Towards Generating Human-like Deep Questions
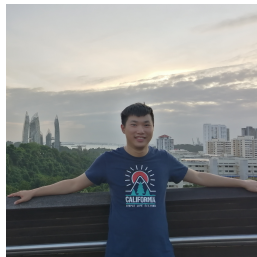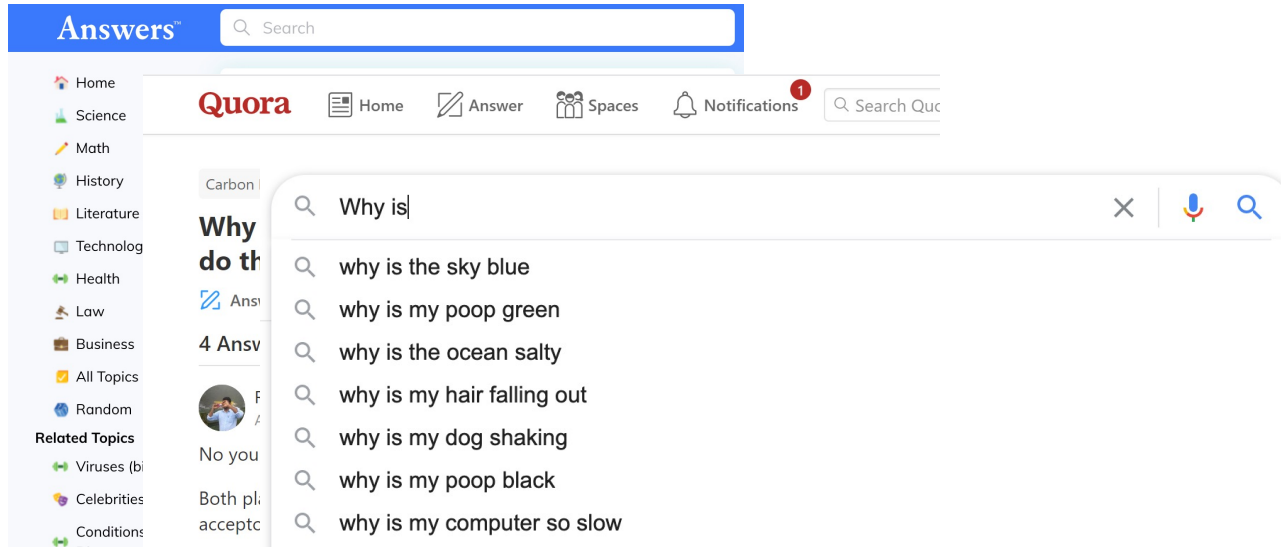
Liangming Pan

Curiosity is the beginning of wisdom

Can machines ask questions like humans?

Question Generation

Context $\rightarrow$ generate $\rightarrow$ Question

❏ Texts

Why do humans take O2 to produce CO2 while plants do the opposite?

❏ Knowledge

Were Bill Gates and Satya Nadella once colleagues?

❏ Images

Who is having a birthday?

Question Generation

What can question generation be used for?
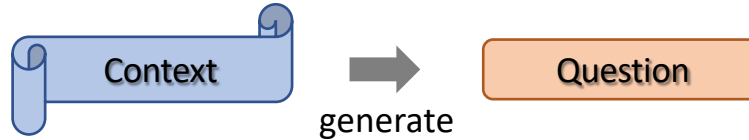
❏ Texts

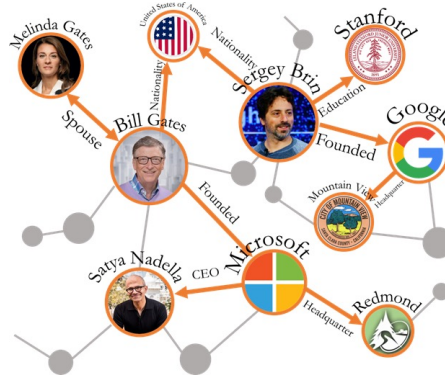WIKIPEDIA
The Free Encyclopedia

Article | Talk

Photosynthesis

From Wikipedia, the free encyclopedia

Photosynthesis is a process used by plants and other organisms to convert light energy into chemical... be released to fuel the organism's metabolic activities. This chemical energy is stored in carbohydrate... from carbon dioxide and water – hence the name *photosynthesis*, from the Greek *phôs* (φῶς), "light",... In most cases, oxygen is also released as a waste product. Most plants, algae, and cyanobacteria per... photoautotrophs. Photosynthesis is largely responsible for producing and maintaining the oxygen con... the energy necessary for life on Earth.[4]

Although photosynthesis is performed differently by different species, the process always begins whe... reaction centres that contain green chlorophyll pigments. In plants, these proteins are held inside org... in leaf cells, while in bacteria they are embedded in the plasma membrane. In these light-dependent r... from suitable substances, such as water, producing oxygen gas. The hydrogen freed by the splitting o... compounds that serve as short-term stores of energy, enabling its transfer to drive other reactions: the... dinucleotide phosphate (NADPH) and adenosine triphosphate (ATP), the "energy currency" of cells.

Why do humans take O2 to produce CO2 while plants do the opposite?

❏ Knowledge

Were Bill Gates and Satya Nadella once colleagues?
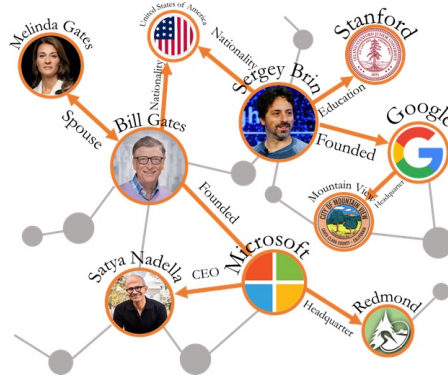
❏ Images

Who is having a birthday?

4

Applications of Question Generation

❏ Education

Generate quiz questions from KG

evaluate
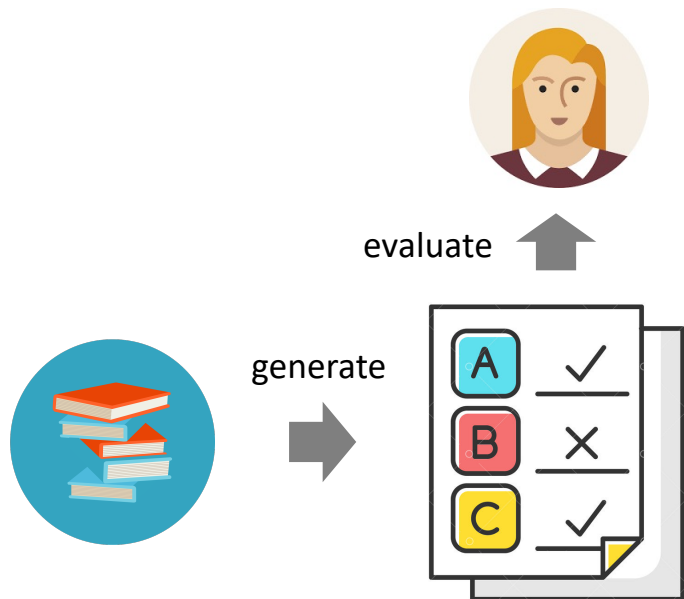
generate

(Seyler et al., ICTIR 2017)

❏ Dialogue

Generate clarification questions

I am working for an employer in Canada. Do I need to carry on paying UK National Insurance

Have you been working abroad 52 weeks or less?
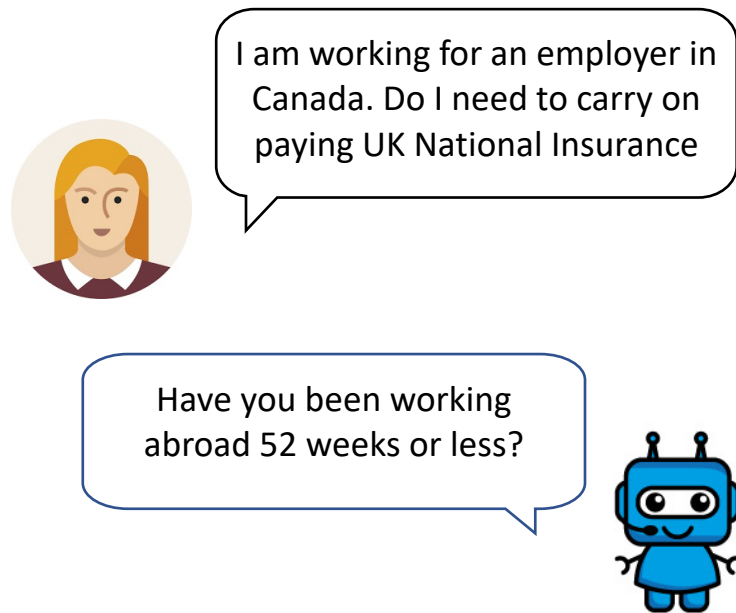
(Marzieh et al., EMNLP 2018)

Applications of Question Generation

❑ Question Answering

Data augmentation for QA
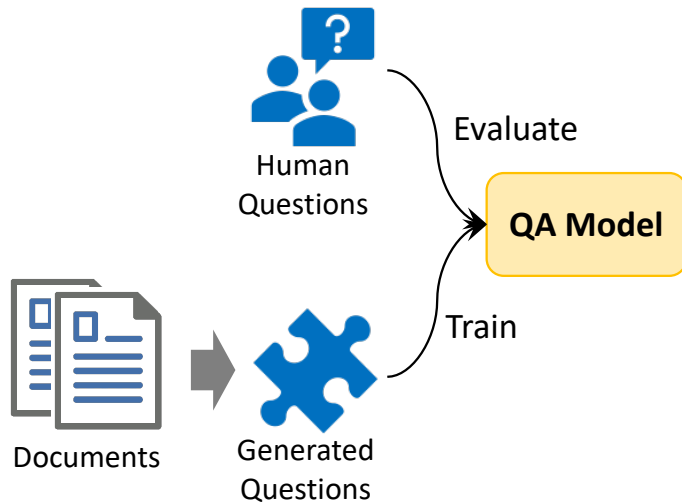
(Duan et al., EMNLP 2017)
(Lewis et al., ACL 2019)
(Puri et al., EMNLP 2020)

❑ Reading Comprehension

QA-style Summarization

Massive Attack (band)

On 21 January 2016, the iPhone application "Fantom" was released. The application was developed by a team including Massive Attack's Robert Del Naja and let users hear parts of four new songs by remixing them in real time, using the phone's location, movement, clock, heartbeat, and camera. On 28 January 2016, Massive Attack released a new EP, Ritual Spirit, which includes the four songs released on Fantom.

Q. What was the iPhone application Fantom?
A. The app... let users hear parts of ... real time,

Q. Who created it?
A. ... team including ... Robert Del Naja

Q. What is Ritual Spirit?
A. On ... Attack released a new EP Ritual Spirit

Q. When did they release a song with "Fantom"?
A. On 28 January 2016,

...

(Krishna and Iyyer, ACL 2019)

Question depth: Bloom's Taxonomy

Produce new or original idea
how would you design a new..? what would happen if you...?

Justify a stand or decision
do you agree that...? what do you think about...?

Draw connections among ideas
what are the parts or features of...? Classify...according to...;

Use information in new situation
how is...an example of...?; how is...related to...?;

Explain ideas or concepts
why is … happen? what does … mean?

Recall facts and basic concepts
who, what, when, where, how..? describe...?

Creating

Evaluating

Analyzing

Applying

Understanding

Remembering

Question depth: Bloom's Taxonomy

Current **faculty** include the anthropologist Marshall Sahlins and **the Shakespeare scholar** David Bevington.

