

WikiQA: A Challenge Dataset for Open-Domain Question Answering

Yi Yang^{*}, **Scott Wen-tau Yih[#]**, Christopher
Meek[#]

^{*}Georgia Institute of Technology, Atlanta

[#]Microsoft Research, Redmond

Open-domain Question Answering

“What are the names of Obama’s daughters?”

- Question Answering with Knowledge Base



- Large-scale knowledge bases (e.g., Freebase)

- KB-specific semantic parsing (e.g., [Berant+ 13], [Yih+ 15])



- Question Answering with Free Text

- High-quality, reliable text (e.g., Wikipedia, news articles)



WIKIPEDIA
The Free Encyclopedia

The New York Times

Answer Sentence Selection

- Given a factoid question, find the sentence in the candidate set that
 - Contains the answer
 - Can sufficiently support the answer

Q: Who won the best actor Oscar in 1973?

S₁: Jack Lemmon was awarded the Best Actor Oscar for Save the Tiger (1973).

S₂: Academy award winner Kevin Spacey said that Jack Lemmon is remembered as always making time for others.

QASent Dataset [Wang+ 07]

- Based on the TREC-QA data
- Standard benchmark for answer sentence selection
 - Dependency tree matching (e.g., [Wang+ 07], [Wang+ 10], [Yao+ 13])
 - Tree kernel SVMs (e.g., [Severyn & Moschitti 13])
 - Latent word alignment with lexical semantic matching (e.g., [Yih+ 13])
 - Deep neural network models (e.g., [Yu+ 14], [Wang+ 15])

Issues with QASent Dataset

- Question distribution

- Contain questions from human editors



- Candidate selection bias

- Output from the systems participating in TREC-QA
- Sentences need to share non-stopwords from the questions

Q: How did Seminole war end?

A: Ultimately, the Spanish Crown ceded the colony to United States rule.

- Excluding questions that have no correct answers

WikiQA Dataset

- Question distribution

- Questions sampled from Bing query logs



- Candidate selection

- Candidate sentences are from summary paragraphs of Wikipedia pages

- **Including** questions that have no correct answers

- *Answer Triggering*: detecting whether a correct answer exists in candidate sentences

Outline

- Introduction
- **WikiQA Dataset**
 - Data construction and annotation
 - Data statistics
- Experiments
- Conclusion

Data Construction – Questions

- Questions sampled from Bing query logs
 - Search queries starting with a WH-word
 - Filter out entity queries (e.g., “how I met your mother”)
 - Select queries issued by at least 5 users and have clicks to Wikipedia



Who wrote second Corinthians 

Web Images Videos Explore

1 210 000 RESULTS Narrow by language ▼ Narrow by region ▼

[Second Epistle to the Corinthians - Wikipedia, the free ...](#)
en.wikipedia.org/wiki/Second_Epistle_to_the_Corinthians ▼

The **Second Epistle to the Corinthians**, often referred to as **Second Corinthians** (and written as **2 Corinthians**), is the eighth book of the New Testament of the Bible.

[Composition](#) · [Structure](#) · [Background](#) · [Content](#) · [Uniqueness](#) · [Scholars](#)

Data Construction – Sentences

- Candidate sentences
 - Use all sentences from the summary paragraph of the Wikipedia page



“Who wrote second Corinthians?”

Second Epistle to the Corinthians

From Wikipedia, the free encyclopedia

The **Second Epistle to the Corinthians**, often referred to as **Second Corinthians** (and written as *2 Corinthians*), is the eighth book of the [New Testament](#) of the Bible. [Paul the Apostle](#) and "Timothy our brother" wrote this epistle to "the church of God which is at [Corinth](#), with all the saints which are in all [Achaia](#)".^[2Cor.1:1]

Sentence Annotation by Crowdsourcing

- *Step 1. Does the short paragraph answer the question?*

Question: Who wrote second Corinthians?

Second Epistle to the Corinthians The Second Epistle to the Corinthians, often referred to as Second Corinthians (and written as 2 Corinthians), is the eighth book of the New Testament of the Bible. Paul the Apostle and “Timothy our brother” wrote this epistle to “the church of God which is at Corinth, with all the saints which are in all Achaia”.

