Infusing Disease Knowledge into BERT for Health Question Answering, Medical Inference and Disease Name Recognition

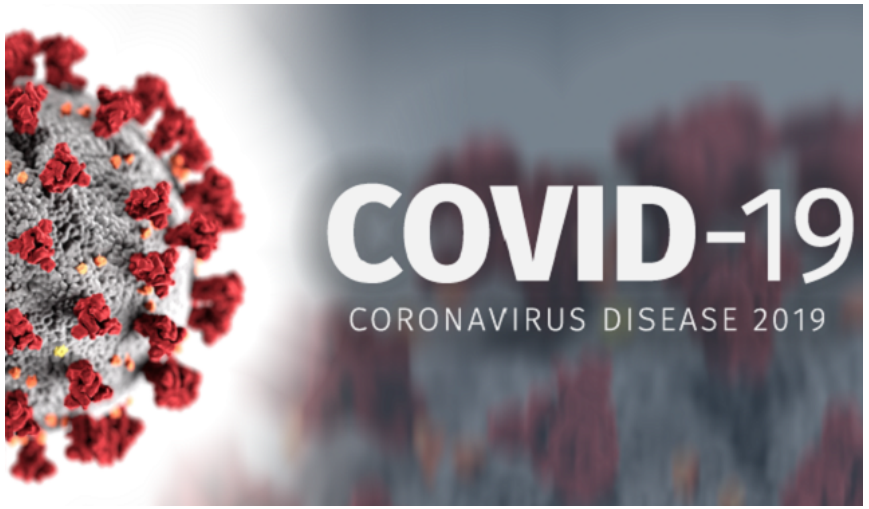
Yun He[1], *Ziwei Zhu*[1], *Yin Zhang*[1], *Qin Chen*[2], *James Caverlee*[1]
[1]*Texas A&M University, College Station, USA*
[2]*Fudan University, Shanghai, China*

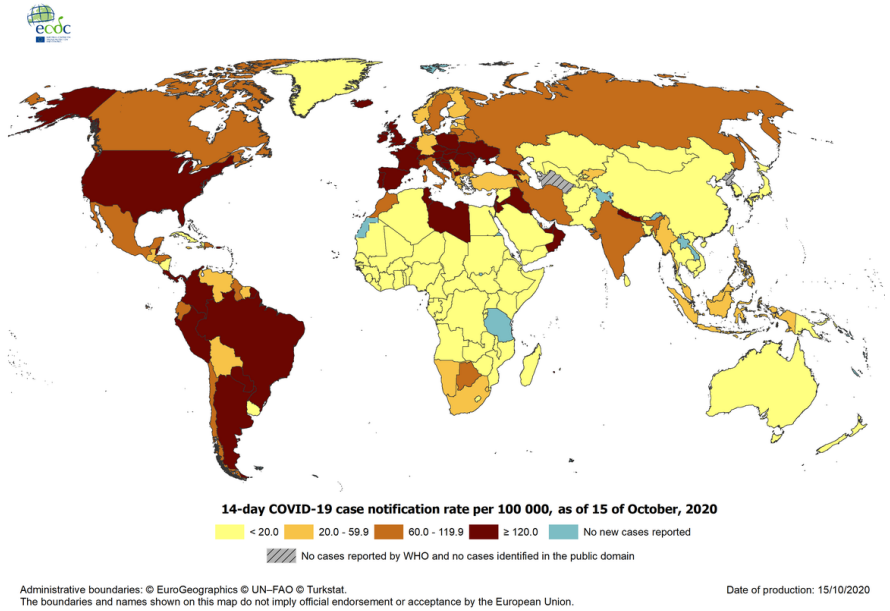
Overview

- Disease knowledge is critical for many health and biomedical related tasks such as consumer health question answering.
- In this paper, we integrate BERT models with disease knowledge for improving these important tasks.
- Our proposed method, Disease Knowledge Infusion Training, achieves new SOTA results in two datasets: MEDIQA-2019 and MEDNLI.

Disease

- Disease has a huge impact on human life.



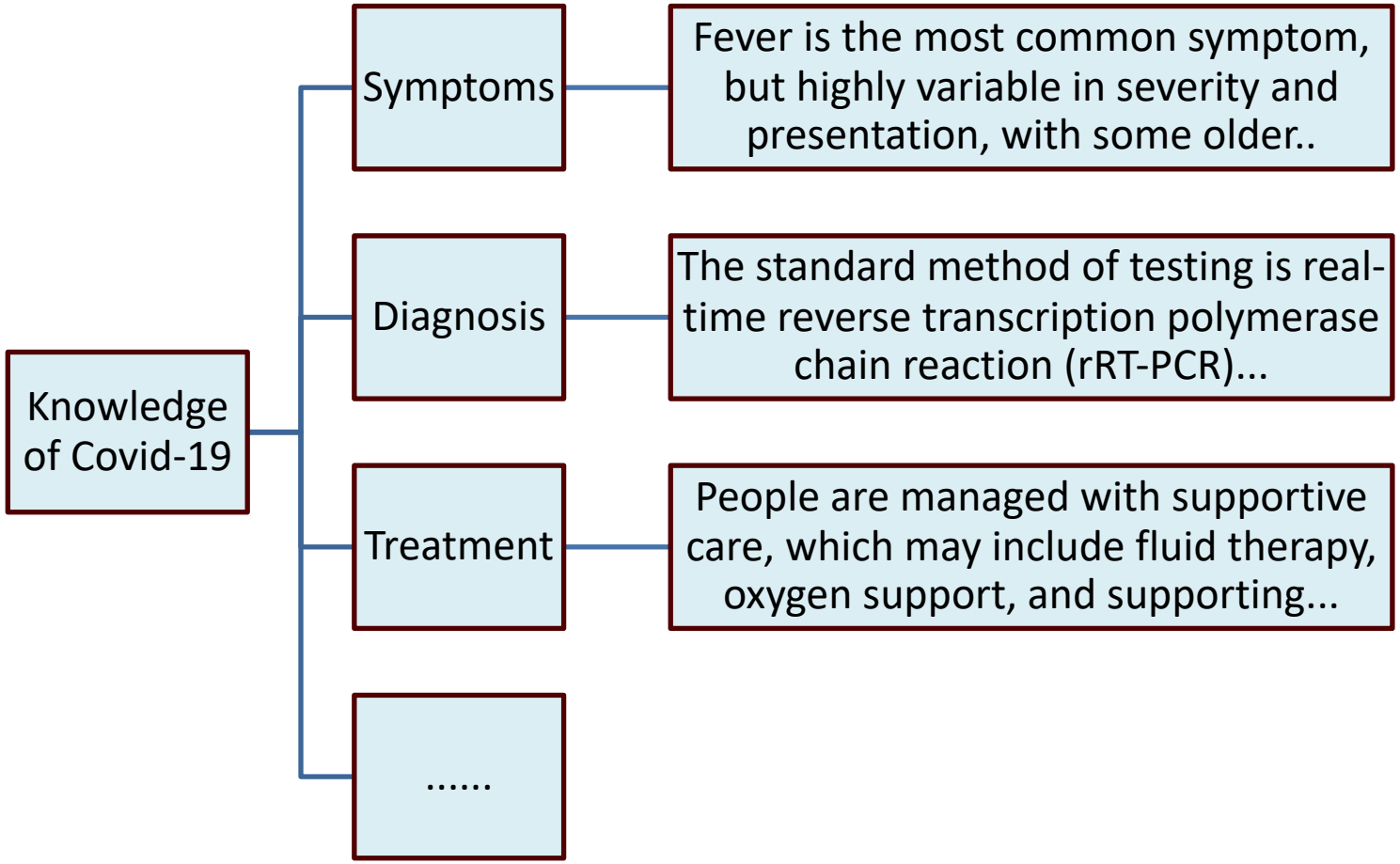


- Disease is one of the fundamental biological entities in biomedical research.

Until October 12, 2020, COVID-19 has killed 1.08M people and disrupted the everyday lives of billions of people worldwide.

Disease Knowledge

- Knowledge of a disease includes information of various aspects of the disease.
- This disease knowledge is critical for many health and biomedical related NLP tasks.



Consumer Health Question Answering

TEXAS A&M UNIVERSITY

Question: …keen to learn **how to get COVID-19 diagnosed**, many thanks

Answer 1: ... real-time reverse transcription polymerase chain reaction...
Answer 2: ... diagnosis of vipoma requires demonstration of diarrhea...
Answer 3: ...affected by this disorder are not able to make lipoproteins…

Label: Answer 1 is the most relevant
Disease Knowledge: Answer 1 is the diagnosis of COVID-19

Disease-descriptive text
Answer 1:"...real-time reverse transcription polymerase chain reaction..."

Semantic Relations

Disease
COVID-19

Aspect
Diagnosis

BERT

Medical Language Inference

Premise: She was **not able to speak, but appeared to comprehend well**

Hypothesis: Patient had **aphasia**

Label: entailment

Disease Knowledge: Premise describes the symptoms of aphasia

Disease-descriptive text

Premise: "...not able to speak, but appeared to comprehend well..."

Semantic Relations

Disease
Aphasia

Aspect
Symptoms

BERT

Disease Name Recognition

Text: **Myotonic dystrophy** (DM) is **caused by a CTG expansion** in the 3 untranslated region of the DM gene.

Label: Myotonic dystrophy

Disease Knowledge: the text contains the cause of Myotonic dystrophy

Disease-descriptive text

Text: "...caused by a CTG expansion..."

Semantic Relations

Disease
Myotonic dystrophy

Aspect
Cause

BERT

Infusing Disease Knowledge into BERT models

TEXAS A&M UNIVERSITY

Disease-descriptive text

"...real-time reverse transcription polymerase chain reaction..."

Semantic Relations

Disease
COVID-19

Aspect
Diagnosis

BERT

I want to know the semantic relations among them so that I can tell which disease and aspect is this disease-descriptive text talking about.

Infusing Disease Knowledge into BERT models

TEXAS A&M UNIVERSITY

Disease-descriptive text
"...real-time reverse transcription polymerase chain reaction..."

BERT is trained to Infer the corresponding disease and aspect from this text.

Ground-truth: Diagnosis

Predicted: Treatment ✗

Predicted: COVID-19 ✓

Ground-truth: COVID-19

By minimizing the loss between the prediction and the ground-truth, BERT can memorize the disease knowledge.

