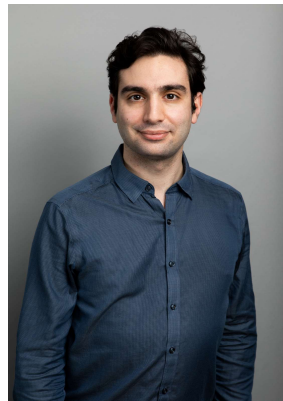




---

# Yes, No or IDK: The Challenge of Unanswerable Yes/No Questions

Elior Sulem, Jamaal Hay and Dan Roth



NAACL 2022

# Yes/No Question Answering



- In order to know whether a claim is true or false given a context, one way is to ask a Yes/No question.

Jane, who is a native of Los Angeles, married a lawyer from NYC.	Did Jane marry a lawyer?	Yes
	Was Jane born in France?	No

- Clark et al., (2019) proposed the **Yes/No QA task** and released the **BoolQ dataset** that includes paragraphs together with naturally occurring questions whose answer is either “Yes” or “No”.

# Yes/No Question Answering



- However, in practical situations, we do not always have the information required to answer the question.
- In these cases, we expect the system to answer “IDK”.

Jane, who is a native of Los Angeles, married a lawyer from NYC.

Did Jane marry a lawyer?

Yes

Was Jane born in France?

No

Did Jane marry in NYC?

IDK

# In this work

---



## 1. Towards a More Natural Task:

- Augmenting BoolQ with unanswerable questions
- Compiling out-of-domain test sets

## 2. Difficulty of the New Task:

- Only training on the task-specific data
- Leveraging other Natural Language Understanding tasks

# Augmenting BoolQ with IDK

---



- Sampling randomly half of the “Yes” questions and half of the “No” questions
- Matching to each of the extracted questions a passage from BoolQ that has the greatest overlapping with the questions in terms of nouns and verbs.
- We apply this process on both training and dev sets.

# Augmenting BoolQ with IDK



- Example:

**Question:** Do they make a new Lombardi Trophy every year?

**Context:** The Vince **Lombardi Trophy** is the **trophy** awarded each **year** to the winning team of the National Football League's championship game, the Super Bowl. The **trophy** is named in honor of NFL coach Vince **Lombardi**.

**Answer:** IDK

# Augmenting BoolQ with IDK



- We obtain new IDK questions (4.7k for train and 1.6k for dev) that we add to the original sets. It represents 33% of the new corpora.
- We call the new corpus **BoolQ<sub>3L</sub>** (for BoolQ with three labels).
- **Validation:**
  - A sample of 200 sentences, 100 from the train set, and 100 from the dev set, validated by two of the authors of the paper: Is the annotation correct? (i.e. Is there an answer?)  
Results:  
94%/ 93% for train (97% absolute inter-annotator agreement)  
95% for dev (100% absolute inter-annotator agreement)

# New Event-Based Corpora

---



- Compiling a test event corpus for yes/no-questions - **ACE-YNQA**, derived from ACE, focusing on time and location (999 Examples)

- **Has-answer:**

“The world bank offered a loan in 1999.”

Did the offer happen in 1999?

**IDK:**

“The world bank offered a loan in 1999.”

Has the loan been paid?



# New Instruction-Based Corpora

---



- Compiling a small Yes/No INSTRUCTIONS corpus (70 examples)

- Has-answer:

“Multiply columns B and C.”

Should a math operation be performed?

IDK:

“Change the font color to green.”

Is the font size 12?

# The Difficulty of the New Task



		Textual Entailment	Binary Textual Entailment
	BoolQ <sub>3L</sub>	MNLI+BoolQ	c(MNLI)+BoolQ
BoolQ <sub>3L</sub> dev	33.64	42.66	43.25
ACE-YNQA	52.02	52.02	54.94
INSTRUCTIONS	26.56	43.75	20.31

**3 LABELS:** Accuracy of a BERT-LARGE based system trained on BoolQ<sub>3L</sub>

	BoolQ	MNLI+BoolQ	c(MNLI)+BoolQ
BoolQ dev	72.88	78.24	79.49
ACE-YNQA <sub>Y/N</sub>	59.53	65.47	68.01
INSTRUCTIONS <sub>Y/N</sub>	65.00	70.00	65.00

**2 LABELS:** Accuracy of a BERT-LARGE based system trained on BoolQ

# Conclusion

---



- In order to use Yes/No QA in realistic scenarios, the ability to answer IDK is necessary.
- We extend the existing Yes/No QA benchmark:
  - Augmenting BoolQ with unanswerable questions
  - Compiling out-of-domain test sets, including a Yes/No QA test set related to event extraction.
- We show the considerable difficulty of 3-label Yes/No QA compared to the 2-label setting.



Thank you for listening

The data, code and pre-trained models can be found at

[http://cogcomp.org/page/publication\\_view/955](http://cogcomp.org/page/publication_view/955)

Jane, who is a native of  
Los Angeles, married a  
lawyer from NYC.

Did Jane marry a lawyer?

Yes

Was Jane born in France?

No

Did Jane marry in NYC?

IDK

[eliors@seas.upenn.edu](mailto:eliors@seas.upenn.edu)

<https://www.cis.upenn.edu/~eliors/>