Sentence Annotation by Crowdsourcing

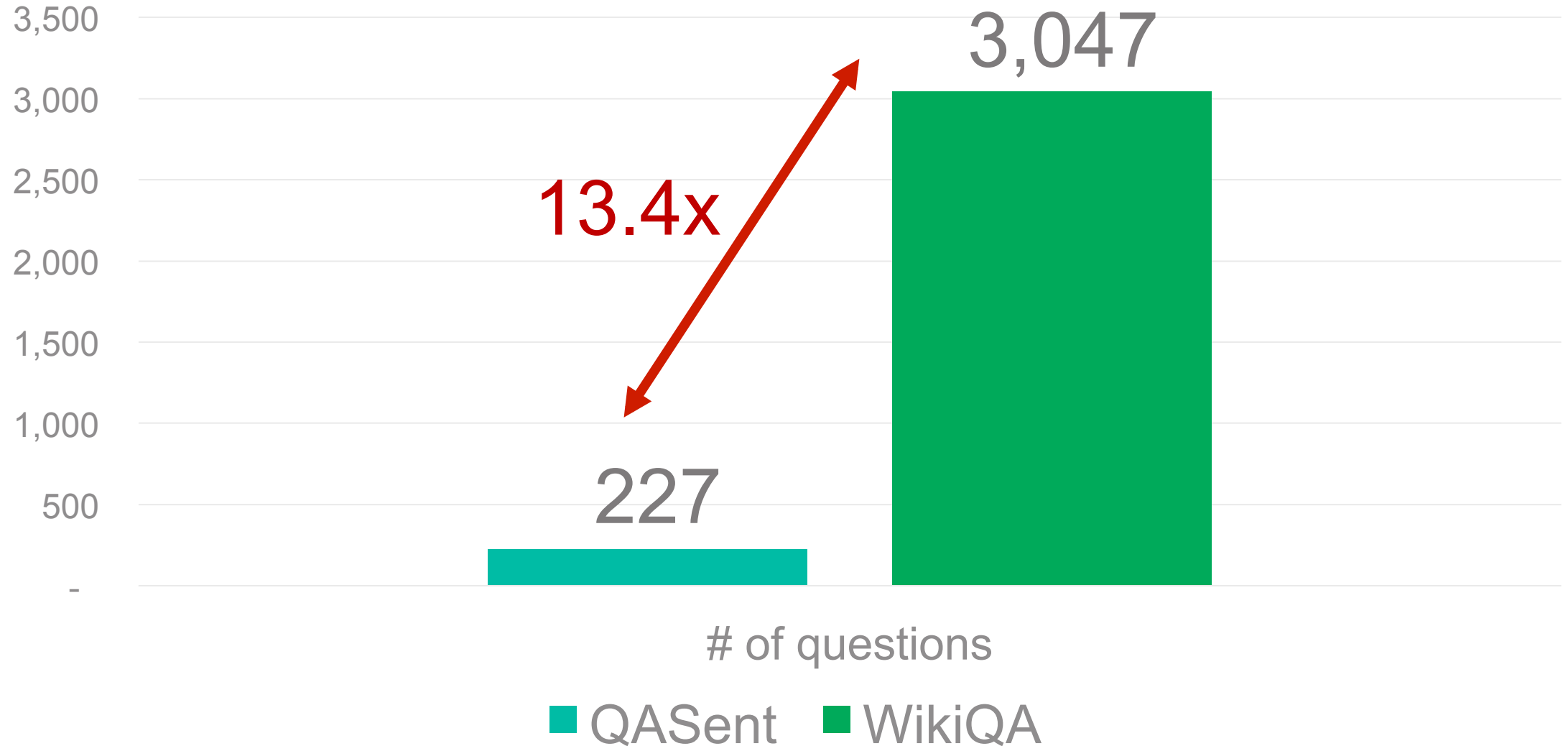
- *Step 2. Check all the sentences that can answer the question **in isolation** (assuming coreferences and pronouns have been resolved correctly)*

Question: Who wrote second Corinthians?