SQuAD

What **Shakespeare scholar** is currently on the **faculty**?
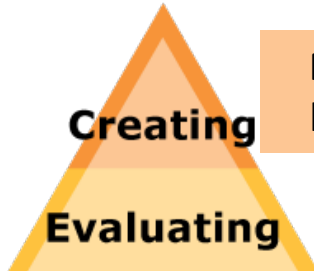
Understanding

Remembering

Explain ideas or concepts
why is … happen? what does … mean?

Recall facts and basic concepts
who, what, when, where, how..? describe…?

Question depth: Bloom's Taxonomy

Creating

Produce new or original idea
how would you design a new..? what would happen if you...?

Evaluating

Justify a stand or decision
do you agree that...? what do you think about...?

TED Ed
LESSONS WORTH SHARING

Science & Technology
Is there a center of the universe?

04:14

545,587 views

Most of the astronomers, scientists, and philosophers had ideas about the universe that we now know were not entirely correct. **In what ways do you think ideas that have turned out to be "wrong" are valuable in developing our understanding of the universe?**

Why asking deep question is hard?

Ancient Greek believes that the earth is the center of the universe.

In 1543, Copernicus proposed that the sun was at the center of the universe.

They are all proved wrong.

Science & Technology
Is there a center of the universe?

04:14

545,587 views

Understanding, Reasoning

Most of the scientists, and philosophers had ideas about the universe that we now know were not entirely correct.

Failures are also valuable in scientific discovery.

Are those "wrong ideas" about the universe center also valuable? If yes, why they are valuable?

Information-Seeking

Prior Knowledge

Why asking deep question is hard?

Ancient Greek believes that the earth is the center of the universe.

In 1543, Copernicus proposed that the sun was at the center of the universe.
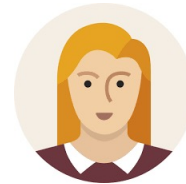
They are all proved wrong.

Science & Technology

Is there a center of the universe?

04:14

545,587 views

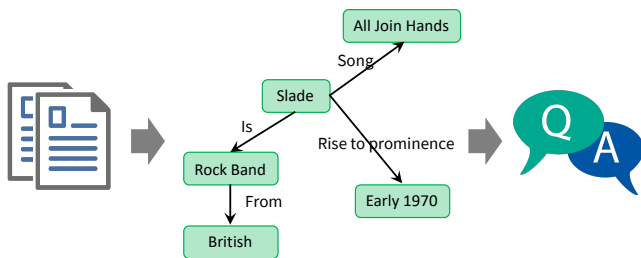Understanding, Reasoning

Prior Knowledge

Information-Seeking

In what ways do you think ideas that have turned out to be "wrong" are valuable in developing our understanding of the universe?

Core Problems (My Interests)

How to **generate** questions that promote deep reasoning?

[Pan et al., ACL 2020]

[Pan et al., NAACL 2021]

How does QG **benefits** downstream NLP applications?
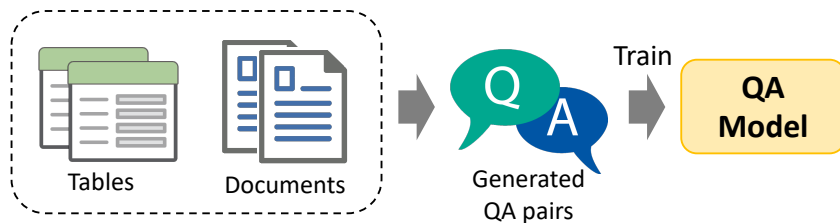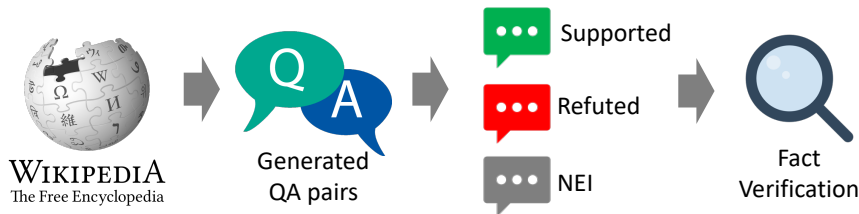

[Pan et al., ACL 2021]

Core Problems (My Interests)

How to **generate** questions that promote deep reasoning?

[Pan et al., ACL 2020]

[Pan et al., NAACL 2021]

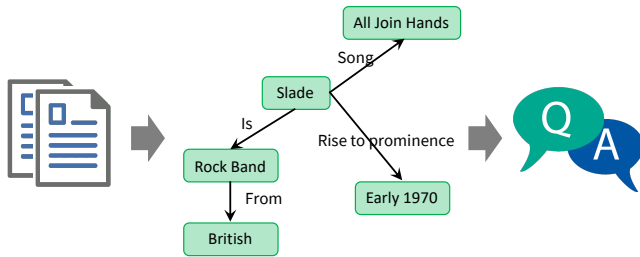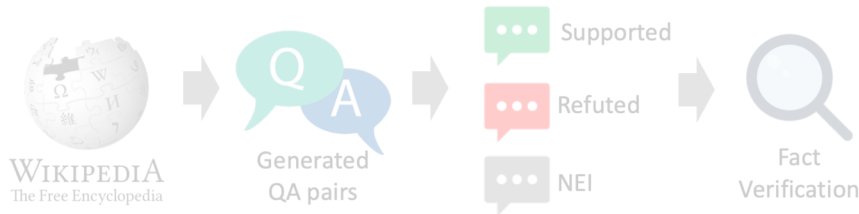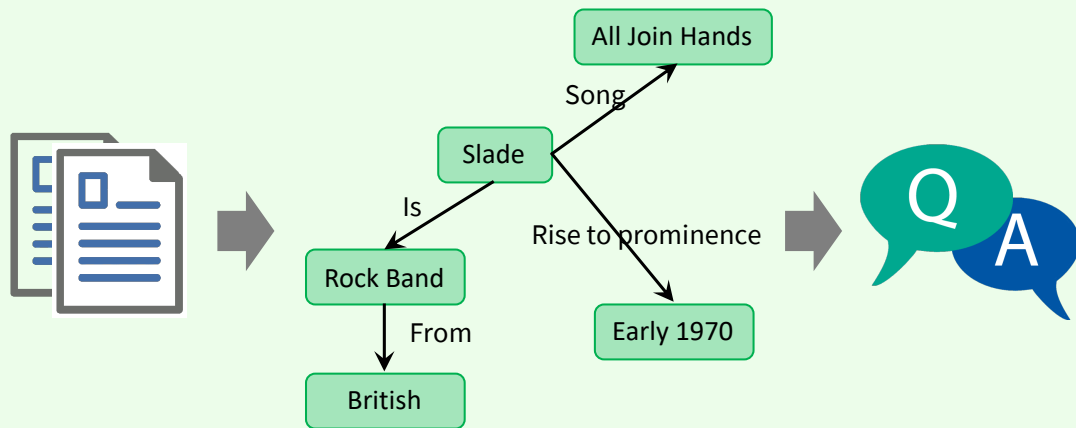How does QG **benefits** downstream NLP applications?

[Pan et al., ACL 2021]

Reasoning with Semantic Graphs

Liangming Pan, Yuxi Xie, Yansong Feng, Tat-Seng Chua, Min-Yen Kan

[ACL 2020] Pan et al: Semantic Graphs for Generating Deep Questions

Multi-hop Question Generation

HotpotQA

A Dataset for Diverse, Explainable Multi-hop Question Answering

> *Paragraph A:* All Join Hands
> "All Join Hands" is a song by the British rock band Slade, released in 1984 as the lead single from the band's twelfth studio album "Rogues Gallery".

> *Paragraph B:* Slade
> Slade are an English glam rock band from Wolverhampton. They rose to prominence during the early 1970s with 17 consecutive top 20 hits and six number ones on the UK Singles Chart.

Question:
When did the rock band that sang "All Join Hands" rise to prominence?
Answer: The early 1970s

(Yang et al., EMNLP 2018)

❑ We choose **multi-hop questions**, a typical type of deep questions, to explore **reasoning in QG**.

❑ To generate a multi-hop question, it requires **the integration and reasoning over different pieces of information**.

❑ In contrast, **shallow QG** only requires local contexts and rarely involves reasoning.

The Seq2Seq Baseline

❑ The key to multi-hop reasoning is to:

   o   Understand the relationship between entities.

   o   Select reasoning chains and perform multi-step reasoning.

Difficult when representing inputs as a sequence of tokens.

Semantic Graphs

How about we represent the documents in a structured way?

Paragraph A: All Join Hands
"All Join Hands" is a song by the British rock band Slade, released in 1984 as the lead single from the band's twelfth studio album "Rogues Gallery".

Paragraph B: Slade
Slade are an English glam rock band from Wolverhampton. They rose to prominence during the early 1970s with 17 consecutive top 20 hits and six number ones on the UK Singles Chart.
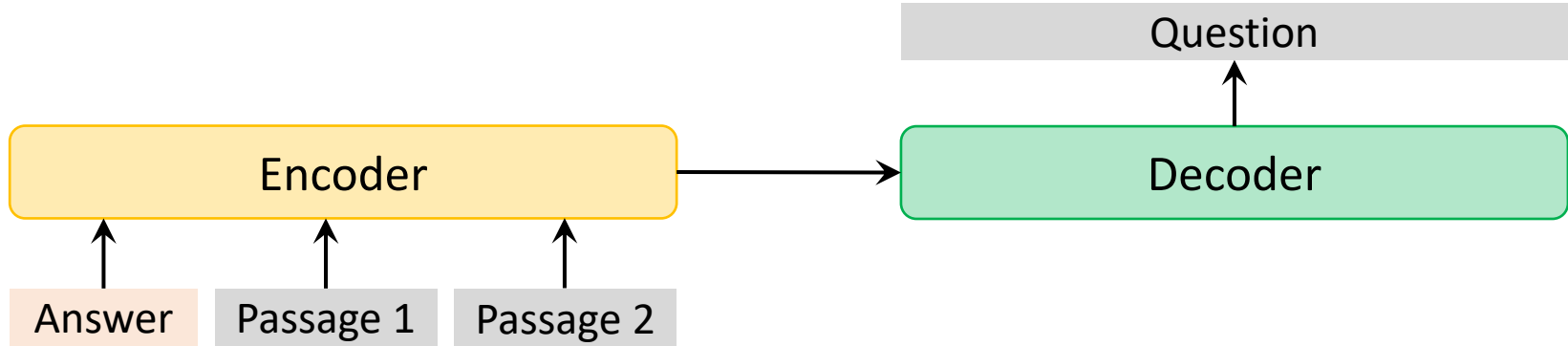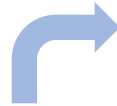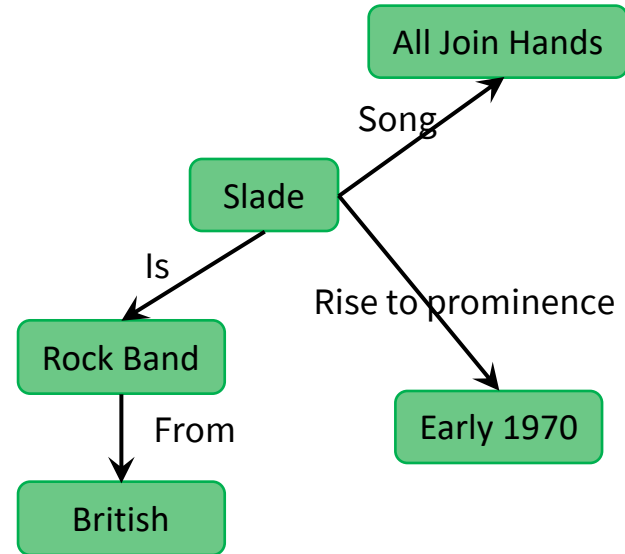
Semantic Graphs

How about we represent the documents in a structured way?

