

Question Answering over Legal Documents

CIS 530 Final Project

Ahmed Ahmed, Danqi Luo, Hannah Gonzalez,
Jiani Huang, Karan Jaisingh

Outline and Timeline

- | | | |
|------------------------|---|----------------------------------|
| 1. Cuad Dataset | } | 2.5 minutes - Jiani |
| 2. Motivating Examples | | |
| 3. Existing Approaches | | |
| 4. Improvement 1 | } | 2 minutes - Ahmed, Hannah, Karan |
| 5. Performance 1 | | |
| 6. Improvement 2 | } | 2 minutes - Jason, Jiani |
| 7. Performance 2 | | |
| 8. Conclusion | | 0.5 minute - Jiani |

CUAD Dataset

THIS CO-BRANDING AGREEMENT (the "Agreement") is made as of **May 22, 2000** (the "Effective Date"), by and between WOMEN.COM NETWORKS, INC., ..., shall remain effective for **two (2) years** from and after the Effective Date (the "Initial Term"). ...

Legal Contract

Effective Date
Expiration Date
Renewal Term
Notice Period To Terminate Renewal
....

Questions

Effective Date	May 22, 2000
Effective Date-Answer	5/22/00
Expiration Date	['shall remain effective for two (2) years from and after the Effective Date']
Expiration Date-Answer	5/22/02
...	...

Answers

We will start form the dataset. We focus on a legal contract review dataset called CUAD.

This dataset has three components: 1) a legal contract in text form, 2) questions regarding to the contract 3) the corresponding answers to the questions.

These questions are designed by legal experts.

The answers are annotated by law school students and reviewed by experts.

Motivating Example

THIS CO-BRANDING AGREEMENT (the "Agreement") is made as of May 22, 2000 (the "Effective Date"), by and between WOMEN.COM NETWORKS, INC., ..., shall remain effective for two (2) years from and after the Effective Date (the "Initial Term"). ...

Legal Contract

Effective Date
Expiration Date
Renewal Term
Notice Period To Terminate Renewal
....

Questions

Effective Date	[May 22, 2000]
Effective Date-Answer	5/22/00
Expiration Date	['shall remain effective for two (2) years from and after the Effective Date']
Expiration Date-Answer	5/22/02
...	...

Explanation

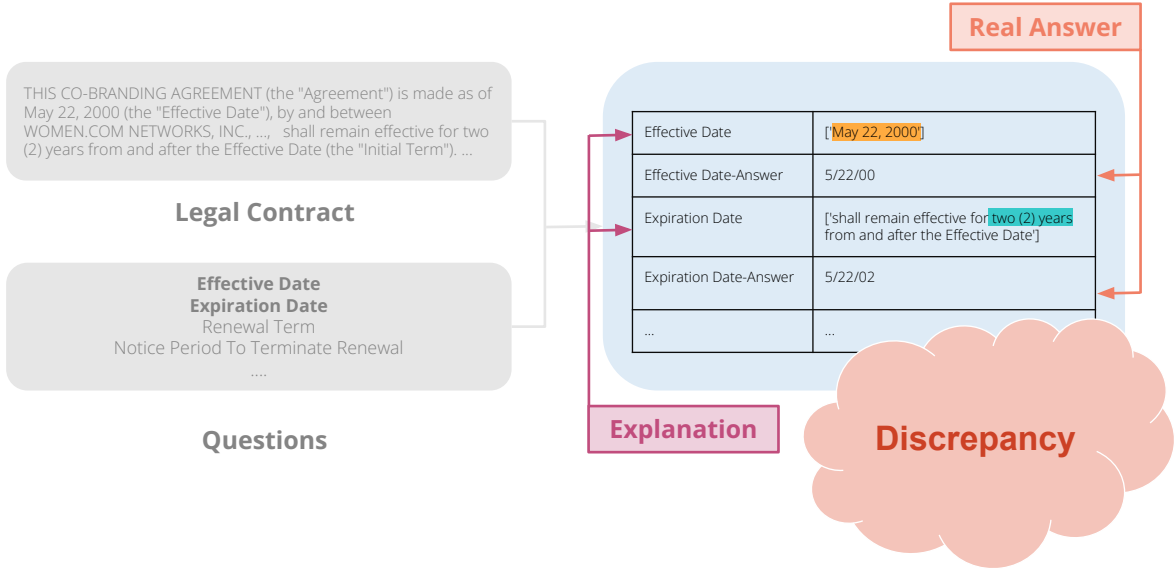
Answers

Real Answer

Its answer to each question has two different forms, a real answer and an explanation to the answer.