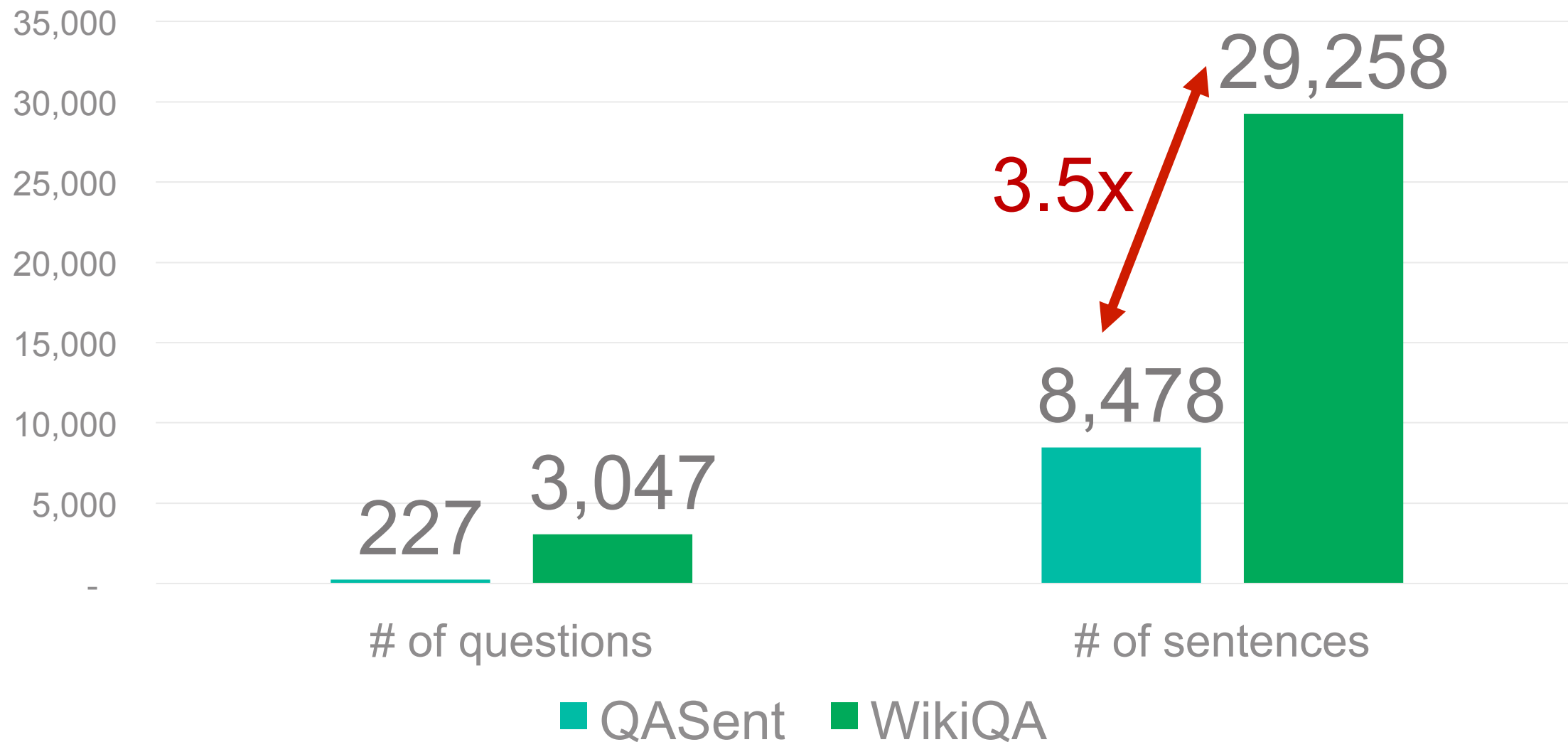
Second Epistle to the Corinthians

- The Second Epistle to the Corinthians, often referred to as Second Corinthians (and written as 2 Corinthians), is the eighth book of the New Testament of the Bible.
- Paul the Apostle and “Timothy our brother” wrote this epistle to “the church of God which is at Corinth, with all the saints which are in all Achaia”.