Paragraph A: All Join Hands
"All Join Hands" is a song by the British rock band Slade, released in 1984 as the lead single from the band's twelfth studio album "Rogues Gallery".

Paragraph B: Slade
Slade are an English glam rock band from Wolverhampton. They rose to prominence during the early 1970s with 17 consecutive top 20 hits and six number ones on the UK Singles Chart.
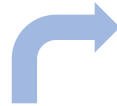
All Join Hands

Slade

Song

Is

Rise to prominence

Rock Band

Early 1970

From

British

Semantic Graphs

❑ Different reasoning chains correspond to different questions.

All Join Hands

Song

Slade

Is

Rock Band

From

British

Rise to prominence

Early 1970

Question: When did the rock band that sang "All Join Hands" rise to prominence?

Answer: The early 1970s

Semantic Graphs

❑ Different reasoning chains correspond to different questions.

Question: Which country is the rock band that sang "All Join Hands" from?

Answer: Britain

Constructing Semantic Graphs

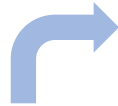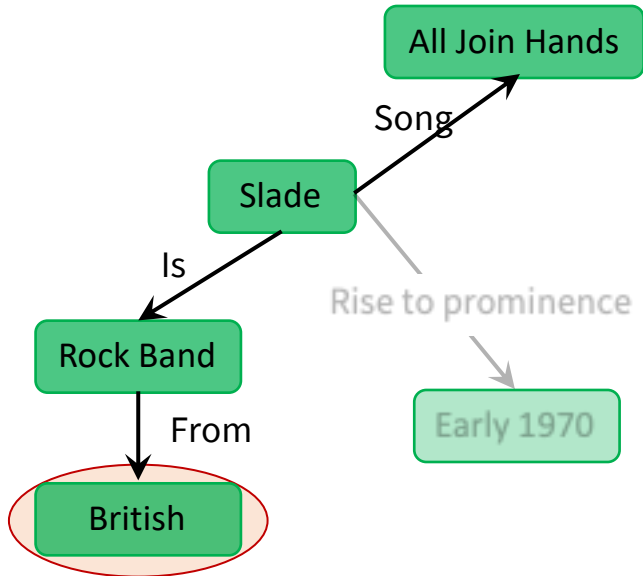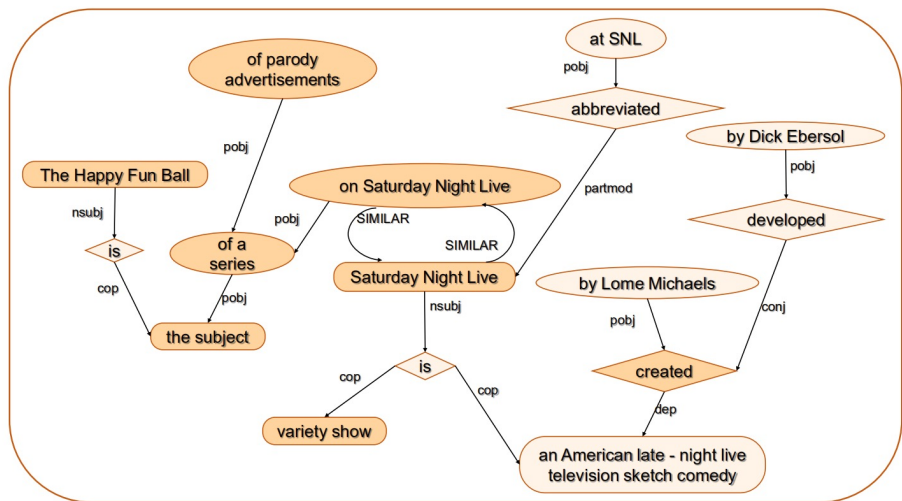**Question** The "Happy Fun Ball" was the **subject** of a **series** of **parody advertisements** on a **show created** by who?
Answer Lorne Michaels

Evidence #1 The "**Happy Fun Ball**" was the **subject** of a **series** of **parody advertisements** on "**Saturday Night Live**" .
Evidence #2 Saturday Night Live (abbreviated as SNL) is an American late - night live television sketch comedy and variety **show created** by Lorne Michaels and developed by Dick Ebersol .

1) Semantic Parsing:
- Dependency Parsing (DP)
- Semantic Role Labeling (SRL)

2) Pruning and Merging
- Based on heuristic rules
- Merge tokens within a phrase as a single node
- Filter out unimportant nodes (for example, punctuation)

3) Connect Similar Nodes
- Merge semantic graphs for different sentences into a unified graph.

Model Architecture

When did the rock band that sang
"All Join Hands" rise to prominence?
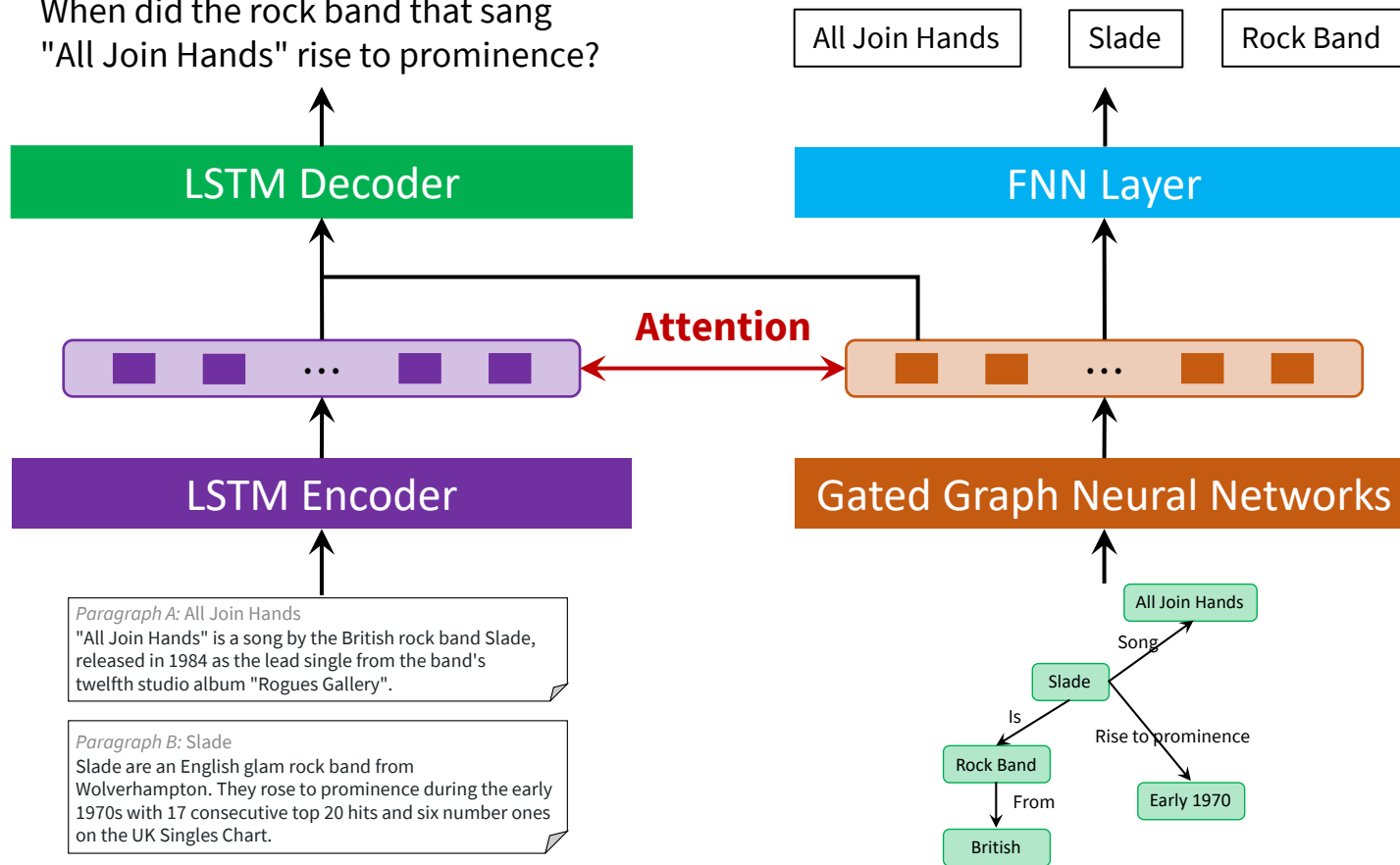
Experimental Results
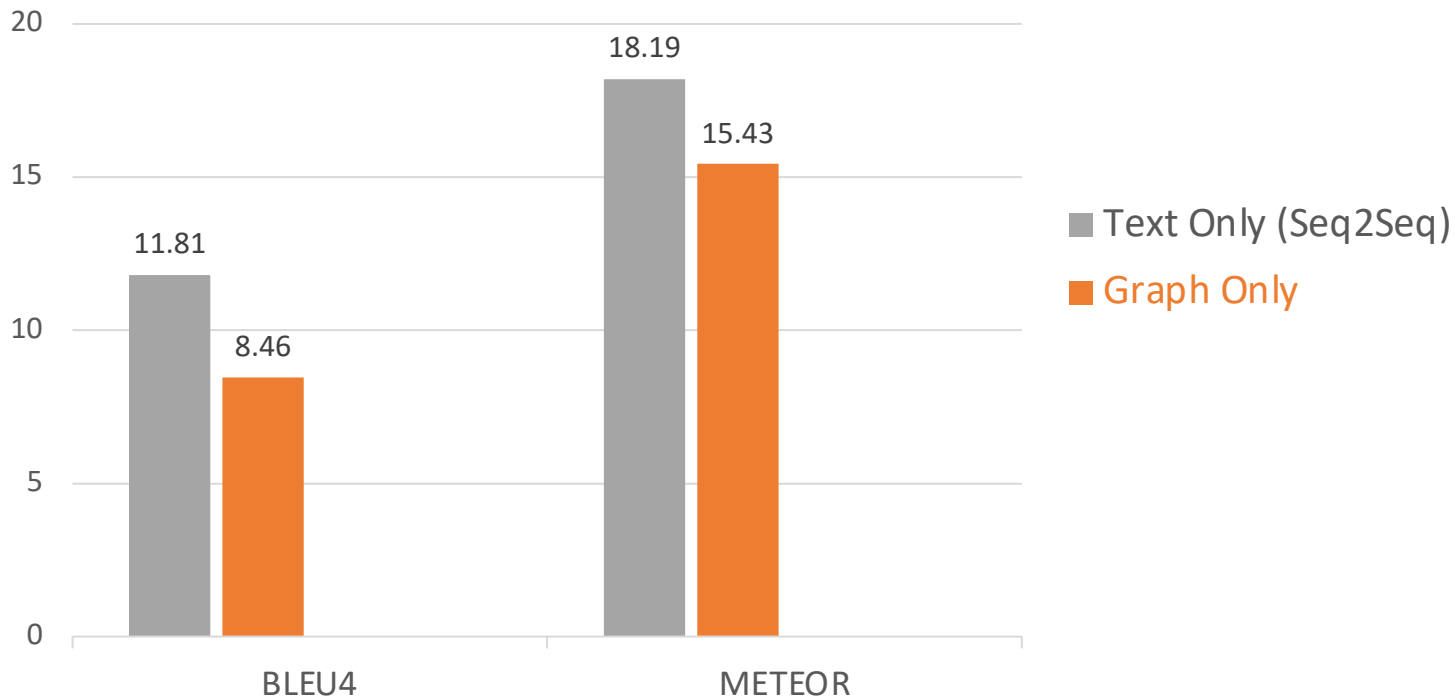
BLEU4 & METEOR on HotpotQA