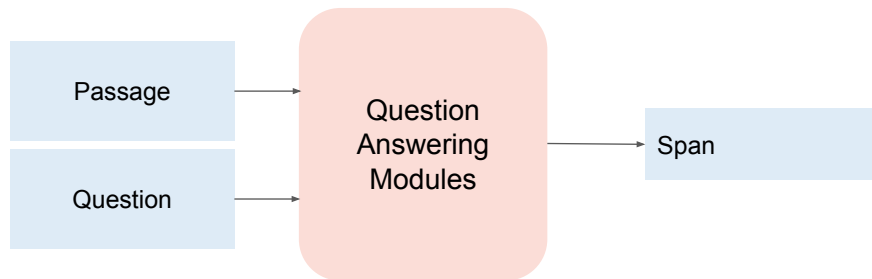
For example, although the expiration date says, May 22nd, 2002, but it is not directly reflected in the explanation. In this case, the explanation only says, the term last for 2 years, starting from the effective date.

Motivating Example



Note that the explanation and the real answer are not the same. We need to perform an extra step here to go from explanation to the real answer.

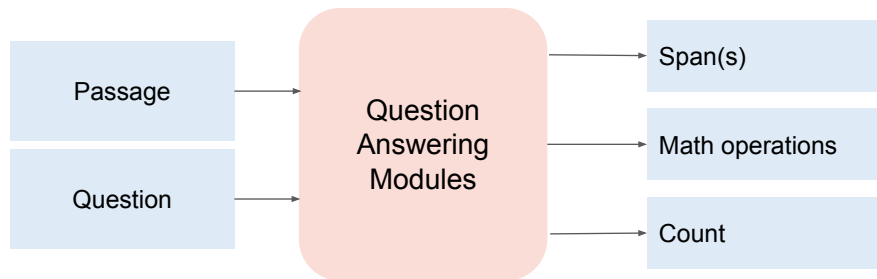
Existing Approaches



Examples includes hugging face QA transformer.

Classic approaches solving the machine reading comprehension problem usually just output a span in the text, such as hugging face QA transformer.

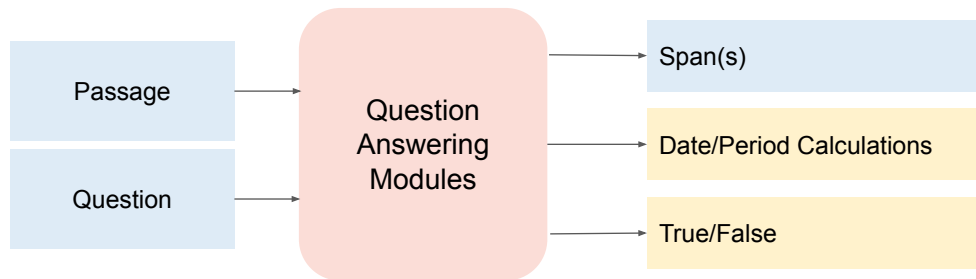
Existing Approaches



Examples includes NeRd, QDGAT. Baseline Model

More recent approaches can generate a variety of the outputs like spans, math operations and counts. For example, the QDGAT framework.

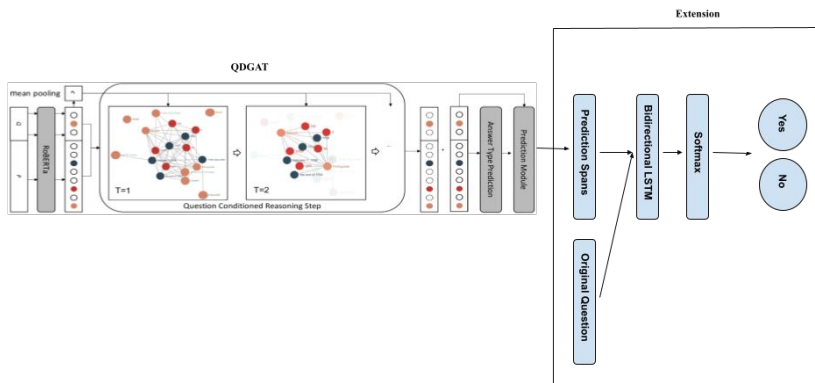
Our Objective



However, we aim to add two other extensions to the question answering process. The date and period calculations, and true or false question and answering.

<Ahmed> will introduce us with the extension 1, answering T/F questions.