Data Statistics: # Questions

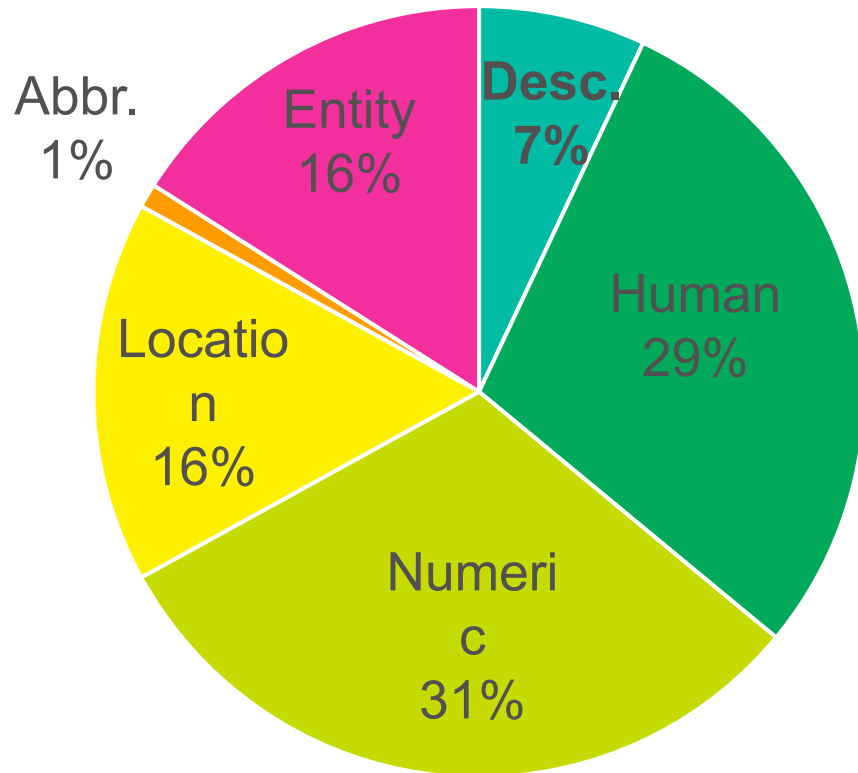


Data Statistics: # Questions & Sentences

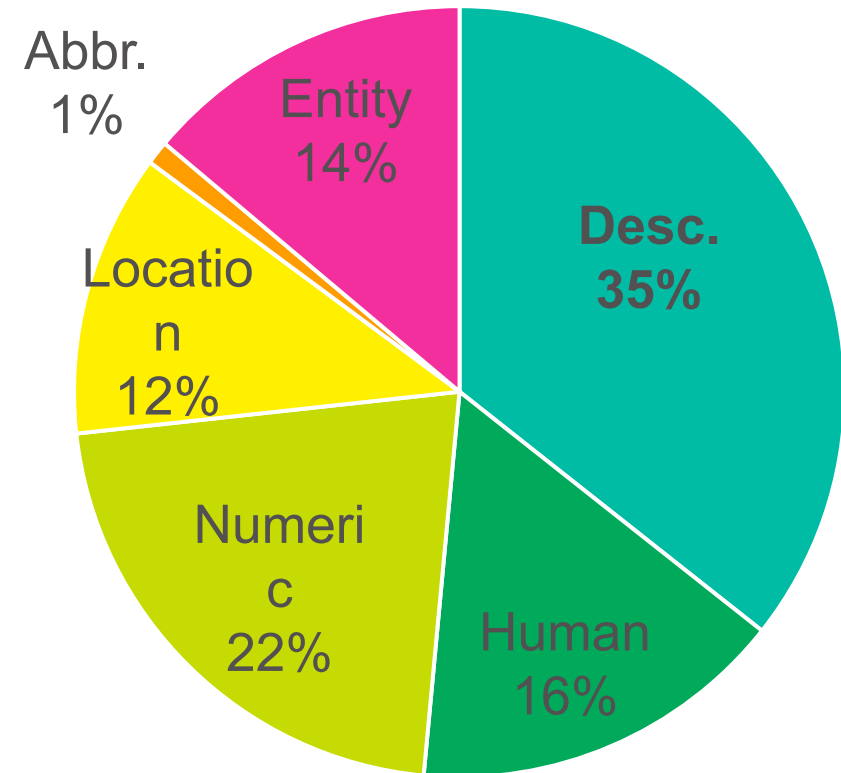


Question Classes (UIUC Question Taxonomy)

QASent



WikiQA



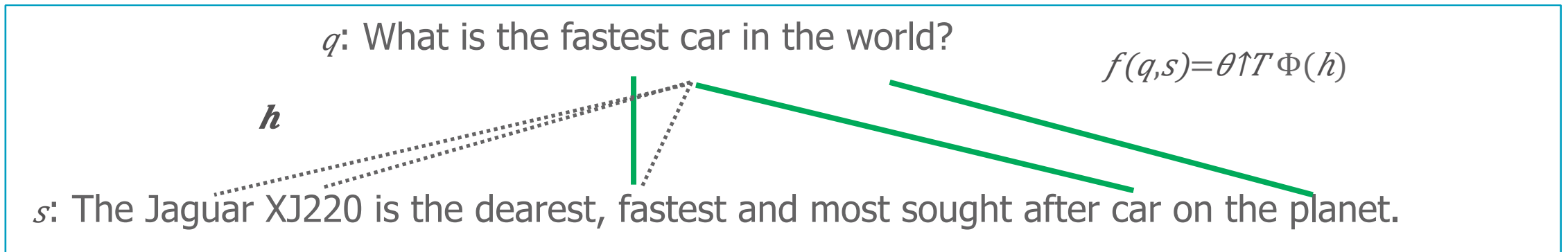
- Category “Description/Definition” has the most questions

Outline

- Introduction
- WikiQA Dataset
- **Experiments**
 - Baseline systems
 - Evaluation on answer sentence selection
 - Evaluation metric & results on answer triggering
- Conclusion