Resources of Disease Terms

TEXAS A&M UNIVERSITY

First, we seek a disease vocabulary that provides disease terms. Several resources include:

Medical Subject Headings MeSH

Unified Medical Language System ®

NCI thesaurus

SNOMED CT The global language of healthcare

As a first step, we choose MeSH, which is a comprehensive controlled vocabulary proposed by the National Library of Medicine.

```
Infections;C01
Aneurysm, Infected;C01.069
Arthritis, Infectious;C01.100
Arthritis, Reactive;C01.100.500
Asymptomatic Infections;C01.125
Bacterial Infections and Mycoses;C01.150
Bacterial Infections;C01.150.252
Bacteremia;C01.150.252.100
```

Examples of MeSH Disease Terms

Proposed Method: Disease Knowledge Infusion Training

TEXAS A&M UNIVERSITY

1. Obtain disease terms from MeSH

WIKIPEDIA
The Free Encyclopedia

2. Obtain Articles of diseases from Wikipedia

COVID-19

1 Signs and symptoms
- 1.1 Complications

2 Cause
- 2.1 Transmission
- 2.2 Virology

3 Pathophysiology
- 3.1 Immunopathology

4 Diagnosis
- 4.1 Pathology

5 Prevention

6 Management

Passage: The standard method of testing is real-time reverse transcription polymerase chain reaction (rRT-PCR)...

3. Extract text from a section as the passage.

Disease: COVID-19 **(title of the Wikipedia article)**

Aspect: Diagnosis **(title of the section)**

4. Extract the weakly-supervised topic disease and aspect.

Auxiliary Sentence:
What is the **diagnosis** of **COVID-19**?

5. Construct an auxiliary sentence that mentions the topic disease and aspect.

New Passage for MLM:
What is the **[MASK]** of **[MASK]**? The standard method of testing is real-time reverse transcription polymerase chain reaction (rRT-PCR)...

6. Concatenate the passage and the auxiliary Sentence. BERT is trained to infer the disease and aspect.

Experiments

We infuse disease knowledge into these BERT models:

- General BERT
  - BERT-base (Wikipedia, BookCorpus)
  - ALBERT (Wikipedia, BookCorpus)
- Biomedical BERT
  - BioBERT (PubMed articles)
  - ClinicalBERT (Clinical notes)
  - BlueBERT (PubMed abstracts and clinical notes)
  - SciBERT (18% of papers from the computer science domain and 82% from the biomedical domain)

Relationship between Our Approach and other BERT Models

TEXAS A&M UNIVERSITY

General BERT:
Pre-training on Wikipedia from scratch

Continual Pre-training BERT with MLM and NSP on biomedical corpora.

BioBERT

Continual Pre-training BioBERT with proposed **Disease Knowledge Infusion Training.**

BioBERT + Disease Knowledge

WIKIPEDIA
The Free Encyclopedia

PubMed

Medical Subject Headings MeSH

WIKIPEDIA
The Free Encyclopedia

Relationship with Biomedical BERT

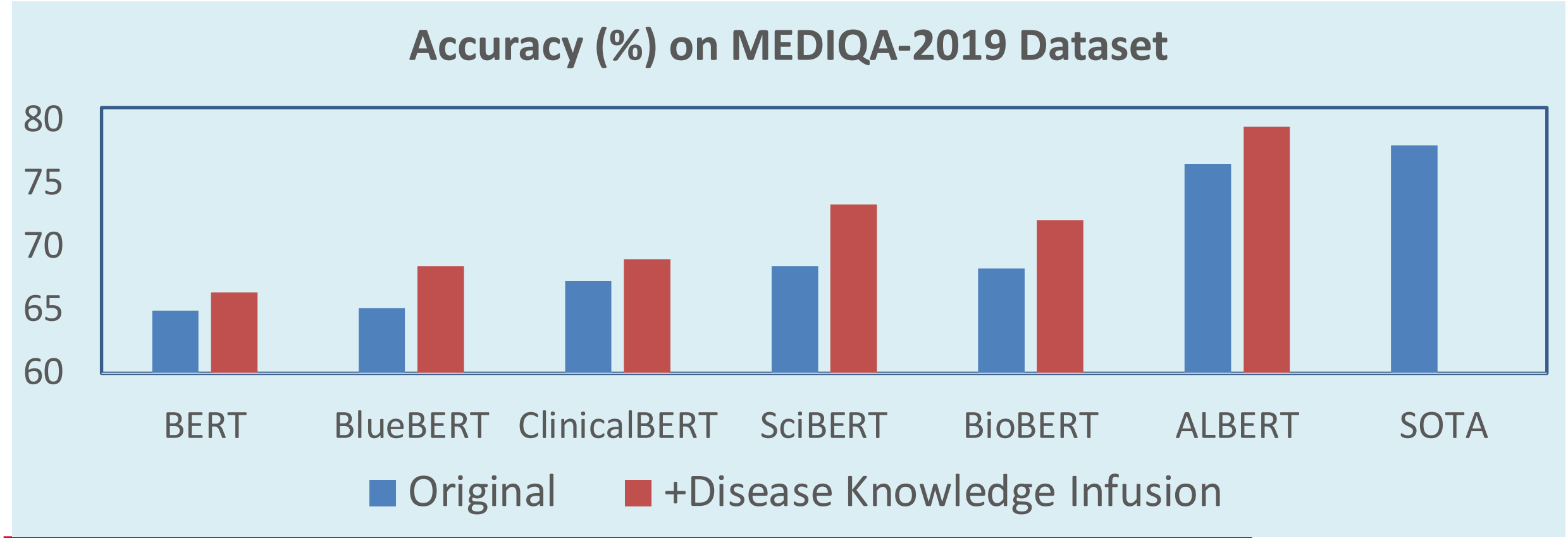
TEXAS A&M UNIVERSITY

- The biomedical BERT models (e.g., ClinicalBERT and BioBERT) capture the **general** syntactic and semantic knowledge of biomedical language.
- Our method is specifically designed for capturing the semantic relations between a disease-descriptive text and its corresponding aspect and disease.
- These biomedical BERT models can be further enhanced by our approach in biomedical NLP tasks.

Consumer Health Question Answering

TEXAS A&M UNIVERSITY

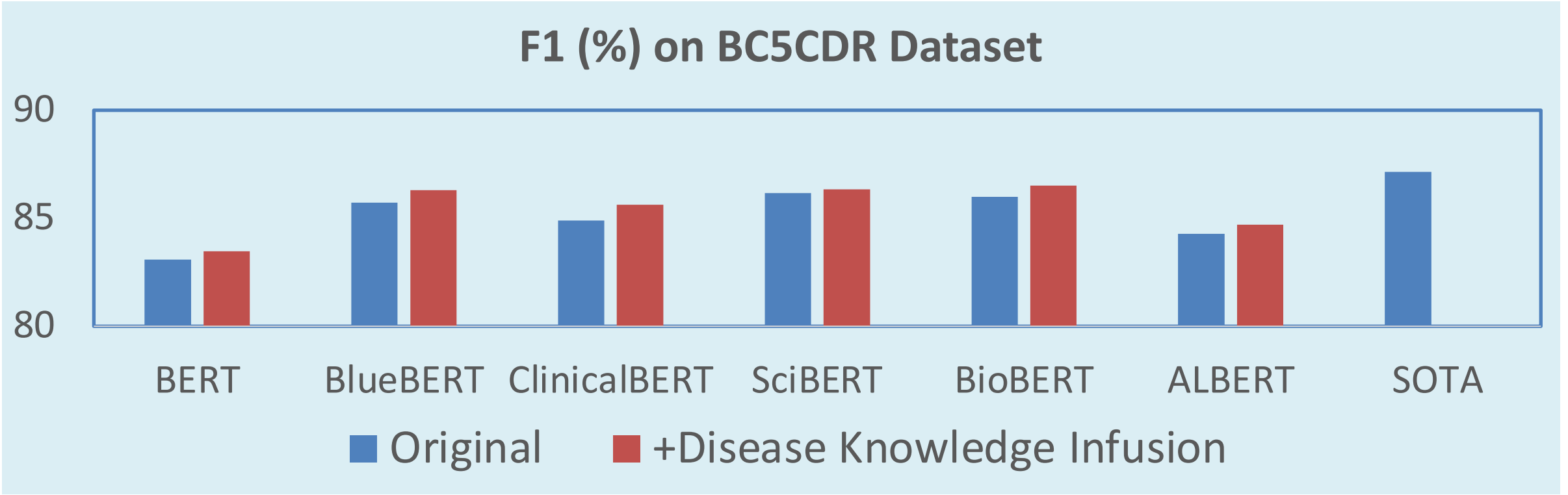
Our approach can enhance BERT models in nearly all cases.
New SOTA results are obtained in MEDIQA-2019 datasets.



Disease Name Recognition

TEXAS A&M UNIVERSITY

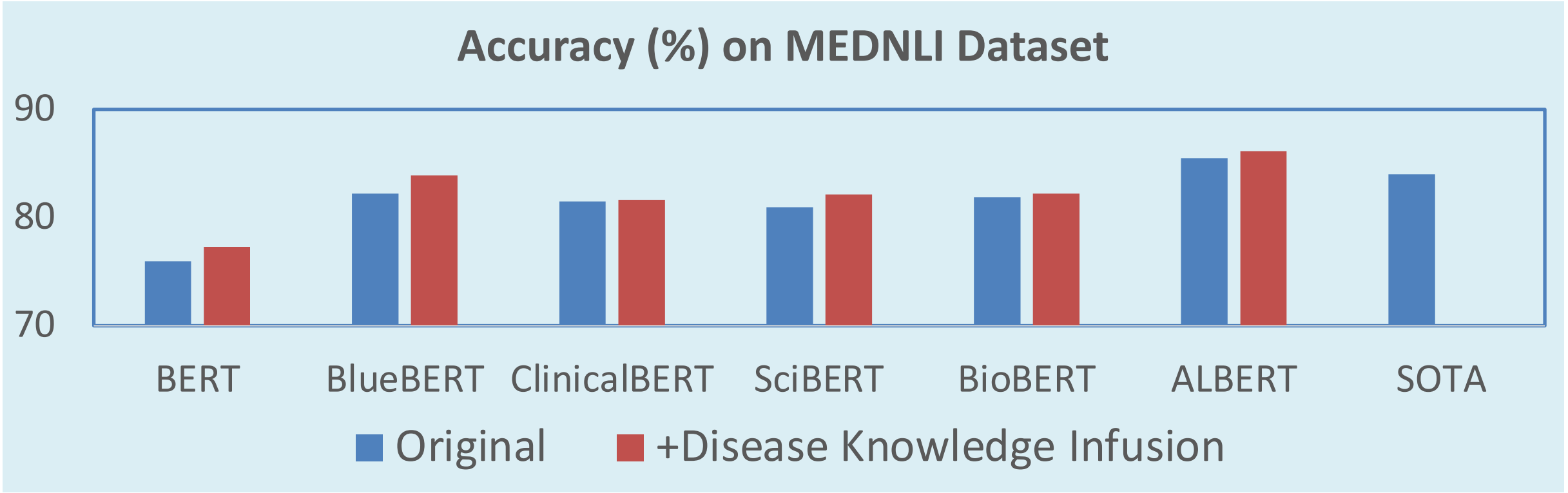
Our approach does not outperform SOTA but it can enhance BERT models in most cases.



Medical Language Inference

TEXAS A&M UNIVERSITY

Our approach can enhance BERT models in nearly all cases.
New SOTA results are obtained in MEDNLI datasets.