Experimental Results

BLEU4 & METEOR on HotpotQA
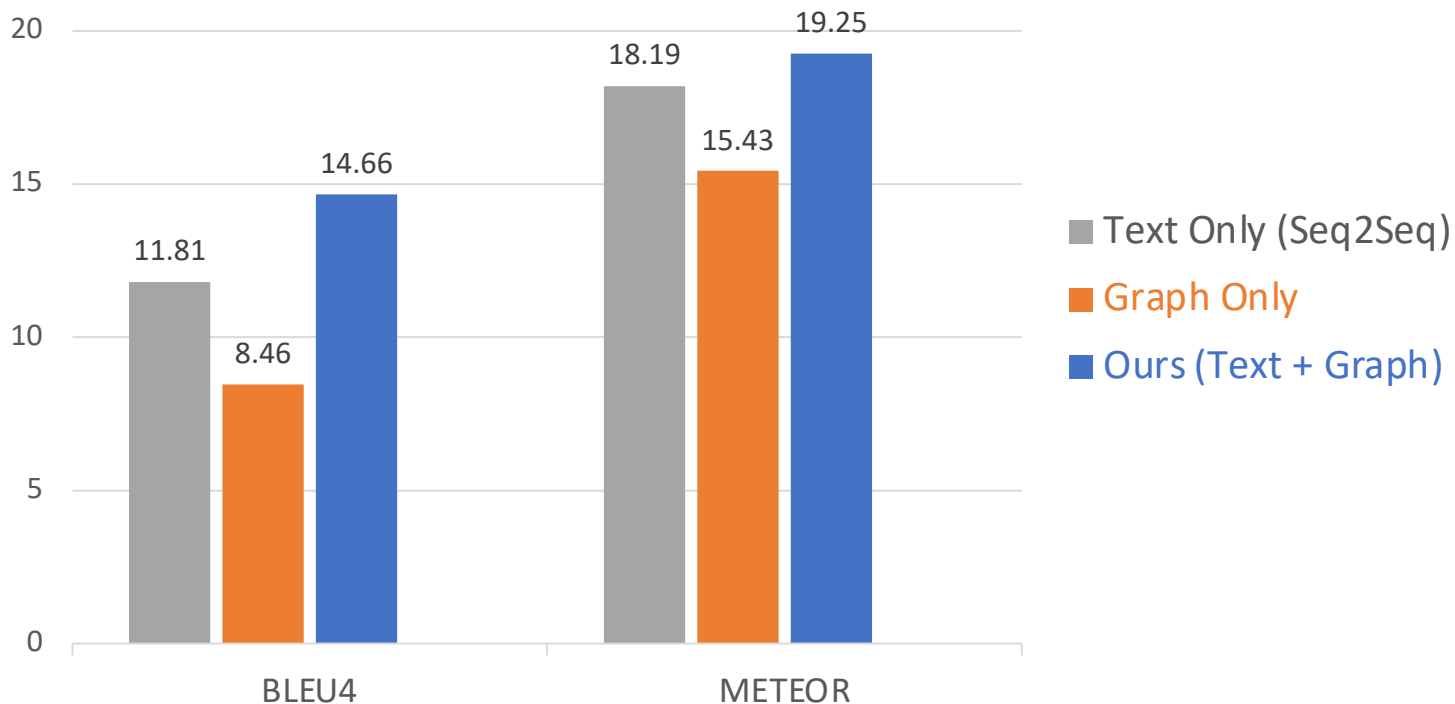
Experimental Results

BLEU4 & METEOR on HotpotQA
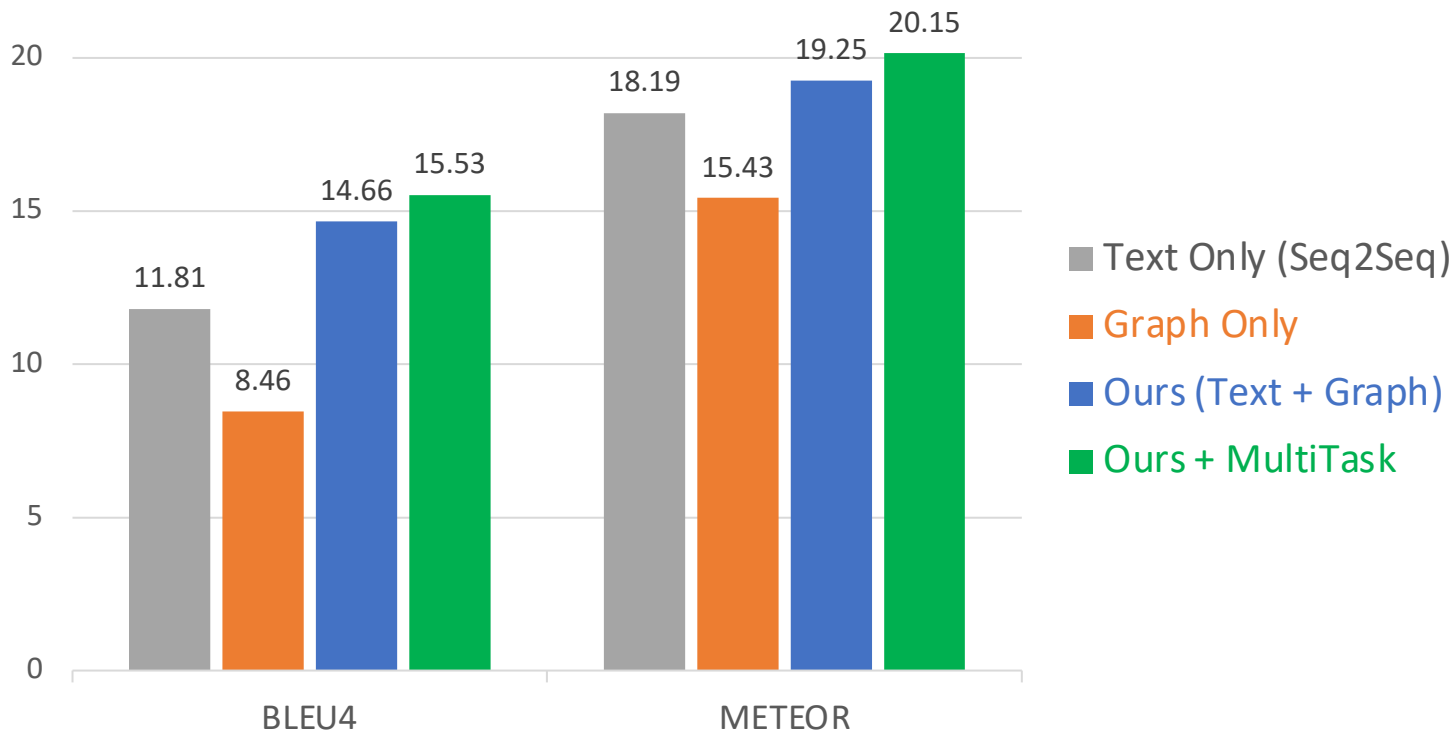
Human Evaluation

- 300 random test samples
- 3 human annotators
- Evaluation metrics: Fluency, Relevance, Complexity
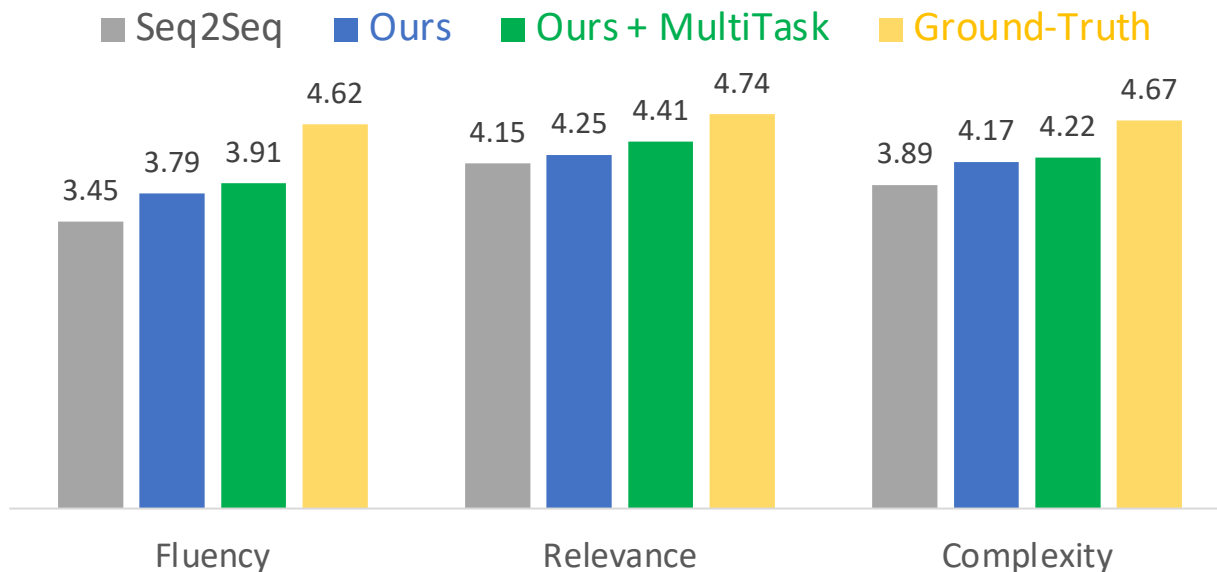
Error Analysis

❑ **Semantic Error**: the question has logic or commonsense error;

❑ **Answer Revealing**: the question reveals the answer;

❑ **Ghost Entity**: the question refers to entities that do not occur in the document;

❑ **Redundant**: the question contains unnecessary repetition;

❑ **Unanswerable**: the question does not have the above errors but cannot be answered by the document.
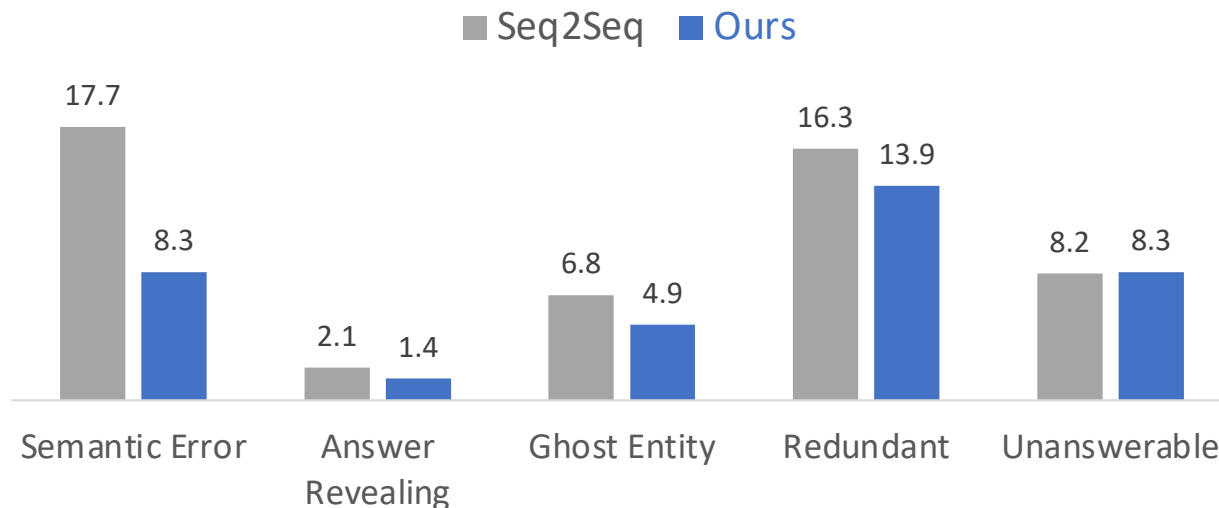
Error Analysis