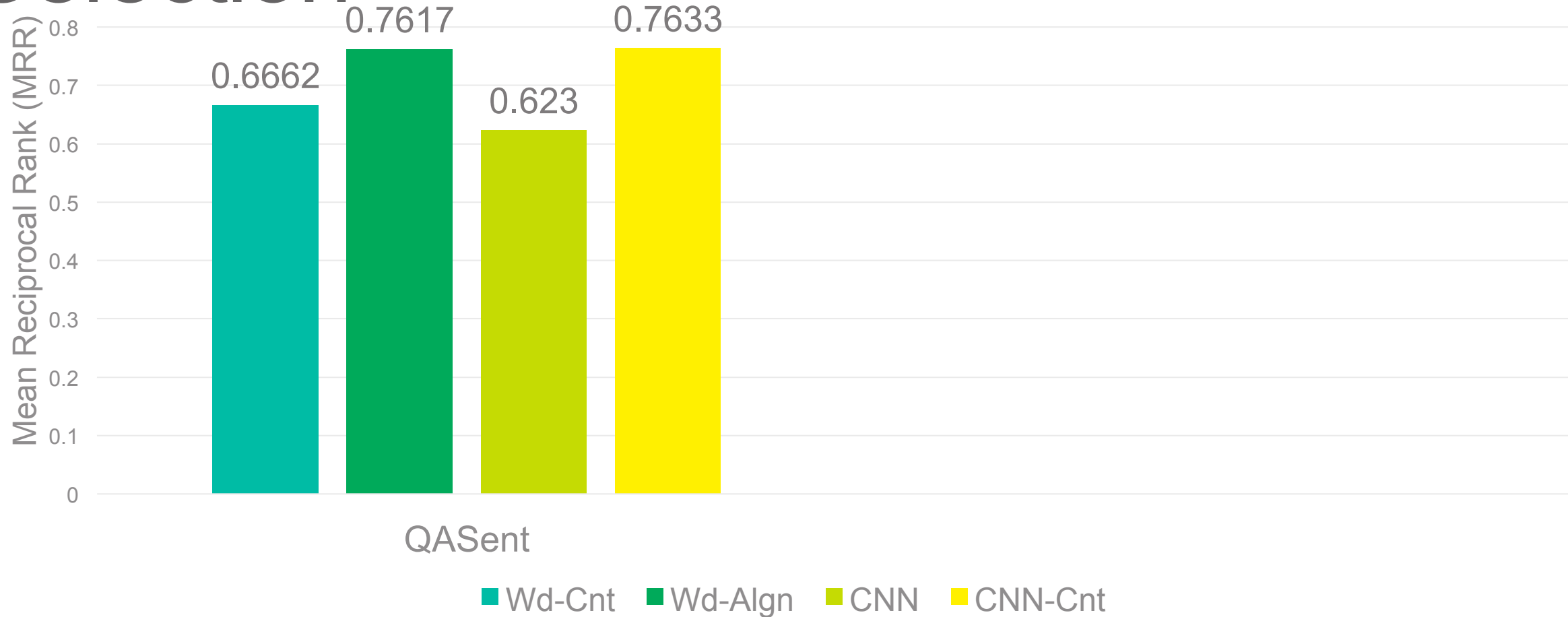
Baseline Systems

- Word matching count (Wd-Cnt)
 - # non-stopwords in Q that also occur in S
- Latent word alignment (Wd-Algn) [Yih+ ACL-13]