Conclusion

- We propose Disease Knowledge Infusion Training to augment BERT-like models with disease knowledge.
- Our approach can enhance both general BERT models (e.g., BERT and ALBERT) and biomedical BERT models (e.g., ClinicalBERT and BioBERT).
- New SOTA results are obtained in MEDIQA-2019 and MEDNLI datasets.

Thanks!

Implicative Reasoning in Conversational Machine Reading

Yifan Gao, The Chinese University of Hong Kong

Nov 4, 2020

Machine Comprehension

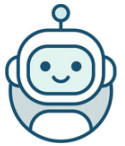
SQuAD: 100,000+ Questions for Machine Comprehension of Text (Rajpurkar et al., 2016)

The current Chief Executive is **Carrie Lam**, who was selected on 26 March 2017, appointed by the Central People's Government with the State Council Decree signed by Premier Li Keqiang, on 11 April 2017 and took office on 1 July 2017.

✓ Literal Answer

Q: Who is the chief executive of Hong Kong?

A: Carrie Lam



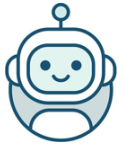
Conversational Question Answering

CoQA: A Conversational Question Answering Challenge (Reddy et al., 2018)

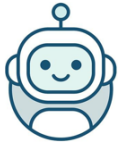
Incumbent **Democratic** President Bill Clinton was ineligible to serve a **third term** due to **term limitations** in the 22nd Amendment of the Constitution, and Vice President Gore was able to secure the Democratic nomination with relative ease.

- ✓ Literal Answer
- ✓ Dialog Understanding

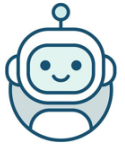
Q: What political party is Clinton a member of?

A: Democratic 

Q: What was he ineligible to serve?

A: third term 

Q: Why?

A: term limitations 

However...

Interpreting Natural Language Rules

The text to read may not contains the literal answer, but it contains a **recipe** to derive it.

Scenario: I am a 34-year-old man from the United States who owns their own business. We are an American small business.

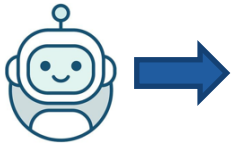
Question: Is the 7(a) Loan Program for me?

Implicative Reasoning in Conversational Machine Reading

Introduction

Scenario: I am a 34-year-old man from the United States who owns their own business. We are an American small business.

Question: Is the 7(a) Loan Program for me?



Rule Text

7(a) loans are the most basic and most used type loan of the Small Business Administration's (SBA) business loan programs. Its name comes from section 7(a) of the Small Business Act, which authorizes the agency to provide business loans to American small businesses. The loan program is designed to assist for-profit businesses that are not able to get other financing from other resources.

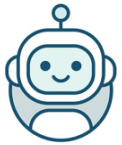
Implicative Reasoning in Conversational Machine Reading

Introduction

Scenario: I am a 34-year-old man from the United States who owns their own business. We are an American small business.

Question: Is the 7(a) Loan Program for me?

Are you a for-profit business?



Yes

Are you able to get financing from other resources?

No

Yes. (You can apply the loan.)

Rule Text

7(a) loans are the most basic and most used type loan of the Small Business Administration's (SBA) business loan programs. Its name comes from section 7(a) of the Small Business Act, which authorizes the agency to provide business loans to ① American small businesses. The ② loan program is designed to assist for-profit businesses ③ that are not able to get other financing from other resources.

Task Definition

ShARC: Shaping Answers with Rules through Conversation

ShARC

(Saeidi et al., 2018)

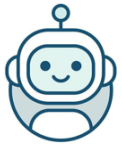
Rule Text: 7(a) loans are the most basic and ...

Scenario: I am a 34-year-old man ....

Question: Is the 7(a) Loan Program for me?

Dialog History

Follow-up Q1: Are you a for-profit business? A1: Yes



Decision Making

Make a prediction among:

Yes, No, Irrelevant, Inquire

- **Yes/No**: Directly answer the question
- **Irrelevant**: unanswerable
- **Inquire**

Question Generation

Ask a follow-up question to clarify the unknown user information

Outline

1. Explicit Memory Tracker with Coarse-to-Fine Reasoning for Conversational Machine Reading, ACL 2020

2. Discern: Discourse-Aware Entailment Reasoning Network for Conversational Machine Reading, EMNLP 2020

The 58th Annual Meeting of
the Association for Computational Linguistics

July 5 – 10
2020

Explicit Memory Tracker with Coarse-to-Fine Reasoning for Conversational Machine Reading

Yifan Gao[1], Chien-Sheng Wu[2], Shafiq Joty[2], Caiming Xiong[2], Richard Socher[2], Irwin King[1], Michael R. Lyu[1], Steven C.H. Hoi[2]

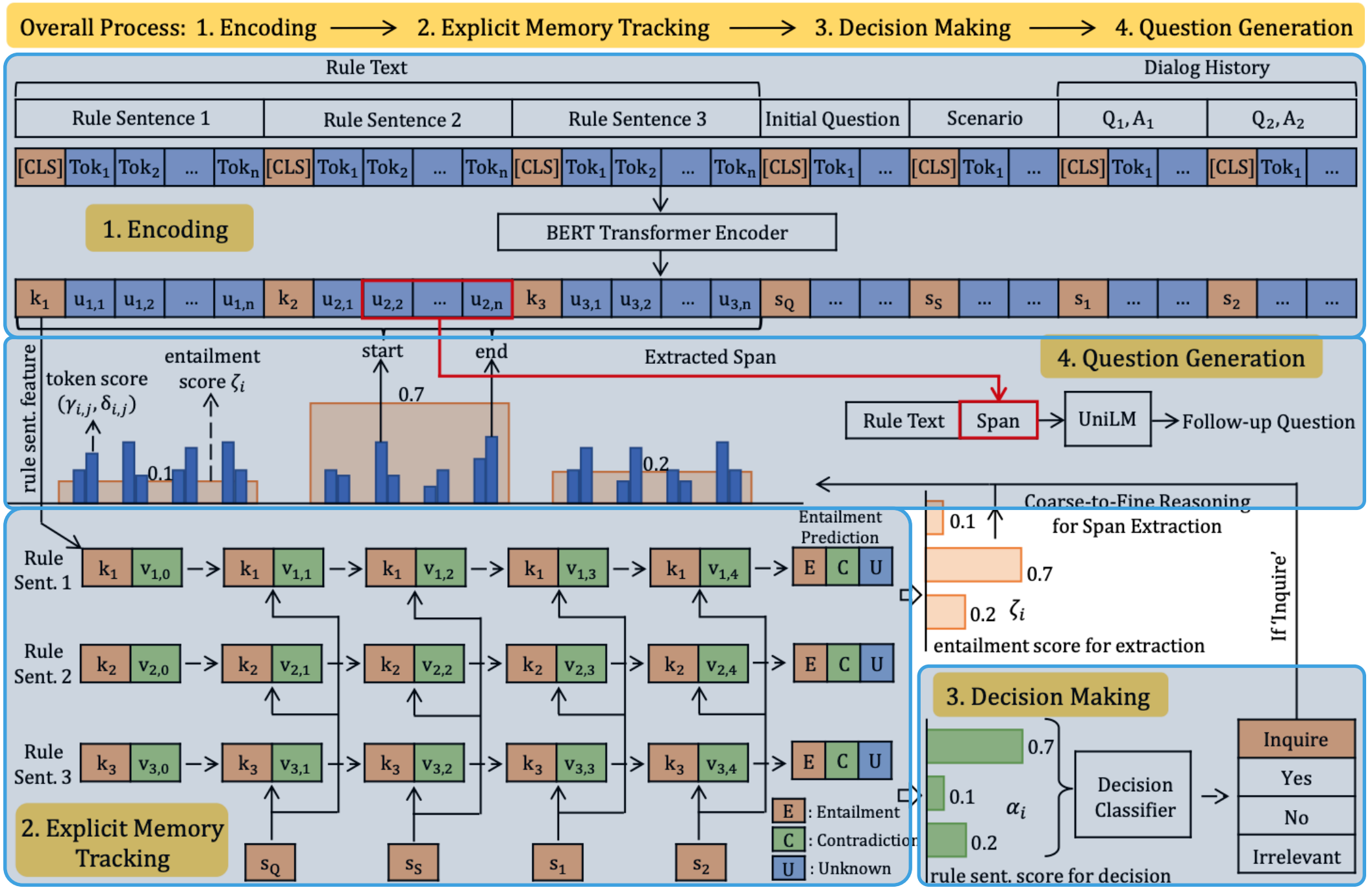
1. The Chinese University of Hong Kong 2. Salesforce Research

Code & Models: https://github.com/Yifan-Gao/explicit_memory_tracker

Proposed Solution

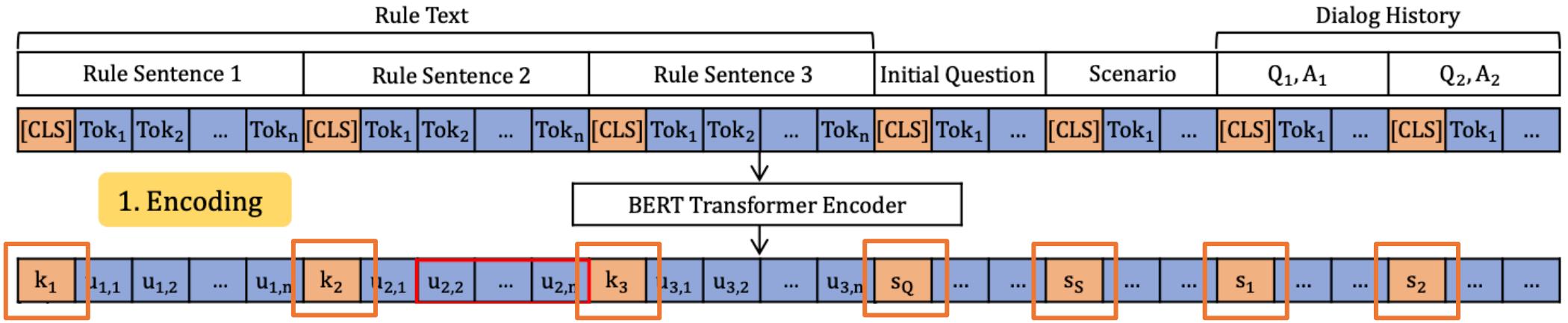
Contributions

- Explicit Memory Tracker (EMT)
  - Explicitly track whether conditions listed in the rule text have been fulfilled or not
- Coarse-to-fine (C2F) Reasoning
  - A coarse-to-fine approach to reason out which part of the rule text is underspecified, and ask a question accordingly
- Our proposed solution achieves new state-of-the-art results on the ShARC benchmark