Extension 1: True/False Questions



Uses the outputs of QDGAT, which are spans to predict a binary outcome of yes/no using a bidirectional LSTM

Extension 1: Performance

- Prediction Imbalance
- Looking Forward
 - Better context embedding
 - More data
 - Use of Attention

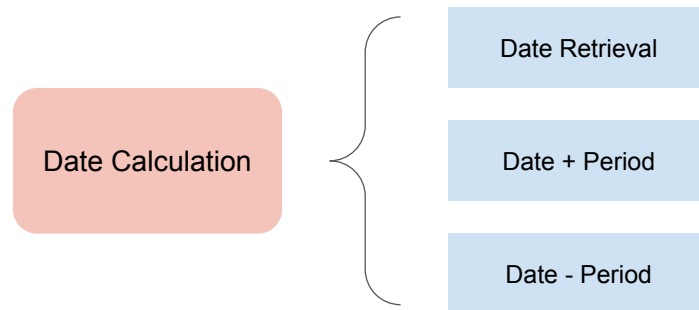
	QDGAT w/o extension 1	QDGAT-extension 1
F1	0.61	0.83

4x more negative instances than positive instances in the data

.78 precision, .88 recall more likely to predict negatives than positives

Full information needed to predict yes/no better is not achieved through QDGAT only

Extension 2: Date/Period Calculations



For our second extension, we add date/period calculations.

The original qdgat baseline does not perform well on date extractions and cannot do arithmetic calculations on dates and periods.

We have three kinds of Date Calculations: Date Retrieval, Date + Period and Date - Period.

Extension 2: Date/Period Calculations

THIS CO-BRANDING AGREEMENT (the "Agreement") is made as of May 22, 2000 (the "Effective Date"), by and between WOMEN.COM NETWORKS, INC., ..., shall remain effective for two (2) years from and after the Effective Date (the "Initial Term"). ...

Phrases	May 22, 2000	Women.com networks, Inc	...	Two (2) years	...
NER Result	DATE	ORG	...	PERIOD	...

Date calculations are especially important for questions for which the answers require some arithmetic operations.
For example, in this case, in order to calculate the Expiration date, we need Effective Date + Renewal Term.

Extension 2: Date/Period Calculations

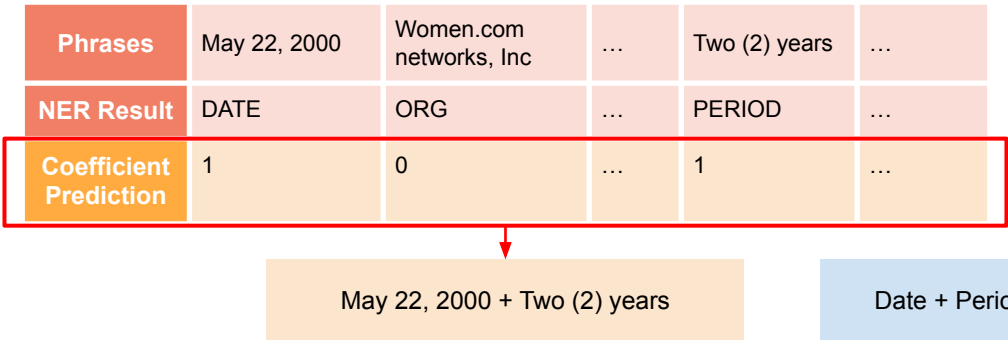
THIS CO-BRANDING AGREEMENT (the "Agreement") is made as of May 22, 2000 (the "Effective Date"), by and between WOMEN.COM NETWORKS, INC., ..., shall remain effective for two (2) years from and after the Effective Date (the "Initial Term"). ...

Phrases	May 22, 2000	Women.com networks, Inc	...	Two (2) years	...
NER Result	DATE	ORG	...	PERIOD	...
Coefficient Prediction	1	0	...	1	...

In order to perform the calculations. We will train a classifier and assign different coefficients (1,0,-1) to each entity.
To perform addition, we will assign coefficients 1 to the entities we want to add together.
In this case, the effective date and renewal term.