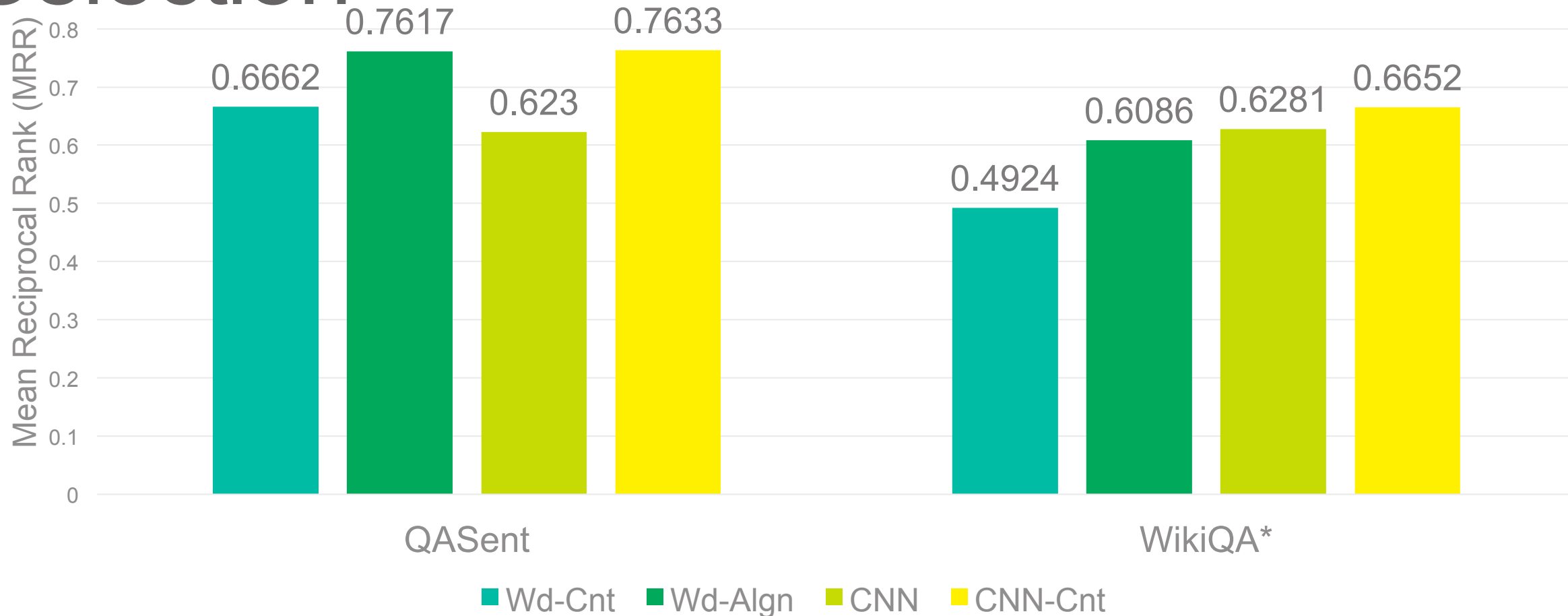


- Convolutional NN (CNN) [Yu+ DLWorkshop-14]
- Convolutional NN & Wd-Cnt (CNN-Cnt)