Overall Process: 1. Encoding → 2. Explicit Memory Tracking → 3. Decision Making → 4. Question Generation
Rule Text
Dialog History
Rule Sentence 1
Rule Sentence 2
Rule Sentence 3
Initial Question
Scenario
Q1, A1
Q2, A2
1. Encoding
BERT Transformer Encoder
rule sent. feature
token score
entailment score ζi
start
end
Extracted Span
0.7
0.1
0.2
4. Question Generation
Rule Text
Span
UniLM
Follow-up Question
Coarse-to-Fine Reasoning for Span Extraction
entailment score for extraction
If 'Inquire'
Entailment Prediction
Rule Sent. 1
Rule Sent. 2
Rule Sent. 3
E : Entailment
C : Contradiction
U : Unknown
2. Explicit Memory Tracking
3. Decision Making
rule sent. score for decision
Decision Classifier
Inquire
Yes
No
Irrelevant

Proposed Solution

Encoding

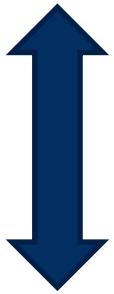


1. Parse the rule text into multiple rule sentences according to rules
2. Insert **[CLS]** token at the start of each rule sentence, initial question, scenario, and question-answer pairs in the dialog history
3. Concatenate all information and feed to BERT for encoding
4. **[CLS]** symbol is treated as the feature representation of the sentence that follows it

Proposed Solution

Explicit Memory Tracking

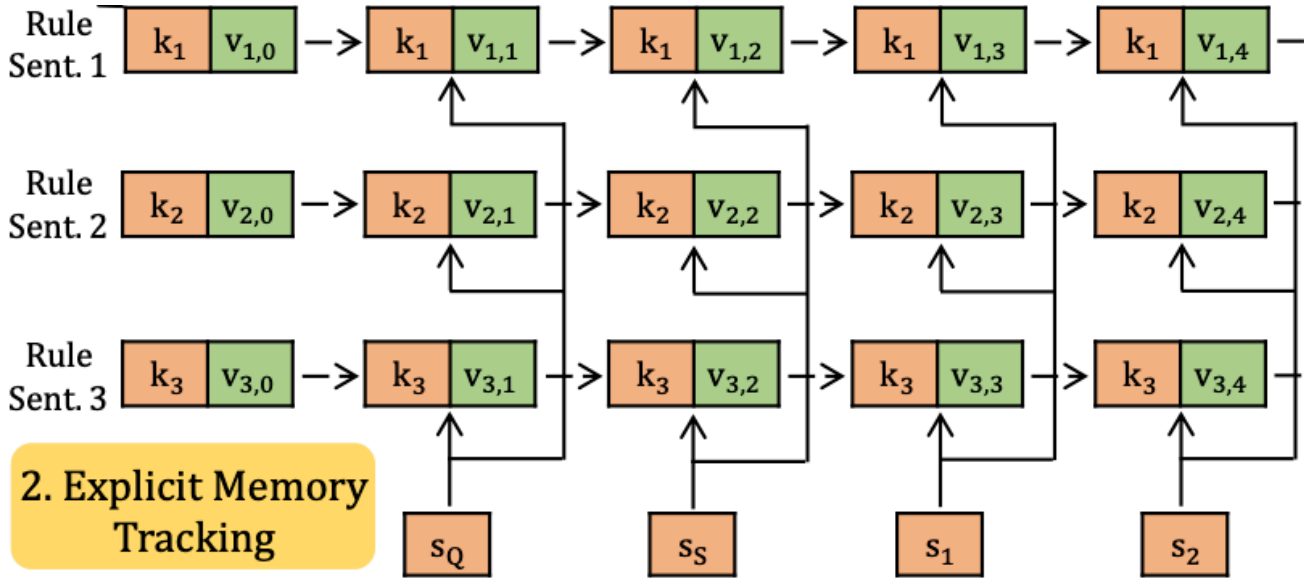
Rule sentences $\mathbf{k}_1, \mathbf{k}_2, \ldots, \mathbf{k}_M$



Find Implications

User provided information:

- Initial question $\mathbf{s}_Q$
- Scenario $\mathbf{s}_S$
- Dialog history $\mathbf{s}_1, \ldots, \mathbf{s}_P$



- We propose Explicit Memory Tracker (EMT), a gated recurrent memory-augmented neural network
- EMT explicitly tracks the states of rule sentences by sequentially reading the user provided information

Proposed Solution

Explicit Memory Tracking

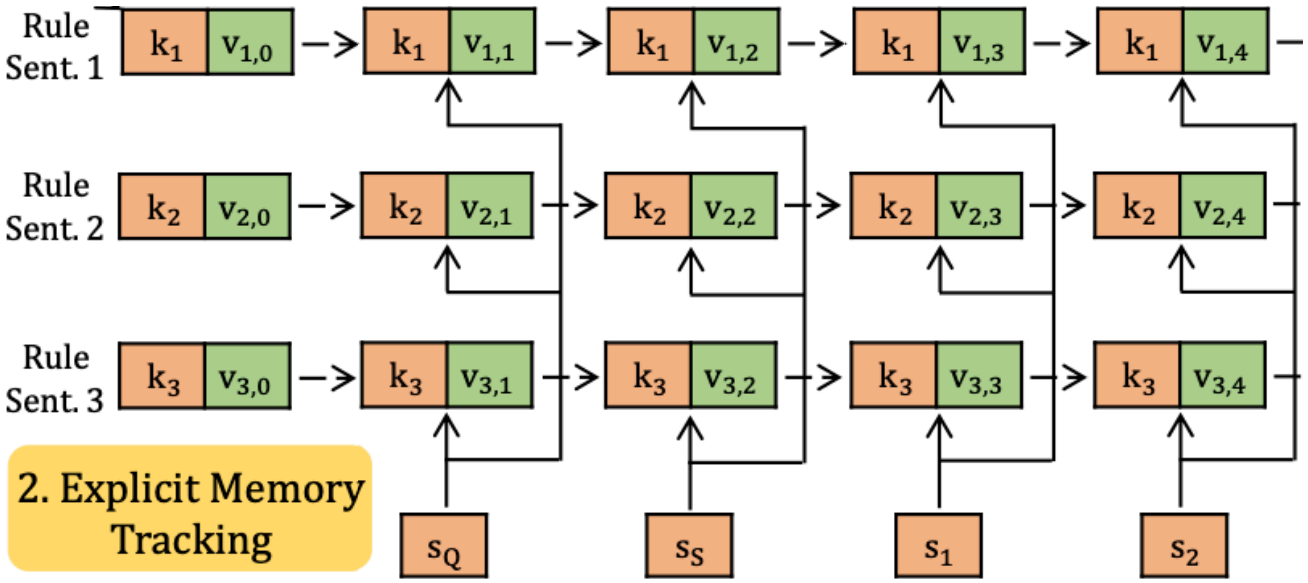
EMT assigns a state $\mathbf{v}_i$ to each key $\mathbf{k}_i$, and sequentially reads user information

At time step t:

$$\tilde{\mathbf{v}}_{i,t} = \mathbf{ReLU}(\mathbf{W}_k\mathbf{k}_i + \mathbf{W}_v\mathbf{v}_{i,t} + \mathbf{W}_s\mathbf{s}_t),$$

$$g_i = \sigma(\mathbf{s}_t^\top \mathbf{k}_i + \mathbf{s}_t^\top \mathbf{v}_{i,t}) \in [0,1],$$

$$\mathbf{v}_{i,t} = \mathbf{v}_{i,t} + g_i \odot \tilde{\mathbf{v}}_{i,t} \in \mathbb{R}^d, \mathbf{v}_{i,t} = \frac{\mathbf{v}_{i,t}}{\|\mathbf{v}_{i,t}\|}$$



Keys and final states of rule sentences are denoted as $(\mathbf{k}_1, \mathbf{v}_1), \ldots, (\mathbf{k}_M, \mathbf{v}_M)$

- Decision Making Module
- Question Generation Module

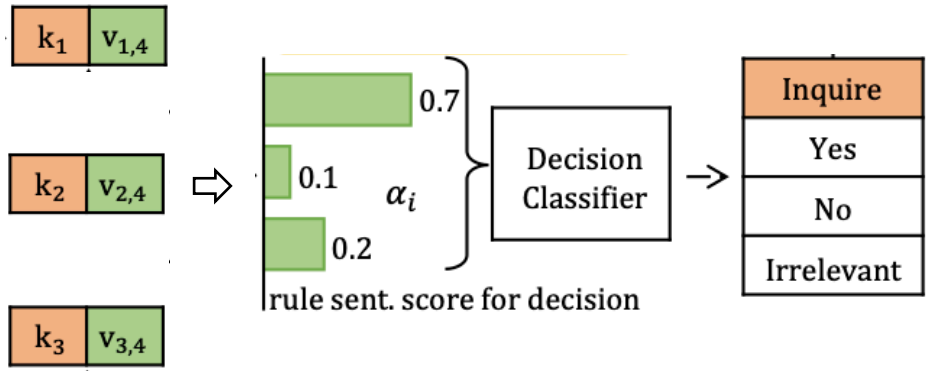
Proposed Solution
Decision Making

Based on the most up-to-date key-value states of rule sentences $(\mathbf{k}_1, \mathbf{v}_1), \ldots, (\mathbf{k}_M, \mathbf{v}_M)$, EMT makes a decision among *Yes, No, Irrelevant, Inquire*

$$\alpha_i = \mathbf{w}_\alpha^\top [\mathbf{k}_i; \mathbf{v}_i] + b_\alpha \in \mathbb{R}^1$$
$$\tilde{\alpha}_i = \text{softmax}(\alpha)_i \in [0, 1]$$
$$\mathbf{c} = \sum_i \tilde{\alpha}_i [\mathbf{k}_i; \mathbf{v}_i] \in \mathbb{R}^d$$
$$\mathbf{z} = \mathbf{W}_z \mathbf{c} + \mathbf{b}_z \in \mathbb{R}^4$$

The decision making module is trained with the following cross entropy loss:

$$\mathcal{L}_{\text{dec}} = -\log \ \text{softmax}(\mathbf{z})_l$$



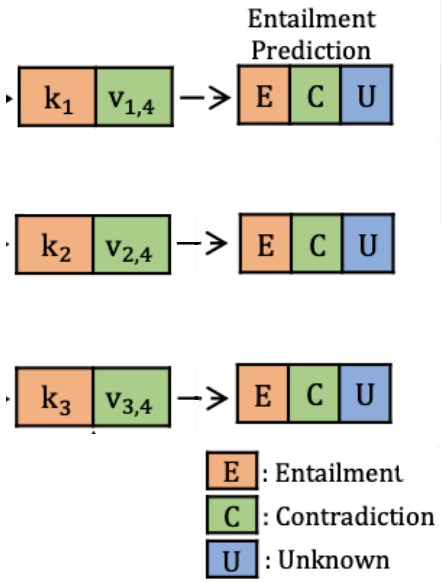
Proposed Solution

Subtask: Entailment State Prediction

- Explicitly track whether a condition listed in the rule has already been satisfied or not
- The possible entailment labels are:
  - Entailment (E)
  - Contradiction (C)
  - Unknown (U)

$$\mathbf{e}_i = \mathbf{W}_e[\mathbf{k}_i; \mathbf{v}_i] + \mathbf{b}_e \in \mathbb{R}^3$$

$$\mathcal{L}_{\text{entail}} = -\frac{1}{M}\sum_{i=1}^{M} \log \text{softmax}(\mathbf{e}_i)_r$$



Proposed Solution

Follow-up Question Generation

When the decision is ‘Inquire’, a follow-up question is required for further clarification.