- ❑ **Semantic Error**: the question has logic or commonsense error;
- ❑ **Answer Revealing**: the question reveals the answer;
- ❑ **Ghost Entity**: the question refers to entities that do not occur in the document;
- ❑ **Redundant**: the question contains unnecessary repetition;
- ❑ **Unanswerable**: the question does not have the above errors but cannot be answered by the document.
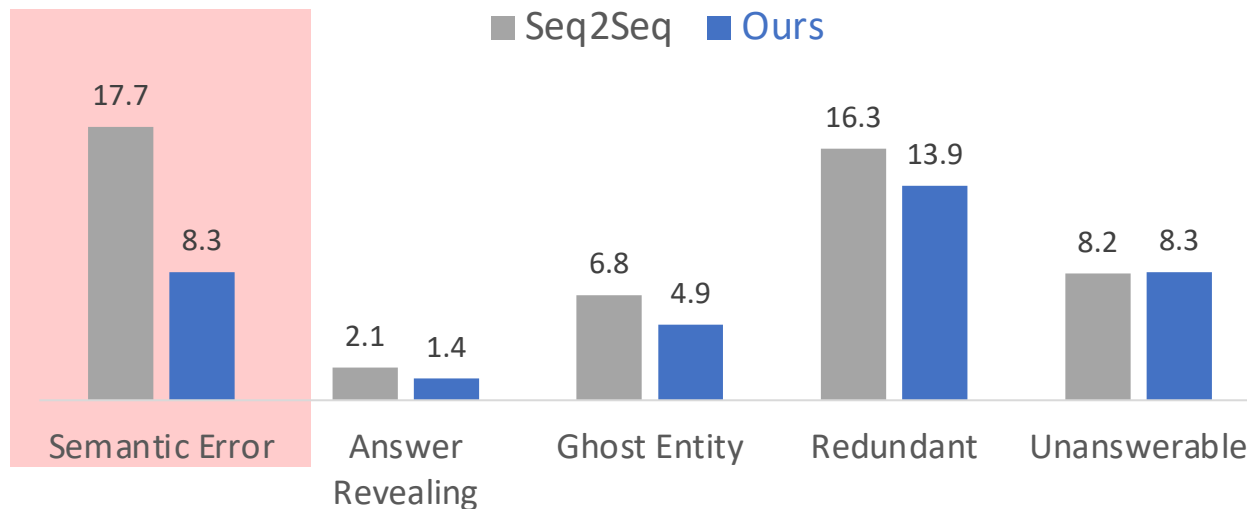
Limitations

❑ **Implicit Reasoning**

- Lack of an explicit modeling of the reasoning process.
- For example: retrieve(A); retrieve(B); Compare(A, B)

❑ **Supervised Learning**

- Requires large-scale human-written multi-hop questions.
- In practice, unsupervised question generation is more desirable.

How can we generate multi-hop questions:
1) with an explicit reasoning process
2) without seeing any human-written examples?

Core Problems (My Interests)
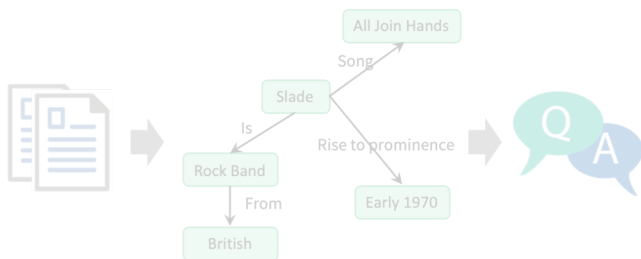
How to **generate** questions that promote deep reasoning?
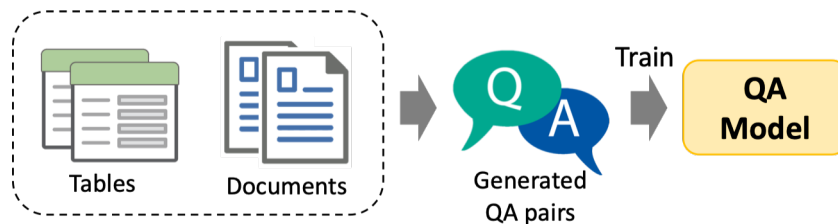
[Pan et al., ACL 2020]

[Pan et al., NAACL 2021]

How does QG **benefits** downstream NLP applications?
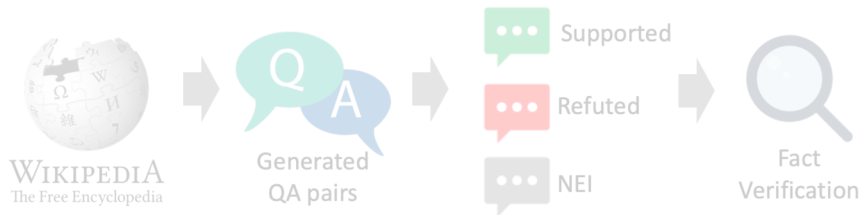
[Pan et al., ACL 2021]

Composing shallow questions into deep questions

Liangming Pan, Wenhu Chen, Wenhan Xiong, Min-Yen Kan, William Yang Wang

[NAACL 2021] Pan et al: Unsupervised Multi-hop Question Answering by Question Generation
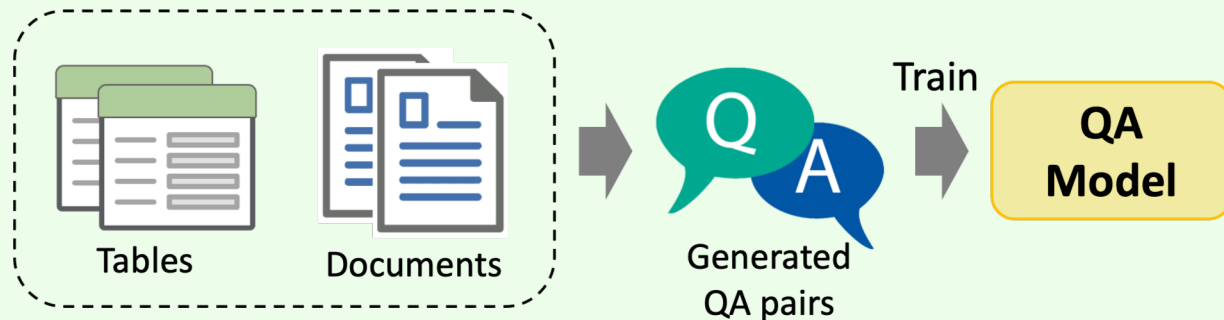
Answer Multi-hop Questions

When did the rock band that sang "All Join Hands" rise to prominence?

Multi-hop Question

Which rock band sang "All Join Hands"?

Paragraph A: All Join Hands
"All Join Hands" is a song by the British rock band Slade, released in 1984 as the lead single from the band's twelfth studio album "Rogues Gallery".

*Ans = **Slade***

When did **Slade** rise to prominence?

Paragraph B: Slade
Slade are an English glam rock band from Wolverhampton. They rose to prominence during the early 1970s with 17 consecutive top 20 hits and six number ones on the UK Singles Chart.

*Ans = **Early 1970s***

Generate Multi-hop Questions

Zero-shot Multi-hop QA

Paragraph A: All Join Hands
"All Join Hands" is a song by the British rock band Slade, released in 1984 ⋯⋯

Paragraph B: Slade
Slade are an English glam rock band from Wolverhampton. They rose to prominence during the early 1970s with ⋯⋯⋯

What rock band sang "All Join Hands"?
Answer: Slade

#1 $FindBridge$ Slade

#3 $QGwithAns$

$QGwithEnt$ #2

Slade sang "All Join Hands"? #4 $QuesToSent$

When did the rock band Slade rose to prominence?
Answer: Early 1970s

#5 $BridgeBlend$

When did the rock band that sang "All Join Hands" rise to prominence?
Answer: Early 1970s

Medal	Championship	Name	Event
Silver	2010 Pruszkow	Tim Veldt	Men's omnium
Bronze	2011 Apeldoorn	Kristen Wild	Women's omnium
Gold	2013 Apeldoorn	Elis Ligtlee	Women's keirin

Kirsten Carlijn Wild (born 15 October 1982) is a Dutch professional racing cyclist, ⋯ ⋯ ⋯. Wild competed in two track cycling events at the 2012 Summer Olympics.

Kirsten Wild **#1** *FindBridge*

#3 *QGwithEnt*

Kirsten Wild of Netherlands won the bronze medal in the 2011 Apeldoorn. **#2** *DescribeEnt*

What is the birthdate of Kirsten Wild?
Answer: 15 October 1982

#4 *BridgeBlend*

What is the birthdate of the athlete that of Netherlands won the bronze medal in the 2011 Apeldoorn?   **Answer**: 15 October 1982

37

The Blending Function

e: Kirsten Wild

q: What is the birthdate of Kirsten Wild? **Answer**: 15 October 1982

s: Kirsten Wild of Netherlands won the bronze medal in the 2011 Apeldoorn.

What is the birthdate of the _____ that of Netherlands won the bronze medal in the 2011 Apeldoorn?

What is the birthdate of the athlete that of Netherlands won the bronze medal in the 2011 Apeldoorn? **Answer**: 15 October 1982

BERT

Evaluation Datasets

❑ HotpotQA

Text + Text

Homogeneous data sources

Paragraph A: All Join Hands
"All Join Hands" is a song by the British rock band Slade, released in 1984 as the lead single from the band's twelfth studio album "Rogues Gallery".

Paragraph B: Slade
Slade are an English glam rock band from Wolverhampton. They rose to prominence during the early 1970s with 17 consecutive top 20 hits and six number ones on the UK Singles Chart.

Question #1:
When did the rock band that sang "All Join Hands" rise to prominence?

(Yang et al., EMNLP 2018)

❑ HybridQA

Table + Text

Heterogeneous data sources

2004 United States Grand Prix

Pos	No	Driver	Constructor	Time	Gap
1	2	Rubens Barrichello	Ferrari	1:10.223	—
2	1	Michael Schumacher	Ferrari	1:10.400	+0.177
3	10	Takuma Sato	BAR - Honda	1:10.601	+0.378
4	9	Jenson Button	BAR - Honda	1:10.820	+0.597
5	3	Juan Pablo Montoya	Williams - BMW	1:11.062	+0.839

Jenson Alexander Lyons Button MBE (born 19 January 1980) is a British racing driver and former Formula One driver. He won the 2009 Formula One World Championship, driving for Brawn GP.
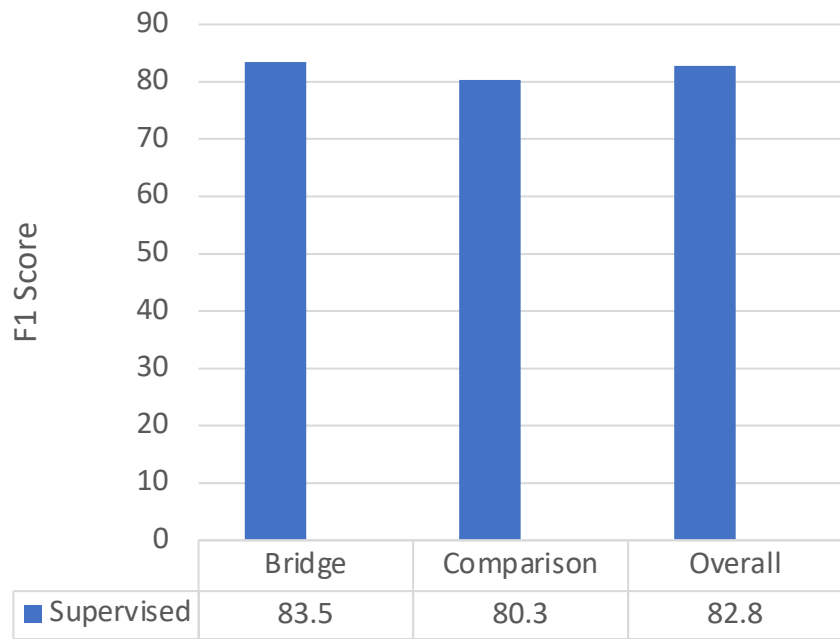
Question #2:
What nationality was the driver who finished in position 4 in the 2004 United States Grand Prix?