Evaluation on Answer Sentence Selection



Evaluation on Answer Sentence Selection

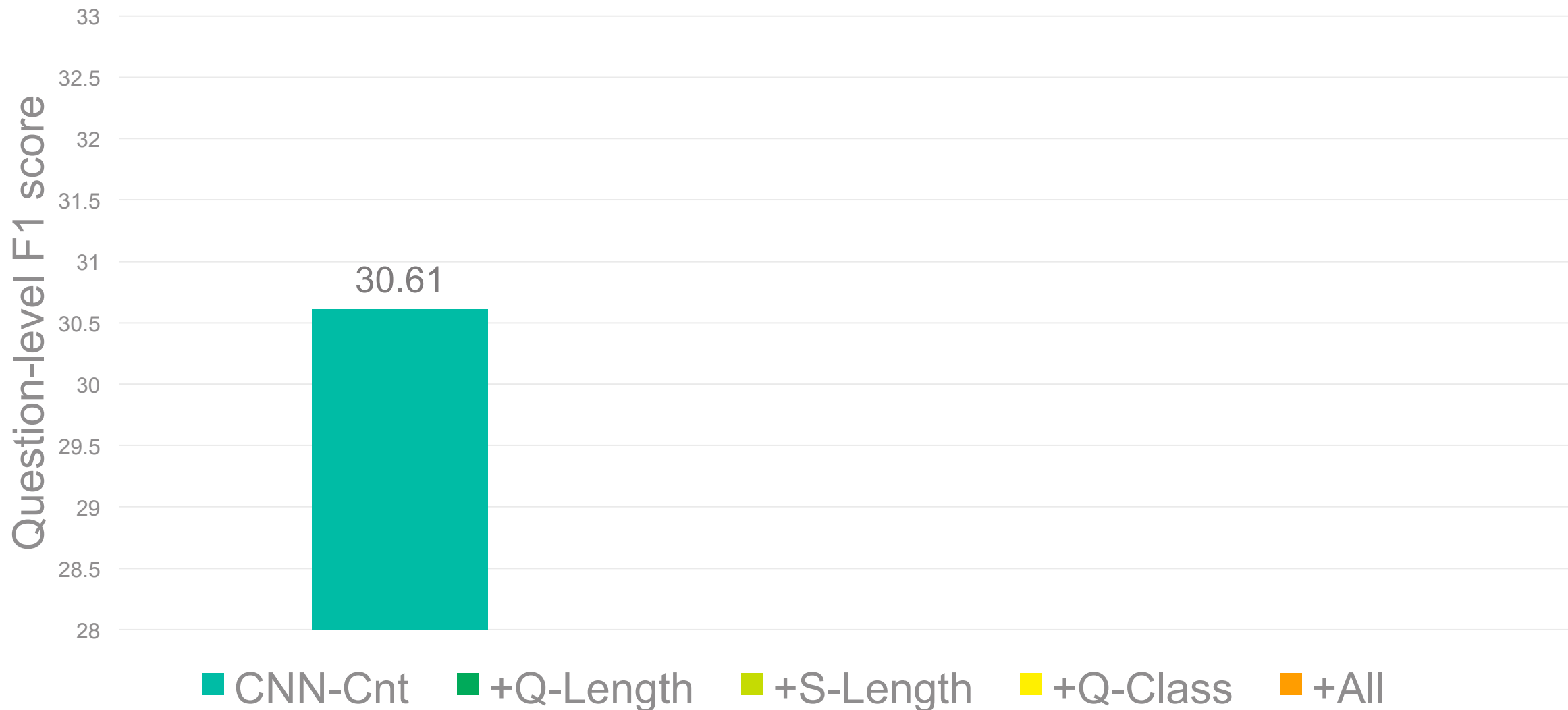


* 1,242 (40.8%) questions that have correct answer sentences in the candidate set

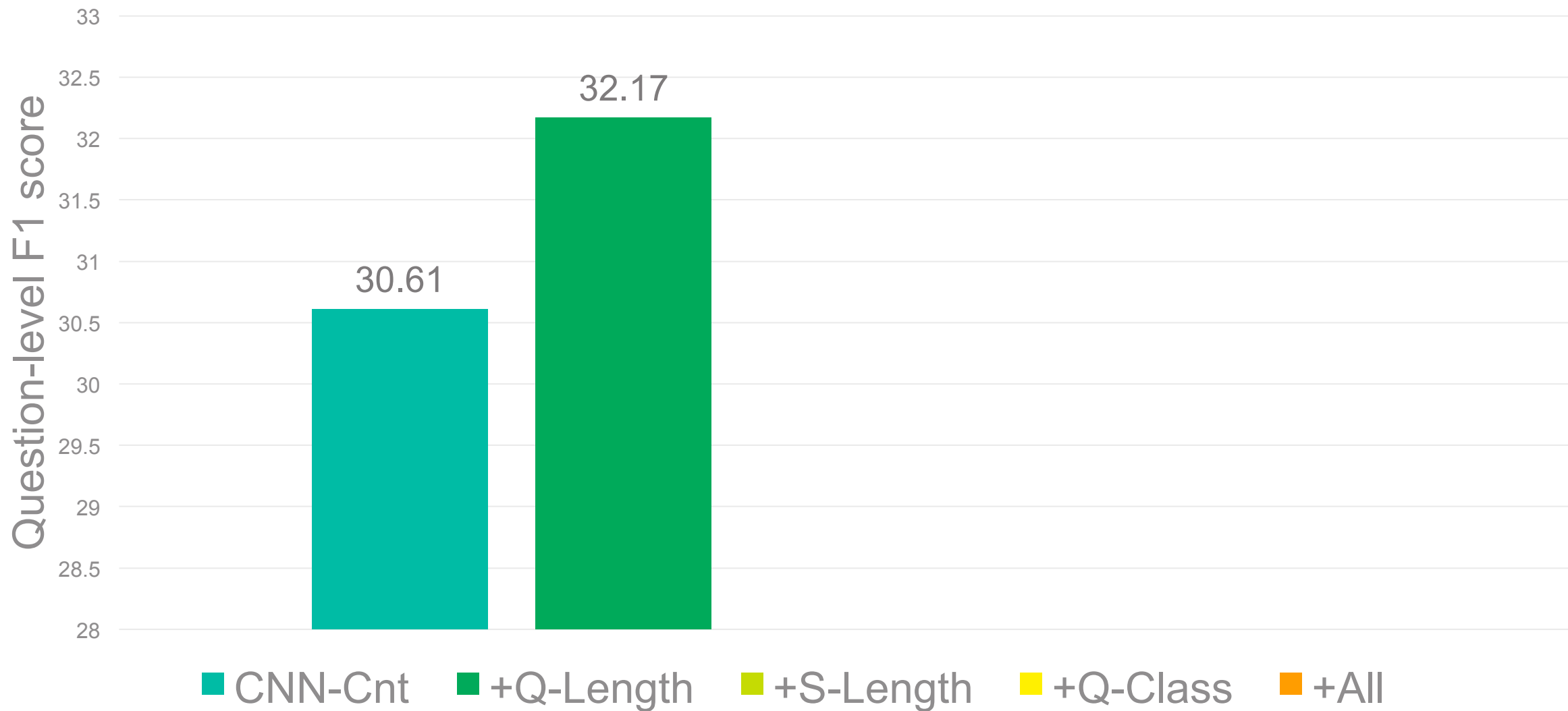
Answer Triggering

- Given a question and a candidate answer sentence set
 - Detect whether there exist correct answers in the sentences
 - Return a correct answer if there exists one
- Evaluation metric
 - Question-level precision, recall and F1 score
- Data – All questions in WikiQA are included in this task