We adopt a two-step approach:

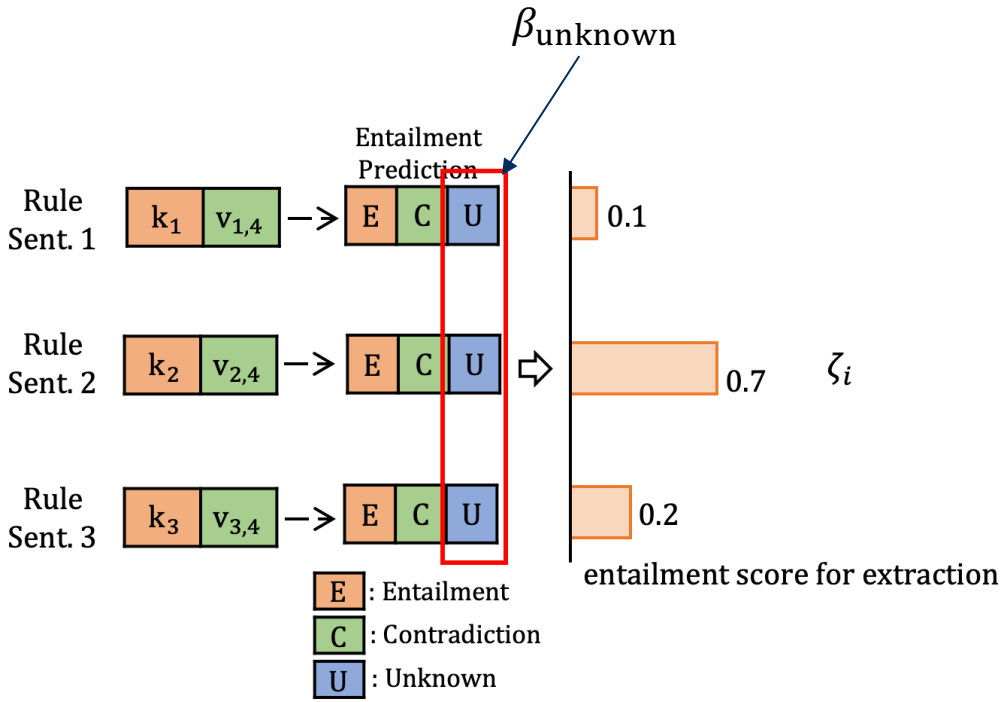
1. Extract a span inside the rule text which contains the underspecified user information
2. Rephrase the extracted span into a follow-up question

Proposed Solution

Follow-up Question Generation: Coarse-to-fine Underspecified Span Extraction

1. Coarse-to-fine Underspecified Span Extraction
   1) Identify underspecified rule sentence ζ_i

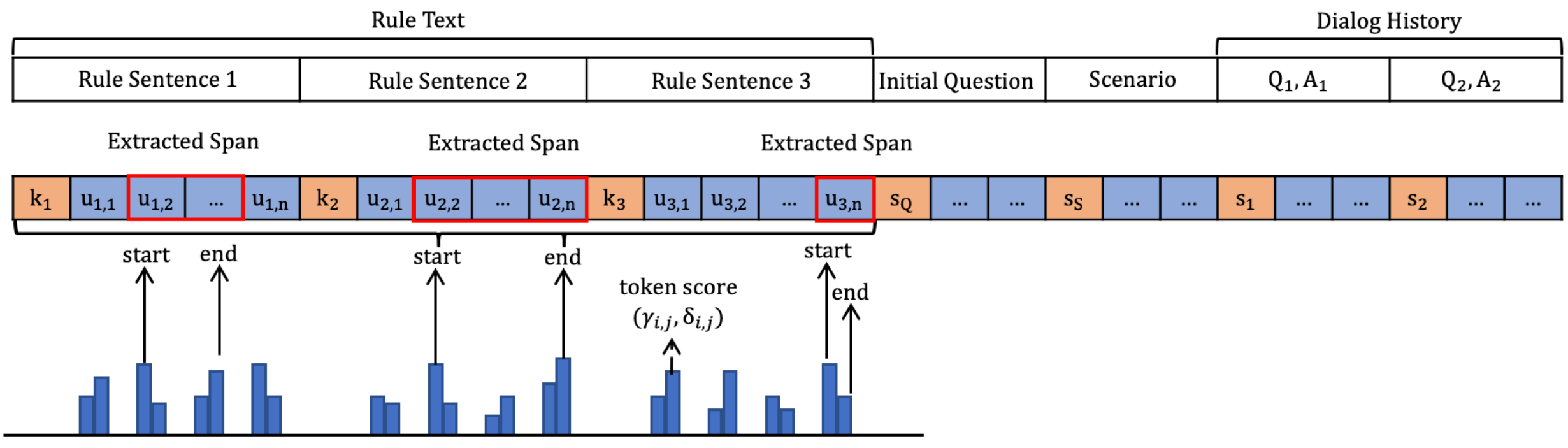
$$\zeta_i = \text{softmax}(\beta_{\text{unknown}})_i \in [0, 1]$$



Proposed Solution

Follow-up Question Generation: Coarse-to-fine Underspecified Span Extraction

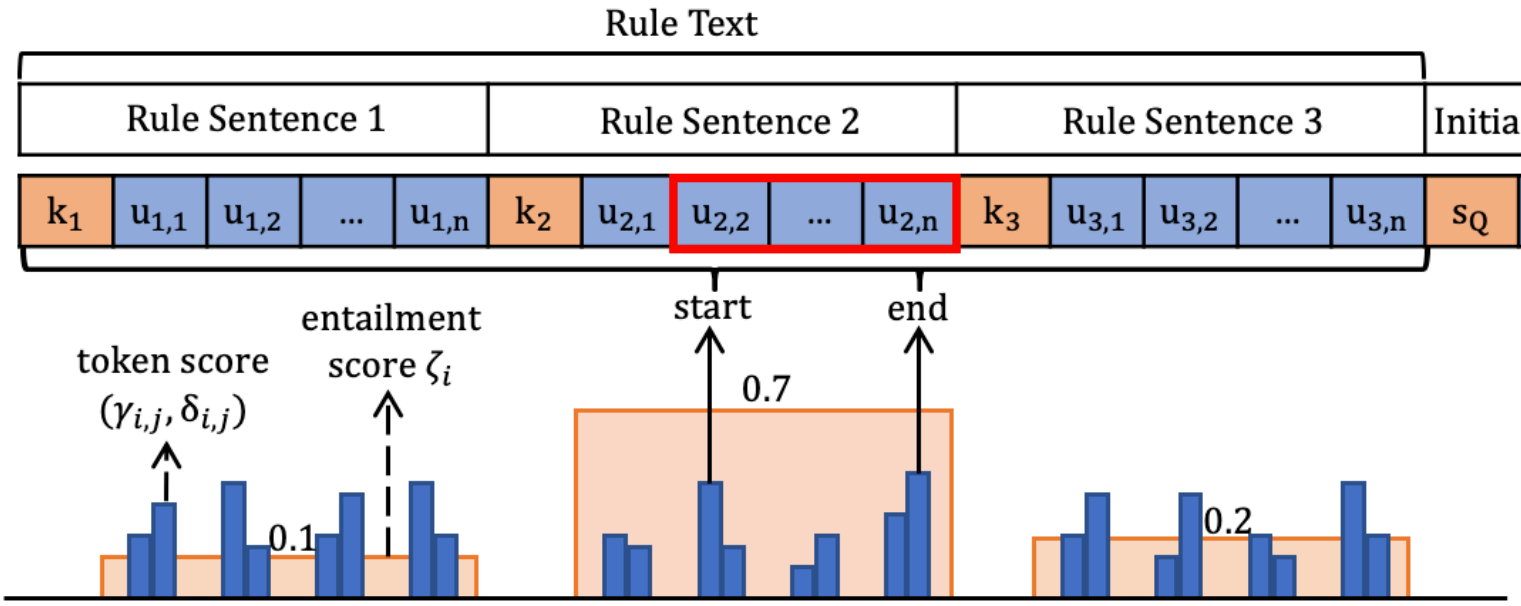
1. Coarse-to-fine Underspecified Span Extraction
   1) Identify underspecified rule sentence ζ_i
   2) Extract a span within each rule sentence $(\gamma_{i,j}, \delta_{i,j})$



Proposed Solution

Follow-up Question Generation: Coarse-to-fine Underspecified Span Extraction

1. Coarse-to-fine Underspecified Span Extraction
   1) Identify underspecified rule sentence ζ_i
   2) Extract a span within each rule sentence $(\gamma_{i,j}, \delta_{i,j})$
   3) Select the span with the highest span score $\zeta_i * (\gamma_{i,j}, \delta_{i,j})$



$$\gamma_{i,j} = \mathbf{w}_s^\top \mathbf{u}_{i,j} * \zeta_i, \quad \delta_{i,j} = \mathbf{w}_e^\top \mathbf{u}_{i,j} * \zeta_i$$

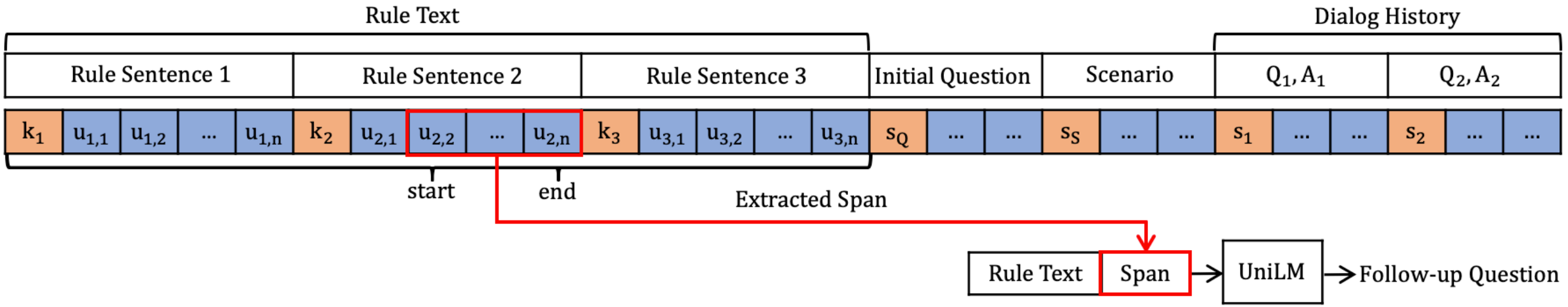
$$\mathcal{L}_{\text{span,s}} = -\mathbb{1}_{l=\text{inquire}} \log \text{softmax}(\gamma)_s$$

$$\mathcal{L}_{\text{span,e}} = -\mathbb{1}_{l=\text{inquire}} \log \text{softmax}(\delta)_e$$

Proposed Solution

Follow-up Question Generation

2. Question Rephrasing
   1) Finetune UniLM (Dong et al, 2019), a pretrained language model
   2) [CLS] rule text [SEP] span [SEP]



Proposed Solution

Overall Loss for EMT

The overall loss is the sum of the decision loss, entailment prediction loss and span extraction loss:

$$\mathcal{L} = \mathcal{L}_{\text{dec}} + \lambda_1 \mathcal{L}_{\text{entail}} + \lambda_2 \mathcal{L}_{\text{span}}$$

Experiments

Experimental Setup

- Dataset:
  - ShARC CMR dataset (Saeidi et
  - Train/Dev/Test dataset sizes are
  - Test set is not public.
  - Leaderboard: https://sharc-data.