(Chen et al., EMNLP 2020)

Supervised QA Performance

HotpotQA

■ Supervised	Bridge	Comparison	Overall
	83.5	80.3	82.8

HybridQA

■ Supervised	In-Table	In-Text	Overall
	58.6	46.4	50

40

Zero-shot QA Performance

HotpotQA

	Bridge	Comparison	Overall
■ Supervised	83.5	80.3	82.8
■ Zero-Shot	72.2	54.4	68.6

HybridQA

	In-Table	In-Text	Overall
■ Supervised	58.6	46.4	50
■ Zero-Shot	40.6	25	30.5

Few-shot QA Performance

❑ **HotpotQA**

❑ **HybridQA**

The F1 score for progressively larger training dataset sizes for finetuning.

Examples of Generated Questions
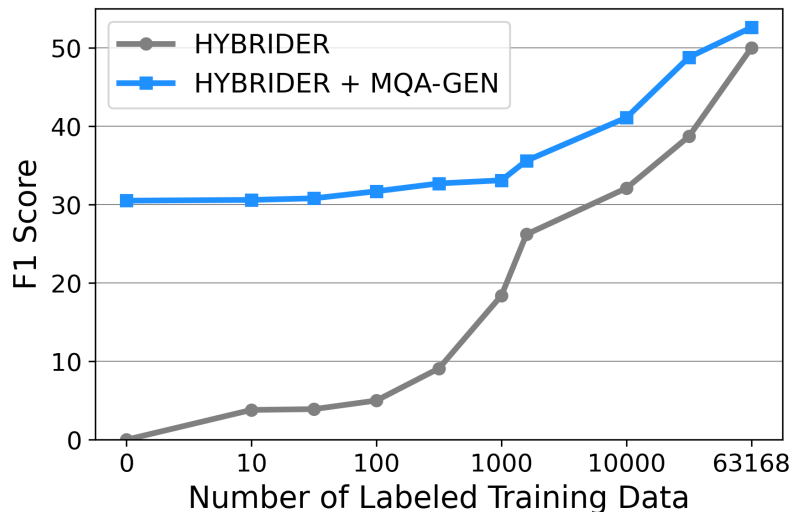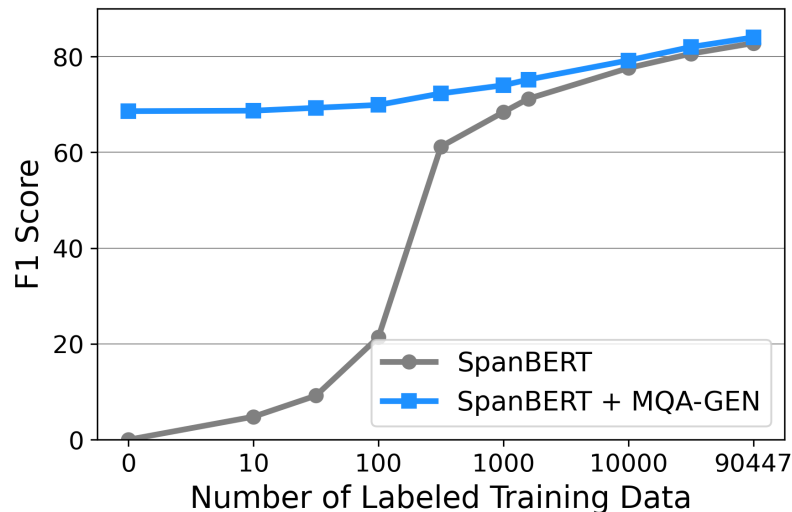
Type	Question	Answer
Table-Text	When did the one that won the Eurovision Song Contest in 1966 join Gals and Pals?	1963
	How many students attend the teams that played in the Dryden Township Conference?	1900
Text-Table	What album did the Oak Ridge Boys release in 1989?	American Dreams
	When was the name that is the name of the bridge that crosses Youngs Bay completed?	Summer
Text-Text	Which Canadian cinematographer is best known for his work on Fargo?	Craig Wrobleski
	What is illegal in the country that is Bashar Hafez al - Assad 's father?	Cannabis
Comparison	Who was born first, Terry Southern or Neal Town Stephenson?	Terry Southern
	Are Beth Ditto and Mary Beth Patterson of the same nationality?	Yes

Core Problems (My Interests)
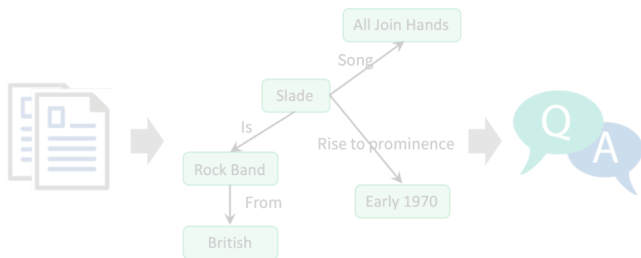
How to **generate** questions that promote deep reasoning?

[Pan et al., ACL 2020]

[Pan et al., NAACL 2021]

How does QG **benefits** downstream NLP applications?

[Pan et al., ACL 2021]

Question Generation for Fact Verification

Liangming Pan, Wenhu Chen, Wenhan Xiong, Min-Yen Kan, William Yang Wang

[ACL 2021] Pan et al: Few-shot Fact Checking with Claim Generation

Fact Checking

Fact Checking

"Immigrants are a drain on the economy"
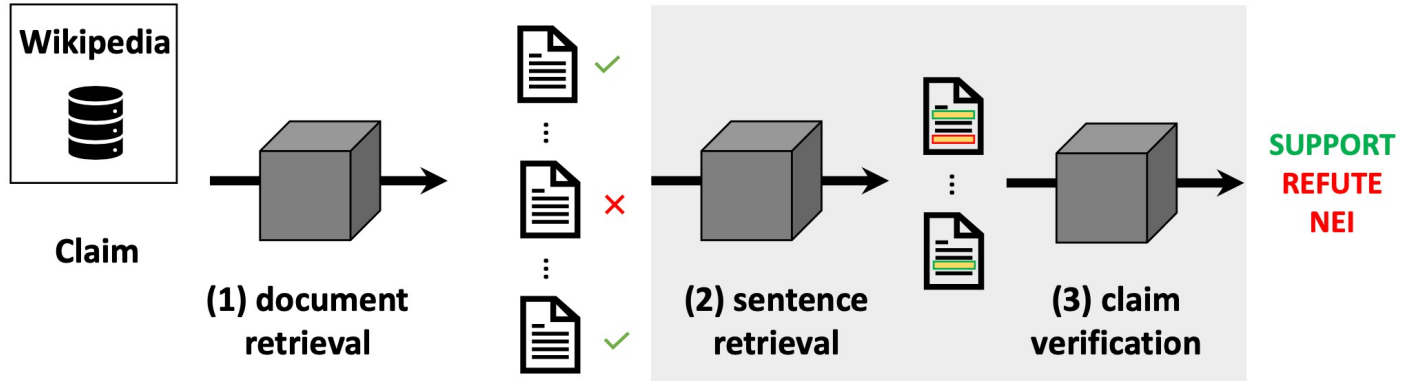
- ❖ Labels: Supports, Refutes, Not Enough Info
- ❖ Pipeline
    - ❖ Document Retrieval
    - ❖ Sentence Retrieval
    - ❖ Claim Verification

Fact Verification

Claim: The Rodney King riots took place in the most populous county in the USA. — Claim

[wiki/Los_Angeles_Riots]
The 1992 Los Angeles riots, also known as the Rodney King riots were a series of riots, lootings, arsons, and civil disturbances that occurred in Los Angeles County, California in April and May 1992.

[wiki/Los_Angeles_County]
Los Angeles County, officially the County of Los Angeles, is the most populous county in the USA. — Evidence

Verdict: Supported — Label

FEVER: a large-scale dataset for Fact Extraction and VERification

James Thorne[1], **Andreas Vlachos**[1], **Christos Christodoulopoulos**[2], and **Arpit Mittal**[2]

[1]Department of Computer Science, University of Sheffield
[2]Amazon Research Cambridge
{j.thorne, a.vlachos}@sheffield.ac.uk
{chrchrs, mitarpit}@amazon.co.uk

Split	SUPPORTED	REFUTED	NEI
Train	80,035	29,775	35,639
Dev	6,666	6,666	6,666
Test	6,666	6,666	6,666

48

Motivation

Training fact verification models require large scale human-annotated (Evidence, Claim, Label) data, such as FEVER.

- It is time-consuming to ask humans to annotate the training data.

- Could we automatically generate (Evidence, Claim, Label) to train the fact verification model?

Zero/Few-shot Fact Verification
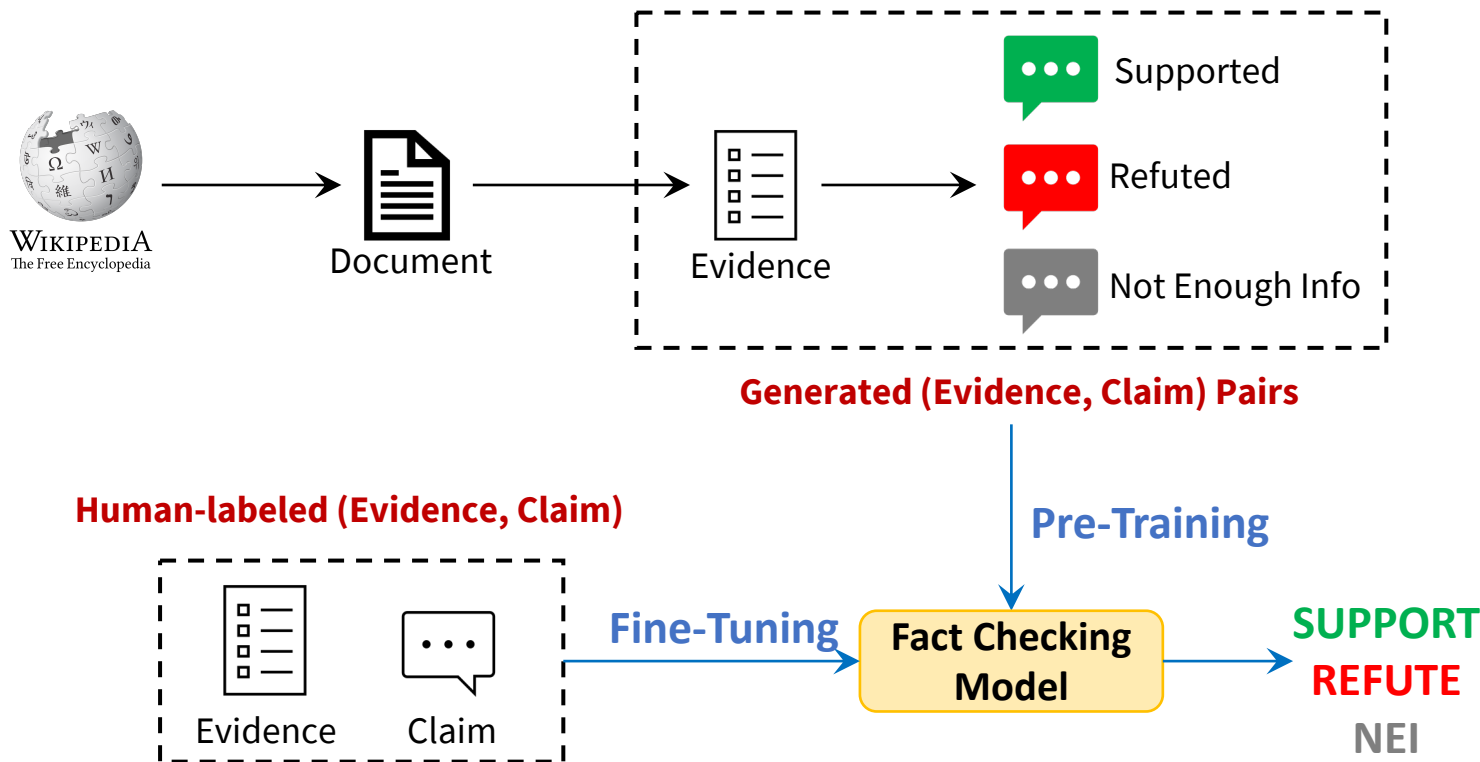
Claim Generation with QG