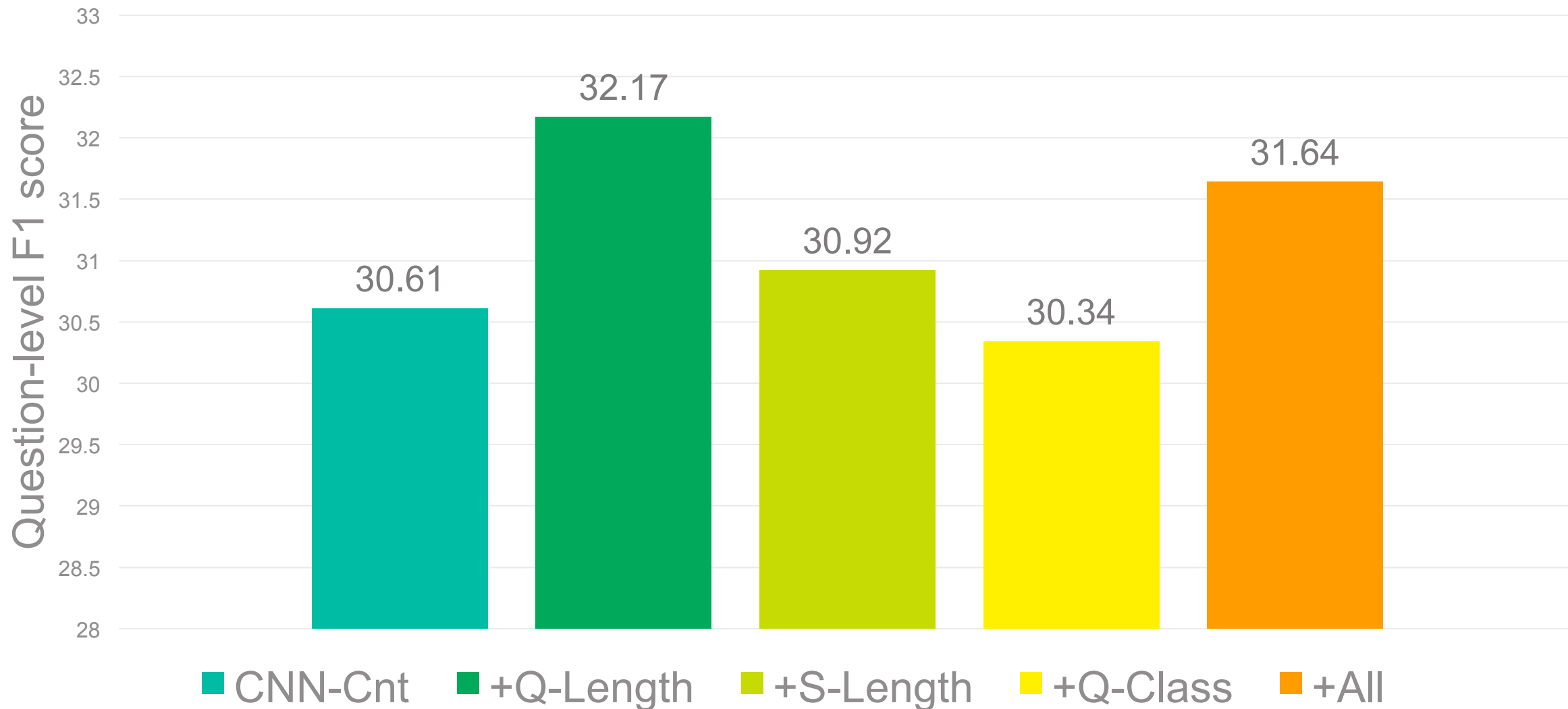
Evaluation on Answer Triggering



Evaluation on Answer Triggering



Evaluation on Answer Triggering



Conclusion (1/2)

- **WikiQA: A new dataset for open-domain QA**
 - Question distribution: questions sampled from Bing query logs
 - Candidate selection: sentences from Wikipedia summary paragraphs
 - Enable “Answer Triggering” by including questions w/o answers
- **Experiments**
 - Different model behaviors on WikiQA and QASent datasets
 - Simple word-matching baseline is no longer strong
 - Answer triggering remains a challenging task

Conclusions (2/2)

- Future Work
 - Investigate advanced semantic matching methods
 - Encoder-decoder semantic matching (e.g., [Sutskever+ NIPS-14])
 - Structured text semantic matching (e.g., [Hu+ NIPS-14])
 - Improve the performance of answer triggering
- Data & Evaluation Script: <http://aka.ms/WikiQ/>
 - Includes “answer phrases” labeled by authors