ShARC: End-to-end Task

#	Model / Reference	Affiliation	Date	Micro Accuracy[%]	Macro Accuracy[%]	BLEU-1	BLEU-4
1	[Anonymous]	[Anonymous]	May 2020	**73.2**	**78.3**	**64.0**	49.1
2	EMT	Salesforce Research & CUHK	Nov 2019	69.4	74.8	60.9	46.0
3	EMT + entailment	Salesforce Research & CUHK	Mar 2020	69.1	74.6	63.9	**49.5**
4	[Anonymous]	[Anonymous]	Dec 2019	69.0	74.6	56.7	42.0
5	E3	University of Washington	Feb 2019	67.6	73.3	54.1	38.7
6	BiSon (single model)	NEC Laboratories Europe	Aug 2019	66.9	71.6	58.8	44.3

Experiments

Experimental Setup

- Dataset:
  - ShARC CMR dataset (Saeidi et
  - Train/Dev/Test dataset sizes are
  - Test set is not public.
  - Leaderboard: https://sharc-data.
- Evaluation Metrics
  - End-to-End Evaluation

ShARC: End-to-end Task

#	Model / Reference	Affiliation	Date	Micro Accuracy[%]	Macro Accuracy[%]	BLEU-1	BLEU-4
1	[Anonymous]	[Anonymous]	May 2020	**73.2**	**78.3**	**64.0**	49.1
2	EMT	Salesforce Research & CUHK	Nov 2019	69.4	74.8	60.9	46.0
3	EMT + entailment	Salesforce Research & CUHK	Mar 2020	69.1	74.6	63.9	**49.5**
4	[Anonymous]	[Anonymous]	Dec	69.0	74.6	56.7	42.0
6	BiSon (single model)	NEC Laboratories Europe	Aug 2019	66.9	71.6	58.8	44.3

If two models have different *Inquire* predictions, the follow-up questions for evaluation will be different, making the comparison unfair.

Experiments

Experimental Setup

- Dataset:
  - ShARC CMR dataset (Saeidi et al. 2018)
  - Train/Dev/Test dataset sizes are 21980/2270/8276.
  - Test set is not public.
  - Leaderboard: https://sharc-data.github.io/leaderboard.html
- Evaluation Metrics
  - End-to-End Evaluation
  - Oracle Question Generation Evaluation
    - We propose a new evaluation perspective.
    - We ask the models to generate follow-up questions *whenever* the ground truth decision is *Inquire*, and compute the BLEU score.

Experiments

Leaderboard Submission

Models	End-to-End Task (Leaderboard Performance)			
	Micro Acc.	Macro Acc.	BLEU1	BLEU4
Seq2Seq (Saeidi et al., 2018)	44.8	42.8	34.0	7.8
Pipeline (Saeidi et al., 2018)	61.9	68.9	54.4	34.4
BERTQA (Zhong and Zettlemoyer, 2019)	63.6	70.8	46.2	36.3
UrcaNet (Sharma et al., 2019)	65.1	71.2	60.5	46.1
BiSon (Lawrence et al., 2019)	66.9	71.6	58.8	44.3
E^3 (Zhong and Zettlemoyer, 2019)	67.6	73.3	54.1	38.7
EMT (our single model)	**69.1**	**74.6**	**63.9**	**49.5**

Table 1: Performance on the blind, held-out test set of ShARC end-to-end task.

Experiments

Class-wise Decision Prediction Accuracy

Models	Yes	No	Inquire	Irrelevant
BERTQA	61.2	61.0	62.6	96.4
E^3	65.9	70.6	60.5	96.4
UrcaNet*	63.3	68.4	58.9	95.7
EMT	**70.5**	**73.2**	**70.8**	**98.6**

Table 2: Class-wise decision prediction accuracy on the development set (*: reported in the paper).

Experiments

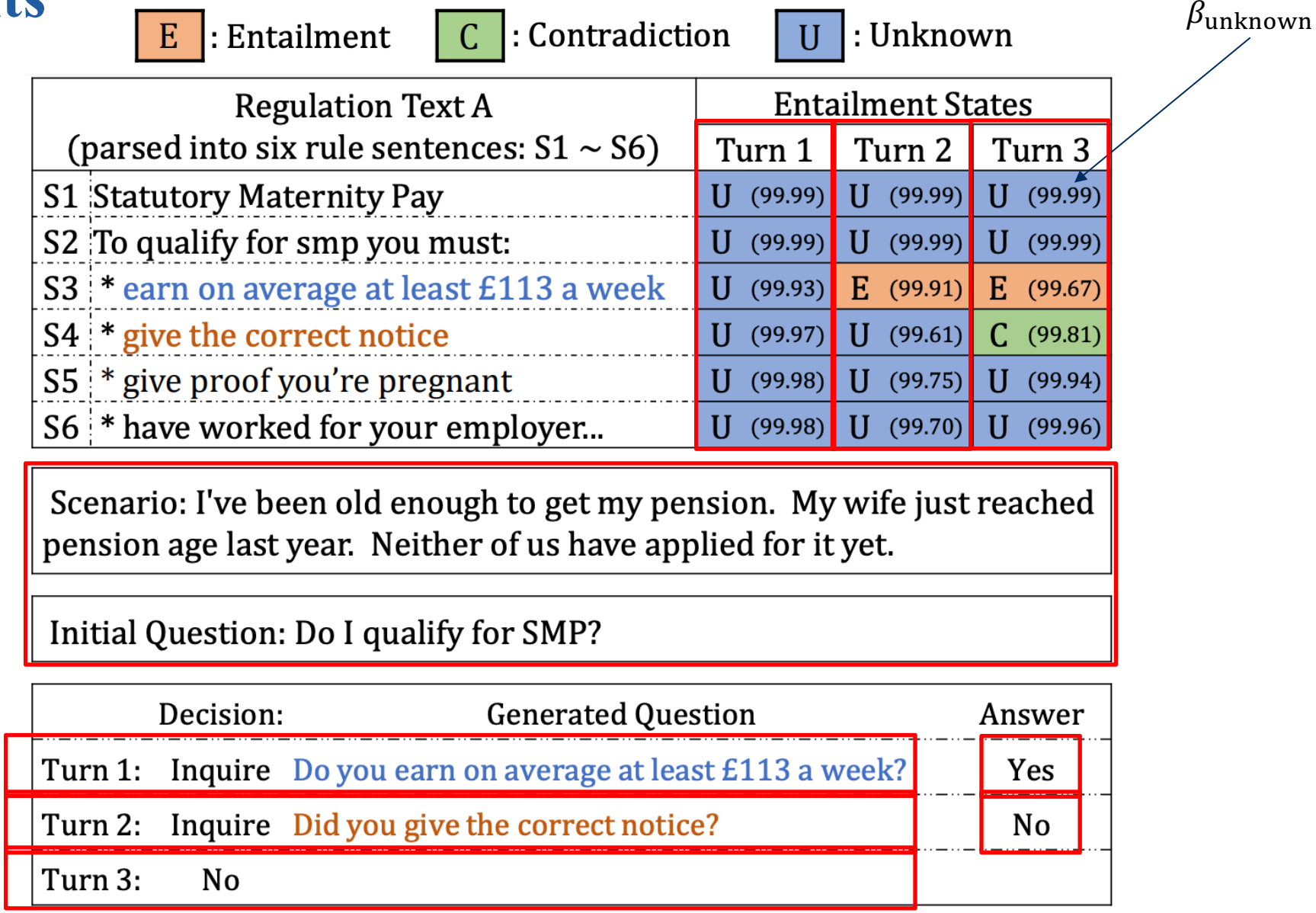
Oracle Question Generation Task

Models	Oracle Question Generation Task			
	Development Set		Cross Validation	
	BLEU1	BLEU4	BLEU1	BLEU4
E^3	52.79±2.87	37.31±2.35	51.75	35.94
E^3+UniLM	57.09±1.70	41.05±1.80	56.94	42.87
EMT	**62.32**±1.62	**47.89**±1.58	**64.48**	**52.40**

Table 3: Performance on Oracle Question Generation Task. We show both results on the development set and 10-fold cross validation. E^3+UniLM replaces the editor of E^3 to our finetuned UniLM.

Experiments

Interpretability



Conclusion

- We propose a new approach called Explicit Memory Tracker (EMT) for conversational machine reading.
- EMT achieved a new state-of-the-art result on the ShARC CMR challenge.
- EMT also gains interpretability by showing the entailment-oriented reasoning process as the conversation flows.

Discern: Discourse-Aware Entailment Reasoning Network for Conversational Machine Reading

Yifan Gao[1] Chien-Sheng Wu[2] Jingjing Li[1] Shafiq Joty[2,3]
Steven C.H. Hoi[2] Caiming Xiong[2] Irwin King[1] Michael R. Lyu[1]

1. The Chinese University of Hong Kong
2. Salesforce Research 3. Nanyang Technological University

Code & Models: https://github.com/Yifan-Gao/Discern

Challenges

1. Document Interpretation
   - Identification of Conditions
   - Determination of Logical Structures

7(a) loans are the most basic and most used type loan of the Small Business Administration's (SBA) business loan programs. Its name comes from section 7(a) of the Small Business Act, which authorizes the agency to provide business loans to ① American small businesses. The loan program is designed to assist ② for-profit businesses that are ③ not able to get other financing from other resources.

"Eligible for 7(a) loans" = (①==True) and (② ==True) and (③ ==True)

Challenges

2. Dialog Understanding
   - Track the user's fulfillment over the conditions
   - Jointly consider the fulfillment states and the logical structure of rules

7(a) loans are the ... provide business loans to ① American small businesses. The loan program is designed to assist ② for-profit businesses that are ③ not able to get other financing from other resources.

"Eligible for 7(a) loans" = (①==True) and (② ==True) and (③ ==True) ⇨ **AND**: examine the fulfillments of all conditions

Scenario: I am a 34-year-old man from the United States who owns their own business. We are an American small business.