Extension 2: Date/Period Calculations

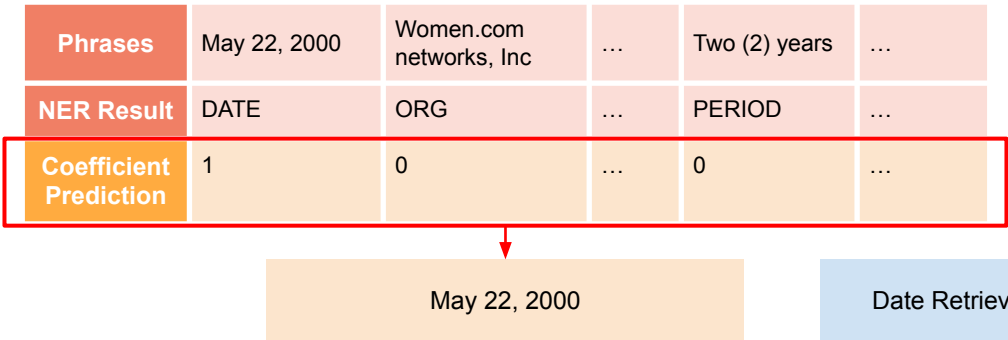
THIS CO-BRANDING AGREEMENT (the "Agreement") is made as of May 22, 2000 (the "Effective Date"), by and between WOMEN.COM NETWORKS, INC., ..., shall remain effective for two (2) years from and after the Effective Date (the "Initial Term"). ...



Now we can get the expiration date (May 22, 2002).

Extension 2: Date/Period Calculations

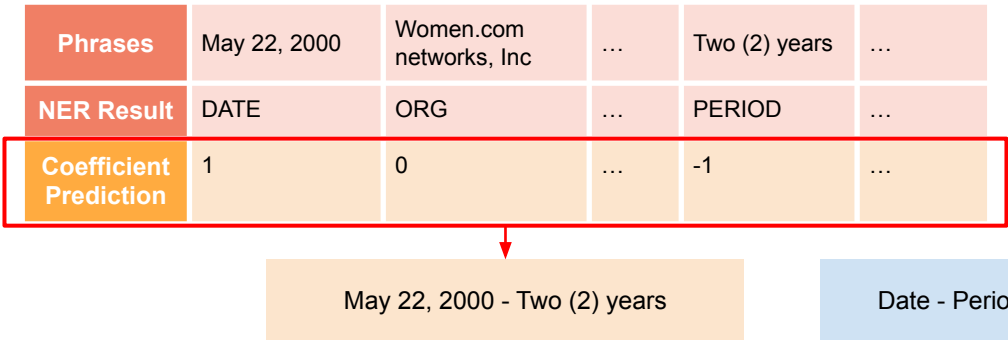
THIS CO-BRANDING AGREEMENT (the "Agreement") is made as of May 22, 2000 (the "Effective Date"), by and between WOMEN.COM NETWORKS, INC., ..., shall remain effective for two (2) years from and after the Effective Date (the "Initial Term"). ...



When we want to answer questions that does not require arithmetic operations, we only assign +1 coefficients to the entity we want. Here, we extract the effective date.


Extension 2: Date/Period Calculations

THIS CO-BRANDING AGREEMENT (the "Agreement") is made as of May 22, 2000 (the "Effective Date"), by and between WOMEN.COM NETWORKS, INC., ..., shall remain effective for two (2) years from and after the Effective Date (the "Initial Term"). ...



In some case, if we want to perform subtractions, we will assign a -1 coefficient to the entity we want to subtract.

Extension 2: Date/Period Calculations



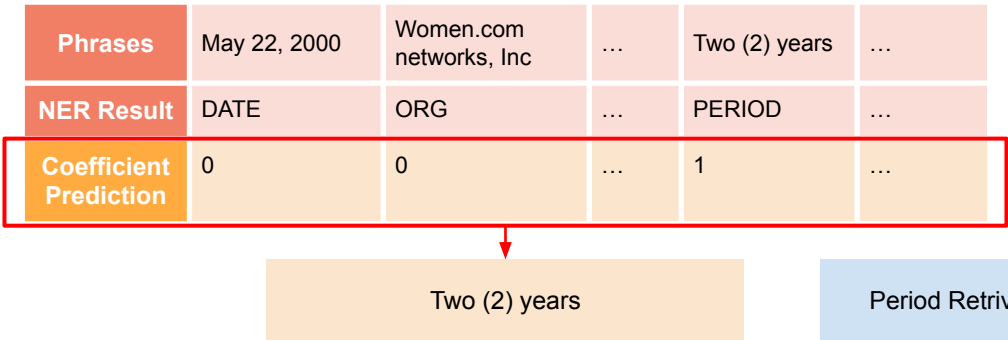
Period Calculation

Period Retrieval

In Legal contracts, we are also asked questions about periods like renewal term and notice period.

Extension 2: Date/Period Calculations

THIS CO-BRANDING AGREEMENT (the "Agreement") is made as of May 22, 2000 (the "Effective Date"), by and between WOMEN.COM NETWORKS, INC., ..., shall remain effective for two (2) years from and after the Effective Date (the "Initial Term"). ...

