general principles , and the boys know it . They know he'll back them up in secret, no matter what they do, just to prove his opinions. Cropper is sly and slippery, and it is hard to corner him."	S: 1 Mr. Cropper was opposed to our hiring you . 2 Not , of course , that he had any personal objection to you , but he is set against female teachers , and when a Cropper is set there is nothing on earth can change him . 3 He says female teachers ca n't keep order . 4 He 's started in with a spite at you on general principles , and the boys know it . 5 They know he 'll back them up in secret , no matter what they do , just to prove his opinions . 6 Cropper is sly and slippery , and it is hard to corner him . '' 7 `` Are the boys big ? '' 8 queried Esther anxiously . 9 `` Yes . 10 Thirteen and fourteen and big for their age . 11 You ca n't whip 'em -- that is the trouble . 12 A man might , but they 'd twist you around their fingers . 13 You 'll have your hands full , I 'm afraid . 14 But maybe they 'll behave all right after all . '' 15 Mr. Baxter privately had no hope that they would , but Esther hoped for the best . 16 She could not believe that Mr. Cropper would carry his prejudices into a personal application . 17 This conviction was strengthened when he overtook her walking from school the next day and drove her home . 18 He was a big , handsome man with a very suave , polite manner . 19 He asked interestedly about her school and her work , hoped she was getting on well , and said he had two young rascals of his own to send soon . 20 Esther felt relieved .
"Are the boys big ?" queried Esther anxiously.	Q: She thought that Mr. _____ had exaggerated matters a little .
"Yes. Thirteen and fourteen and big for their age. You can't whip 'em -- that is the trouble. A man might, but they'd twist you around their fingers. You'll have your hands full, I'm afraid. But maybe they'll behave all right after all."	C: Baxter, Cropper, Esther, course, fingers, manner, objection, opinion, right, spite.
Mr. Baxter privately had no hope that they would, but Esther hoped for the best. She could not believe that Mr. Cropper would carry his prejudices into a personal application. This conviction was strengthened when he overtook her walking from school the next day and drove her home. He was a big, handsome man with a very suave, polite manner. He asked interestedly about her school and her work, hoped she was getting on well, and said he had two young rascals of his own to send soon. Esther felt relieved. She thought that Mr. Baxter had exaggerated matters a little.	d: Baxter

	CNN	Daily Mail	CBT CN	CBT NE	BookTest
# queries	380,298	879,450	120,769	108,719	14,140,825
Max # options	527	371	10	10	10
Avg # options	26.4	26.5	10	10	10
Avg # tokens	762	813	470	433	522
Vocab. size	118,497	208,045	53,185	53,063	1,860,394

语料规模较CBT增加了60倍



	Named entity		Common noun		
	valid	test	valid	test	
Humans (query) (Hill et al., 2015)	NA	52.0	NA	64.4	} CBT training data
Humans (context+query) (Hill et al., 2015)	NA	<b>81.6</b>	NA	<b>81.6</b>	
LSTMs (context+query) (Hill et al., 2015)	51.2	41.8	62.6	56.0	
Memory Networks (Hill et al., 2015)	70.4	66.6	64.2	63.0	
AS Reader (single model)	73.8	68.6	68.8	63.4	
<b>AS Reader (avg ensemble)</b>	74.5	70.6	71.1	<b>68.9</b>	
<b>AS Reader (greedy ensemble)</b>	76.2	<b>71.0</b>	72.4	67.5	
GA Reader (ensemble) (Dhingra et al., 2016)	75.5	71.9	72.1	69.4	
EpiReader (ensemble) (Trischler et al., 2016b)	76.6	71.8	73.6	70.6	
<b>IA Reader (ensemble)</b> (Sordoni et al., 2016)	76.9	<b>72.0</b>	74.1	<b>71.0</b>	
<b>AoA Reader (single model)</b> (Cui et al., 2016a)	77.8	<b>72.0</b>	72.2	69.4	
<b>AS Reader (single model)</b>	80.5	76.2	83.2	80.8	} BookTest training data
<b>AS Reader (greedy ensemble)</b>	82.3	<b>78.4</b>	85.7	<b>83.7</b>	

# 小结

- 传统基于符号表示的特征工程方法
  - 难以适应大规模、开放域的数据需求
- 基于分布式表示的Deep Learning的方法
  - 专注于表示学习
  - 忽略了深度推理



谢谢! Q&A!