⇨ **① == True**

Fulfillment State: (①==True) and (② == ?) and (③ == ?) ⇨ **Decision: Inquire**

Follow-up Q1: Are you a for-profit business?

Challenges

2. Dialog Understanding
   - Track the user's fulfillment over the conditions
   - Jointly consider the fulfillment states and the logical structure of rules

7(a) loans are the ... provide business loans to ① American small businesses. The loan program is designed to assist ② for-profit businesses that are ③ not able to get other financing from other resources.

"Eligible for 7(a) loans" = (①==True) and (② ==True) and (③ ==True) ⇨ **AND**: examine the fulfillments of all conditions

Scenario: I am a 34-year-old man from the United States who owns their own business. We are an American small business. ⇨ **① == True**

Follow-up Q1: Are you a for-profit business? A1: Yes ⇨ **② == True**

Fulfillment State: (①==True) and (② == True) and (③ == ?) ⇨ **Decision: Inquire**

Follow-up Q2: Are you able to get financing from other resources?

Challenges

2. Dialog Understanding
   - Track the user's fulfillment over the conditions
   - Jointly consider the fulfillment states and the logical structure of rules

7(a) loans are the ... provide business loans to ① American small businesses. The loan program is designed to assist ② for-profit businesses that are ③ not able to get other financing from other resources.

"Eligible for 7(a) loans" = (①==True) and (② ==True) and (③ ==True) ⇨ **AND**: examine the fulfillments of all conditions

Scenario: I am a 34-year-old man from the United States who owns their own business. We are an American small business. ➡ **① == True**

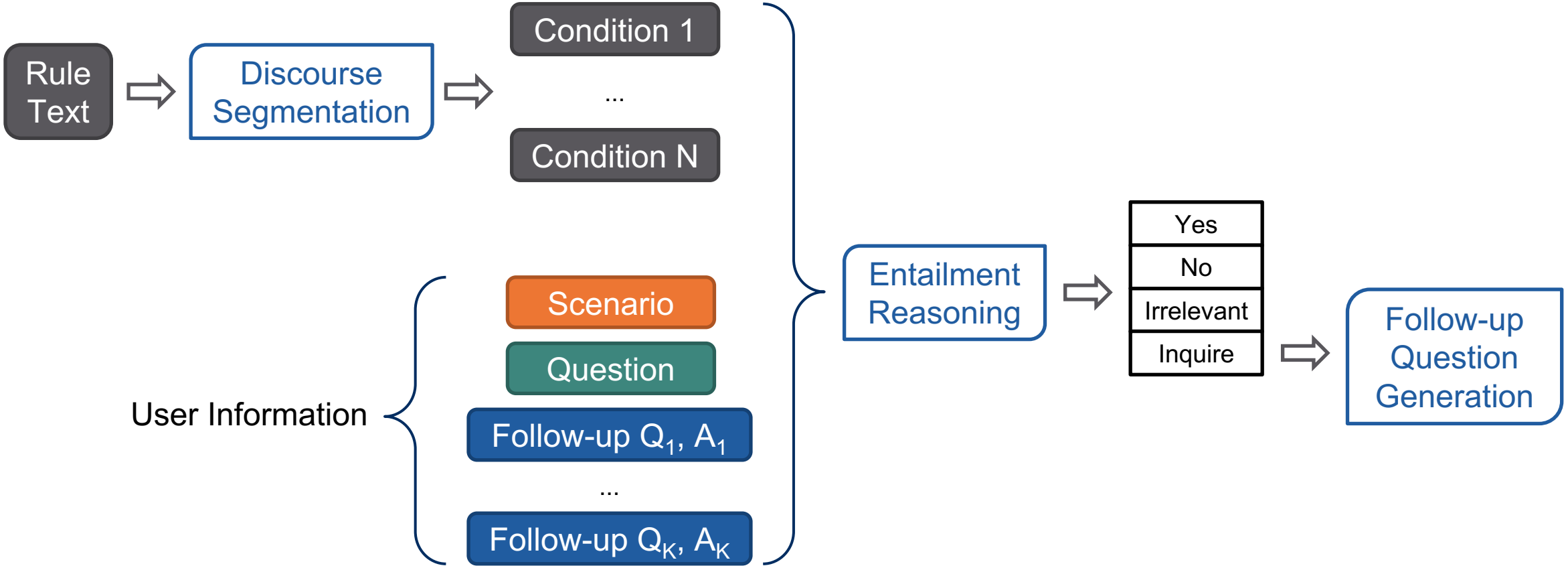
Follow-up Q1: Are you a for-profit business? A1: Yes ➡ **② == True**

Follow-up Q2: Are you able to get financing from other resources? A2: No ➡ **③ == True**

Fulfillment State: (①==True) and (② == True) and (③ == True ⇨ **Decision: Yes** (You can apply the loan.)

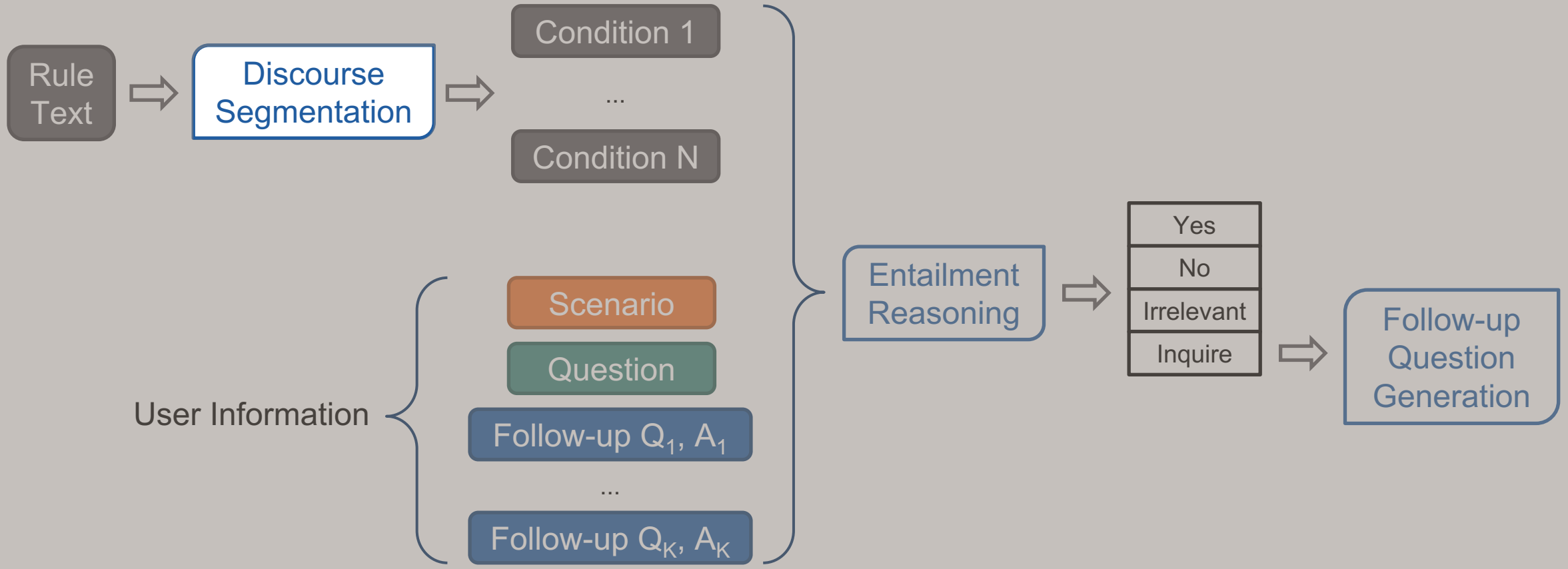
Discern: Discourse-Aware Entailment Reasoning Network for Conversational Machine Reading

Overview



Discern: Discourse-Aware Entailment Reasoning Network for Conversational Machine Reading

Overview



Discern: Discourse-Aware Entailment Reasoning Network for Conversational Machine Reading

Rule Segmentation

- Goal
  - Understand the logical structure of the rule text
  - Parse the rule into individual conditions for entailment reasoning
- Challenges
  - Sentence splitting is not enough: one rule sentence may contain several in-line conditions
- Solution: Discourse Segmentation
  - In the Rhetorical Structure Theory (RST) of discourse parsing (Mann and Thompson, 1988), texts are split into clause-like units called elementary discourse units (EDUs)

Rule Text: If a worker has taken more leave than they're entitled to, their employer must not take money from their final pay unless it's been agreed beforehand in writing.

↓

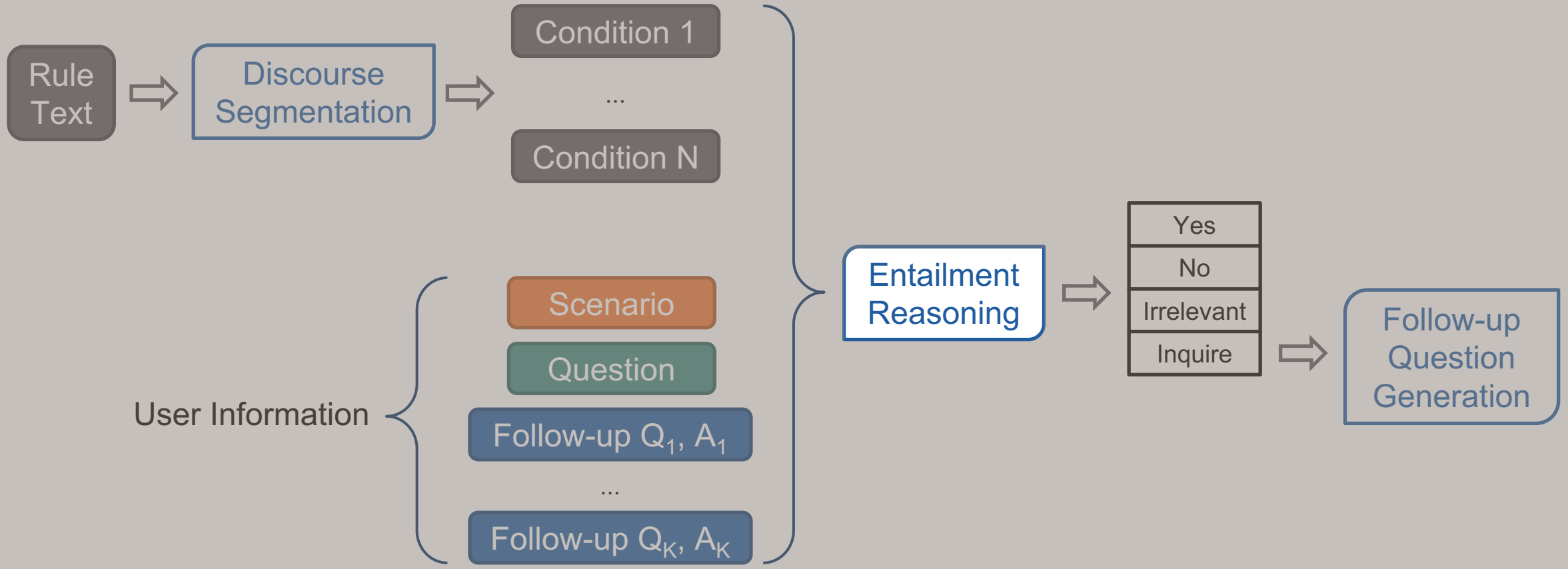
Discourse Segmentation

↓

[If a worker has taken more leave than they're entitled to,]$_{\text{EDU1}}$ [their employer must not take money from their final pay]$_{\text{EDU2}}$ [unless it's been agreed beforehand in writing.]$_{\text{EDU3}}$