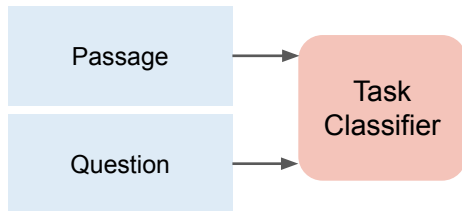


In this case, we don't need to perform arithmetic operations, thus we only assign +1 coefficients to the period entity we want.

Extension 2: Date/Period Calculations

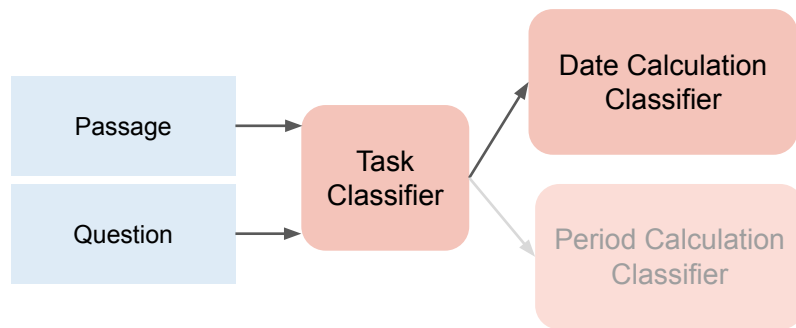
Date / Period Calculation can be represented as a **classification** network

Extension 2: Date/Period Calculations



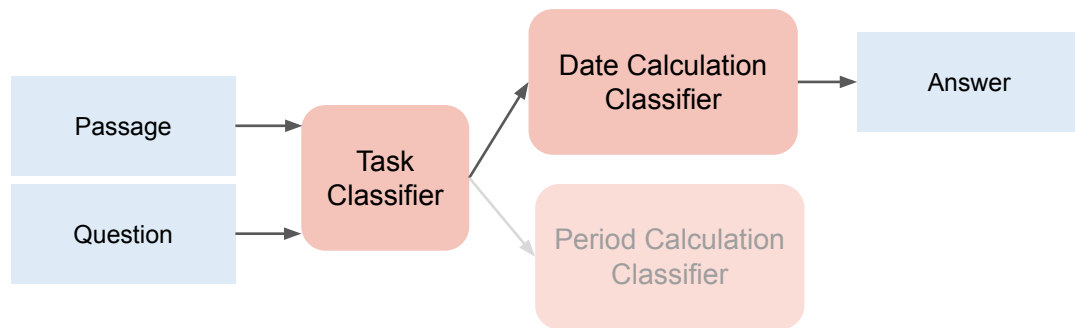
The passage and questions are passed into a task classifier to determine the type of question we want to answer.

Extension 2: Date/Period Calculations



The task classifier will tell us whether it is a date question or a period question.

Extension 2: Date/Period Calculations



If it is a date question, we will further decide whether we want addition, subtraction or just retrieval.
And then predict the answer.

Dataset

Filter out the data points which cannot achieve the correct answer with date and period calculation.

Valid datapoint number: 169

(Stanford Core NLP package does not perform well on the NER task)

Our dataset still comes from the CUAD dataset.

However, since the Stanford Core NLP package does not perform well on our NER task.

We only get 169 valid datapoint for our training and testing.

Extension 2: Performance

	QDGAT w/o extension 2	QDGAT-extension 2
F1	0.12	0.22

The second extension increased the original F1 score from 0.12 to 0.22.

Extension 2: Performance

	QDGAT w/o extension 2	QDGAT-extension 2
F1	0.12	0.22

Using much more labels

While using fewer labels and is faster to train.

Extension 2: Performance

QDGAT w/o extension 2	QDGAT-extension 2	Ground Truth
6th	4/6/1999	4/6/1999
13	None	9/1/2004
10th	1/10/2018	1/18/2010

The output of extension 2 is significantly more readable and complete than without using the extension.

The answer without extension 2 cannot output a valid date, but part of the date majority of the time as a single number.

(Almost no exact match and unable to predict periods in readable form)

Conclusion

- LSTM
 - Improved F1 score
 - Slight favoring towards predicting negative class
 - Ability to handle binary output
 - Good set up for future extensions
- Date/Period Calculation
 - Improved F1 score
 - Complete and readable date predictions
 - Arithmetic operations between dates and periods
 - Ability to handle period questions

Questions?

Thank you!