Examples of Generated Claims

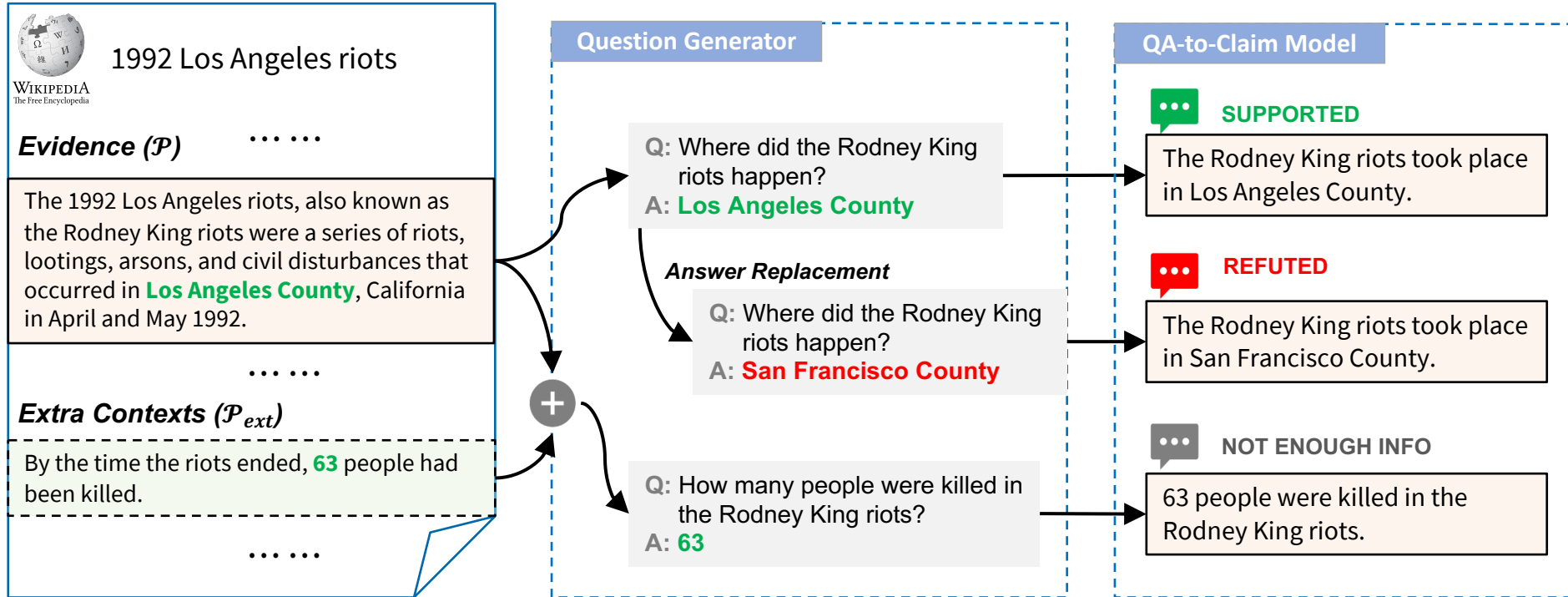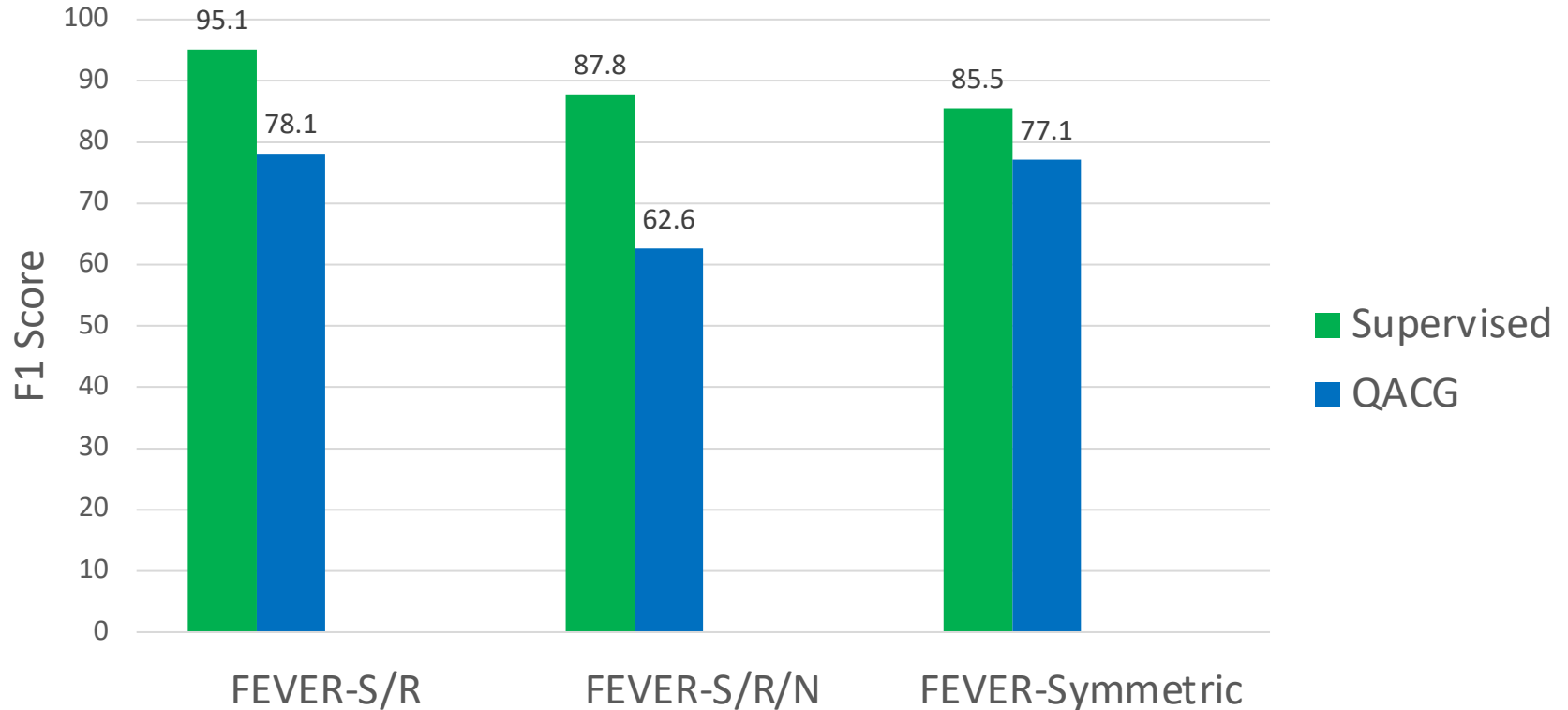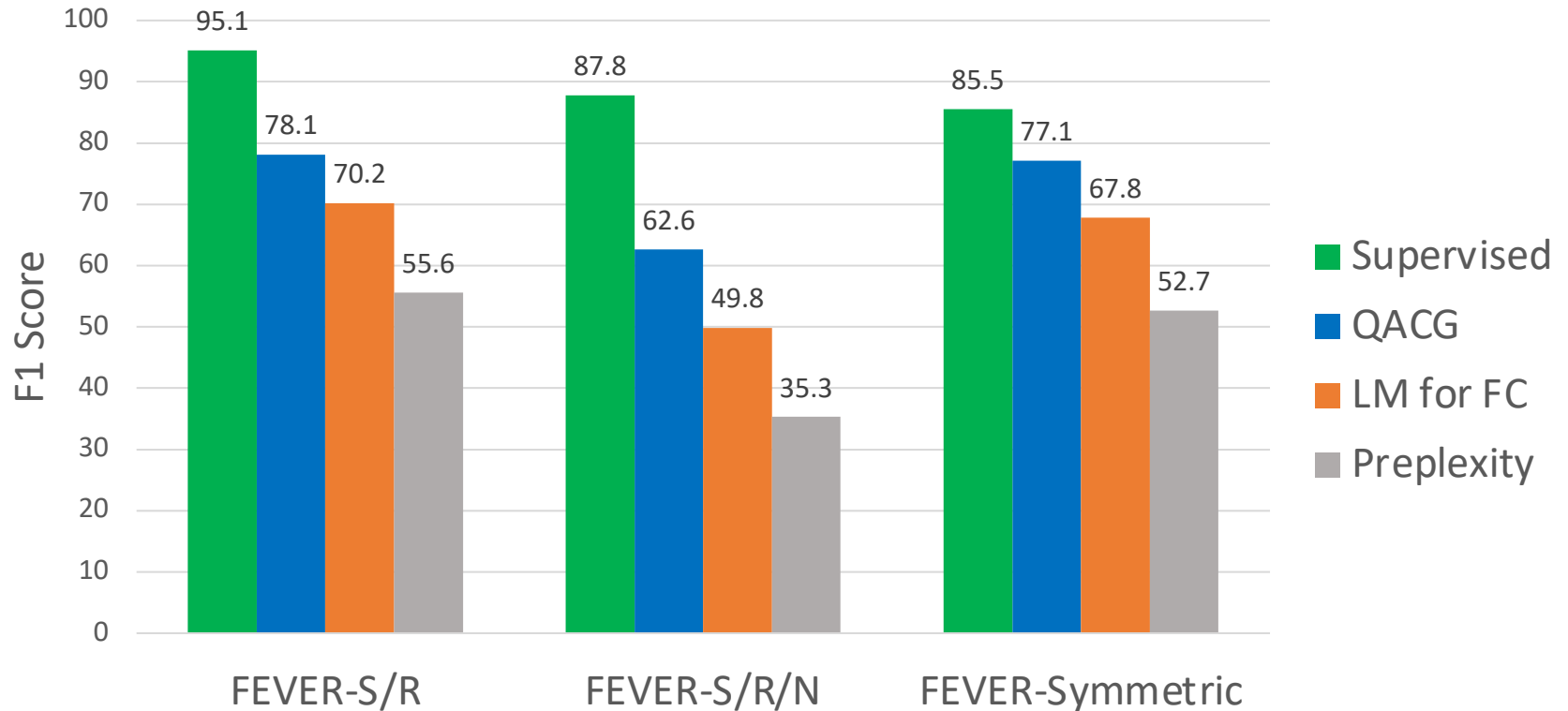| Evidence	Generated Claim
Budapest is cited as one of the most beautiful cities in Europe, ranked as the most liveable Central and Eastern European city on EIU's quality of life index, ranked as "the world's second best city" by Conde Nast Traveler, and "Europe's 7th most idyllic place to live" by Forbes.	**SUPPORTED claims** Budapest is ranked as the most liveable city in central Europe. Budapest ranks 7th in terms of idyllic places to live in Europe. **REFUTED claims** Budapest ranks in 11th in terms of idyllic places to live in Europe. Budapest is ranked the most liveable city in Asia. **NEI claims** Budapest is one of the largest cities in the European Union. Budapest is the capital of Hungary.
Alia Bhatt received critical acclaim for portraying emotionally intense characters in the road drama Highway (2014), which won her the Filmfare Critics Award for Best Actress, and the crime drama Udta Punjab (2016), which won her the Filmfare Award for Best Actress.	**SUPPORTED claims** Bhatt won the Filmfare Award for Best Actress in Udta Punjab. Bhatt received the Filmfare Critics Award for her role in Highway. **REFUTED claims** Alia Bhatt won the Best Original Screenplay award in Highway. 2 States (2014) won Alia Bhatt the Filmfare Award for Best Actress. **NEI claims** Alia Bhatt made her acting debut in the 1999 thriller Sangharsh. Bhatt played her first leading role in Karan Johar's romantic drama.

Zero-shot Fact Verification

F1 Score

	FEVER-S/R	FEVER-S/R/N	FEVER-Symmetric
Supervised	95.1	87.8	85.5
QACG	78.1	62.6	77.1

Zero-shot Fact Verification

Bar chart titled "Zero-shot Fact Verification" showing F1 Score on the y-axis (0–100) across three datasets: FEVER-S/R, FEVER-S/R/N, and FEVER-Symmetric. Four methods compared: Supervised (green), QACG (blue), LM for FC (orange), Preplexity (gray).

- FEVER-S/R: Supervised 95.1, QACG 78.1, LM for FC 70.2, Preplexity 55.6
- FEVER-S/R/N: Supervised 87.8, QACG 62.6, LM for FC 49.8, Preplexity 35.3
- FEVER-Symmetric: Supervised 85.5, QACG 77.1, LM for FC 67.8, Preplexity 52.7

Few-shot Fact Verification

Discussions

❖ Our model requires a good question generator.

❖ To generate deep claims, you need to generate deep questions.

| Model	Answerable Rate	FV Performance P / R / F_1
NQG++	63.0%	62.2 / 62.4 / 62.3
BART	89.5%	76.3 / 76.0 / 76.1

Table 2: *Answerable Rate*: the ratio of answerable questions generated by the NQG++ and the BART model. *FV Performance*: the zero-shot fact verification performance on the FEVER-Symmetric.

|	
Evidence:	Roman Atwood is best known for his vlogs, where he posts updates about his life.
Claim:	Roman Atwood is a content creator.
Evidence:	In 2004, Slovenia entered NATO and the European Union.
Claim:	Slovenia uses the euro.
Evidence:	He has traveled to Chad and Uganda to raise awareness about conflicts in the regions.
Claim:	Ryan Gosling has been to a country in Africa.

Table 5: Examples of claims in FEVER that require commonsense or world knowledge (underlined).
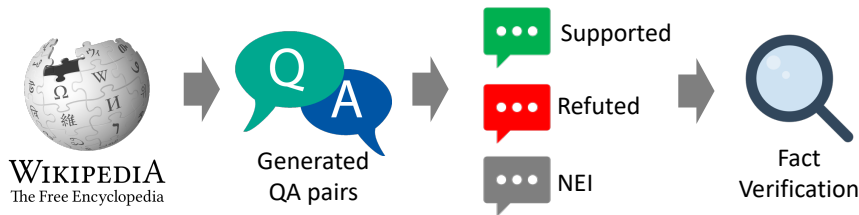
Summary
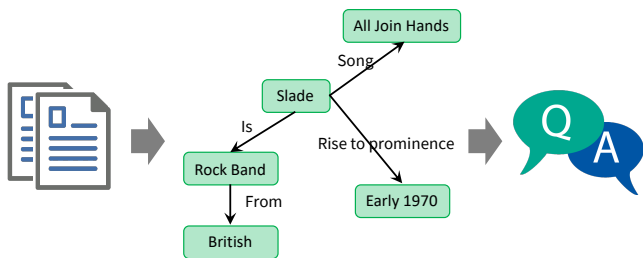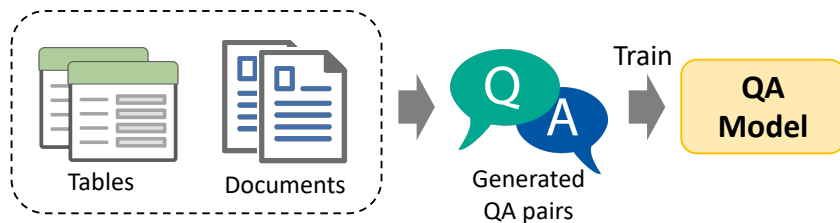
[Pan et al., ACL 2020]

[Pan et al., NAACL 2021]

How does QG **benefits** downstream NLP applications?

[Pan et al., ACL 2021]

Summary

We are still far from real human-raised deep questions

Science & Technology
Is there a center of the universe?

04:14

545,587 views

Most of the astronomers, scientists, and philosophers had ideas about the universe that we now know where not entirely correct. **In what ways do you think ideas that have turned out to be "wrong" are valuable in developing our understanding of the universe?**

Understanding, Reasoning   Prior Knowledge   Information-Seeking

63

Future Direction: Prior Knowledge

Photosynthesis in green plants converts water, carbon dioxide, and minerals into oxygen and energy-rich organic compounds.

Human takes in oxygen and breath out carbon dioxide to maintain life.

Sunlight is important for plants to maintain life.

Why do humans take O2 to produce CO2 while plants do the opposite?

Does sunlight take any part in the photosynthesis?

Future Direction: Information-Seeking

I am sitting in a restaurant with a friend. He orders: 'Tournedos, medium rare ...' I order spaghetti. Soon my friend is served baked chicken with French fries and I am served spaghetti; we are both happy with our dishes.

Now, suppose I would like to understand what really happened.

What question should I ask?

- Didn't you order tournedos?
- You really don't mind that they changed your order?
- Oh, have you changed your order in the meantime?
- Why are you in the restaurant?
- Is the spaghetti delicious?

Know more about QG

Survey for Neural Question Generation

- https://arxiv.org/pdf/1905.08949.pdf

Recent Advances in Neural Question Generation

Liangming Pan, Wenqiang Lei, Tat-Seng Chua and Min-Yen Kan
School of Computing
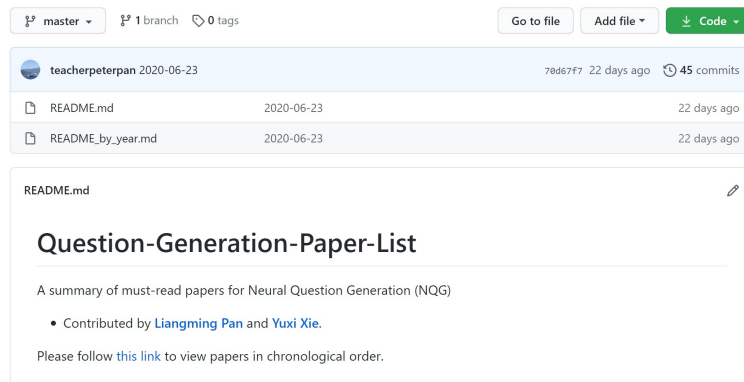National University of Singapore
Singapore 117417
{e0272310,wenql,kanmy,chuats}comp.nus.edu.sg

Question Generation Paper List

- https://github.com/teacherpeterpan/Question-Generation-Paper-List

Collaborators

Yuxi Xie
xieyuxi@u.nus.edu
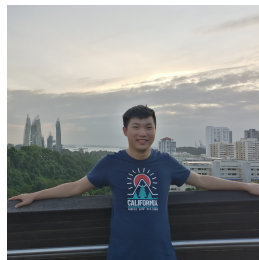
Liangming Pan
e0272310@u.nus.edu
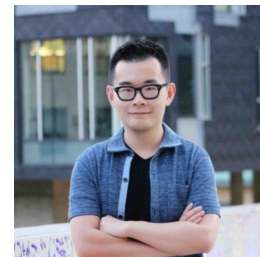
Min-Yen Kan
kanmy@comp.nus.edu.sg

Wenhu Chen
wenhuchen@ucsb.edu

Wenhan Xiong
xwhan@cs.ucsb.edu

William Wang
william@cs.ucsb.edu

References

[1] *Knowledge Questions from Knowledge Graphs*. Dominic Seyler, Mohamed Yahya, Klaus Berberich. ICTIR 2017.

[2] *Interpretation of Natural Language Rules in Conversational Machine Reading*. Marzieh Saeidi, Max Bartolo, Patrick Lewis, Sameer Singh, Tim Rocktäschel, Mike Sheldon, Guillaume Bouchard, Sebastian Riedel. EMNLP 2018.

[3] *Question Generation for Question Answering*. Nan Duan, Duyu Tang, Peng Chen, Ming Zhou. EMNLP 2017.

[4] *Unsupervised Question Answering by Cloze Translation*. Patrick Lewis, Ludovic Denoyer, Sebastian Riedel. ACL 2019.

[5] *Training Question Answering Models From Synthetic Data*. Raul Puri, Ryan Spring, Mostofa Patwary, Mohammad Shoeybi, Bryan Catanzaro. EMNLP 2020.

[6] *Generating Question-Answer Hierarchies*. Kalpesh Krishna and Mohit Iyyer. ACL 2019.

[7] *Recent Advances in Neural Question Generation*. Liangming Pan, Wenqiang Lei, Tat-Seng Chua, Min-Yen Kan. arXiv 2019.

[8] *SQuAD: 100,000+ Questions for Machine Comprehension of Text*. Pranav Rajpurkar, Jian Zhang, Konstantin Lopyrev, Percy Liang. EMNLP 2016.

References

[9] *Semantic Graphs for Generating Deep Questions*. Liangming Pan, Yuxi Xie, Yansong Feng, Tat-Seng Chua, Min-Yen Kan. ACL 2020.

[10] *HotpotQA: A Dataset for Diverse, Explainable Multi-hop Question Answering*. Zhilin Yang, Peng Qi, Saizheng Zhang, Yoshua Bengio, William W. Cohen, Ruslan Salakhutdinov, Christopher D. Manning. EMNLP 2018.

[11] *Exploring Question-Specific Rewards for Generating Deep Questions*. Yuxi Xie, Liangming Pan, Dongzhe Wang, Min-Yen Kan, Yansong Feng. COLING 2020.

[12] *HybridQA: A Dataset of Multi-Hop Question Answering over Tabular and Textual Data*. Wenhu Chen, Hanwen Zha, Zhiyu Chen, Wenhan Xiong, Hong Wang, William Wang. ENNLP 2020.

[13] *Unsupervised Multi-hop Question Answering by Question Generation*. Liangming Pan, Wenhu Chen, Wenhan Xiong, Min-Yen Kan, William Yang Wang. arXiv 2020.

Thanks!

Any questions?

Liangming Pan
Email: liangmingpan@u.nus.edu