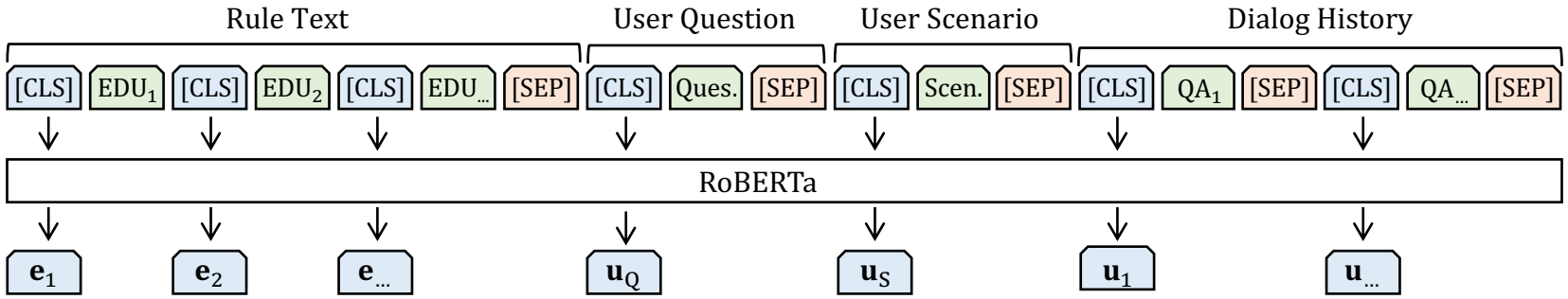
Discern: Discourse-Aware Entailment Reasoning Network for Conversational Machine Reading

Overview



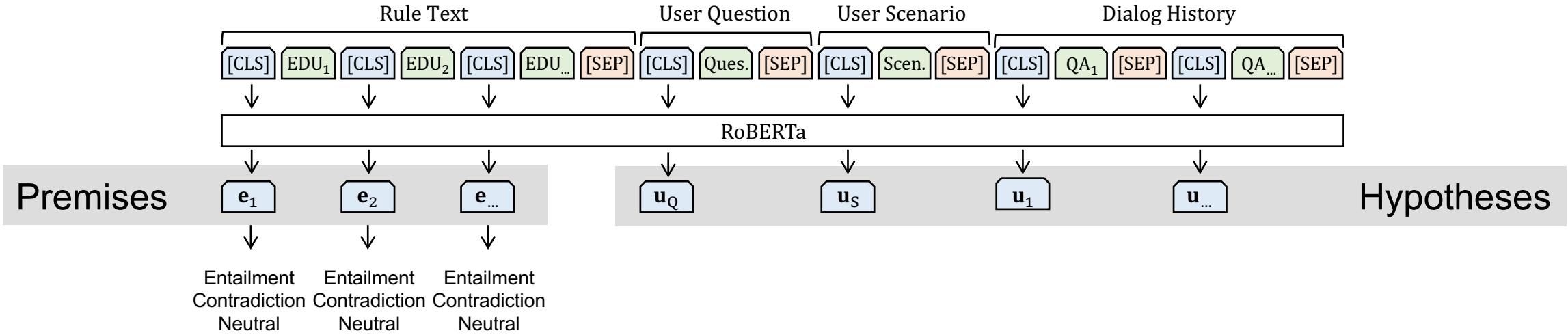
Discern: Decision Making via Entailment Reasoning

Encoding



Discern: Decision Making via Entailment Reasoning

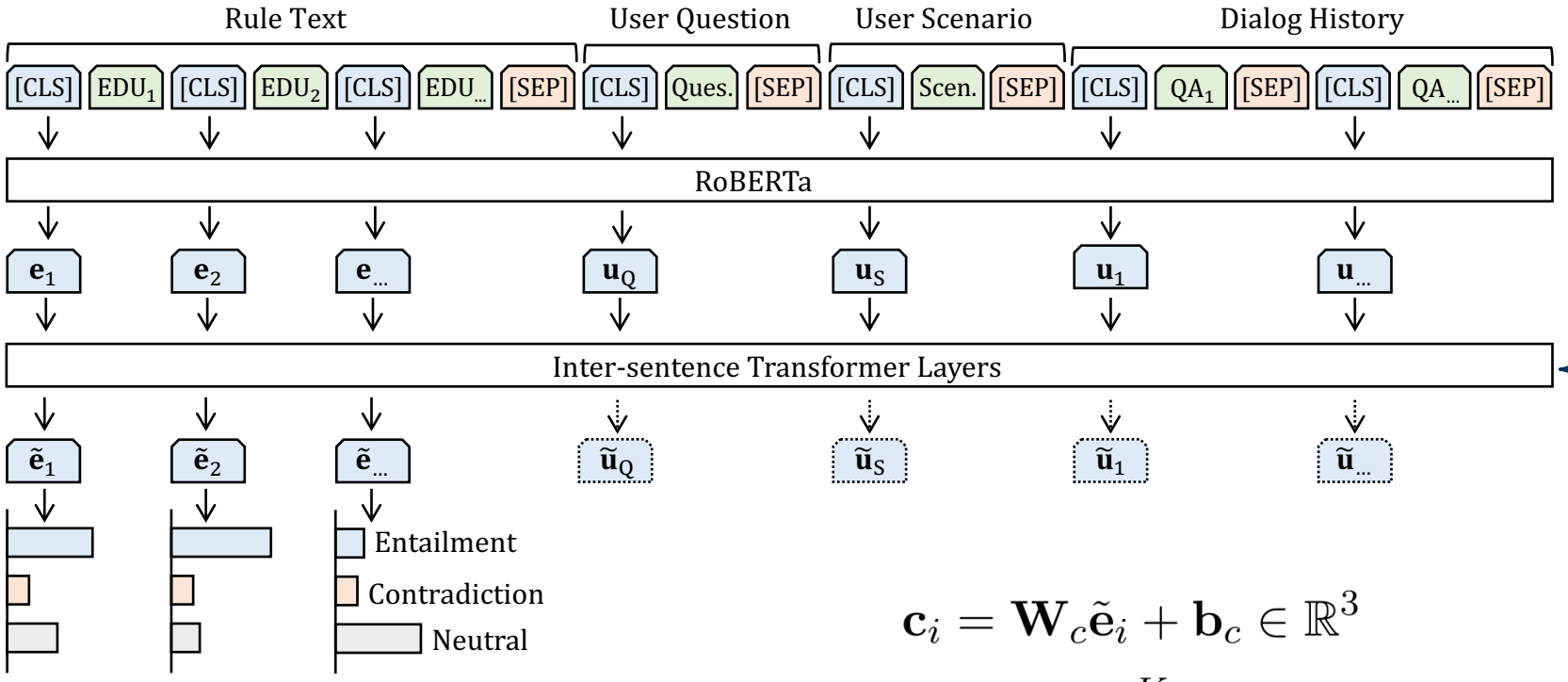
Entailment Prediction



Multi-Sentence Entailment Prediction

Discern: Decision Making via Entailment Reasoning

Entailment Prediction

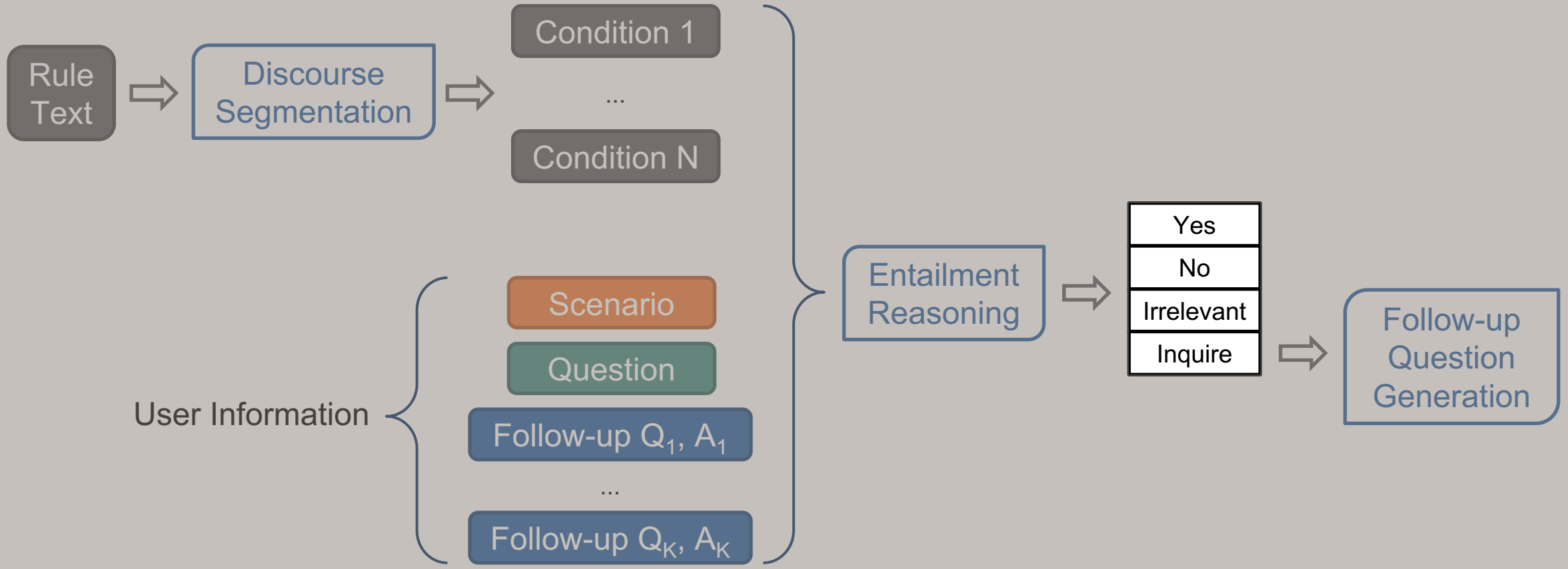


$$\mathbf{c}_i = \mathbf{W}_c \tilde{\mathbf{e}}_i + \mathbf{b}_c \in \mathbb{R}^3$$

$$\mathcal{L}_{\text{entail}} = -\frac{1}{K} \sum_{i=1}^{K} \log \text{softmax}(\mathbf{c}_i)_r$$

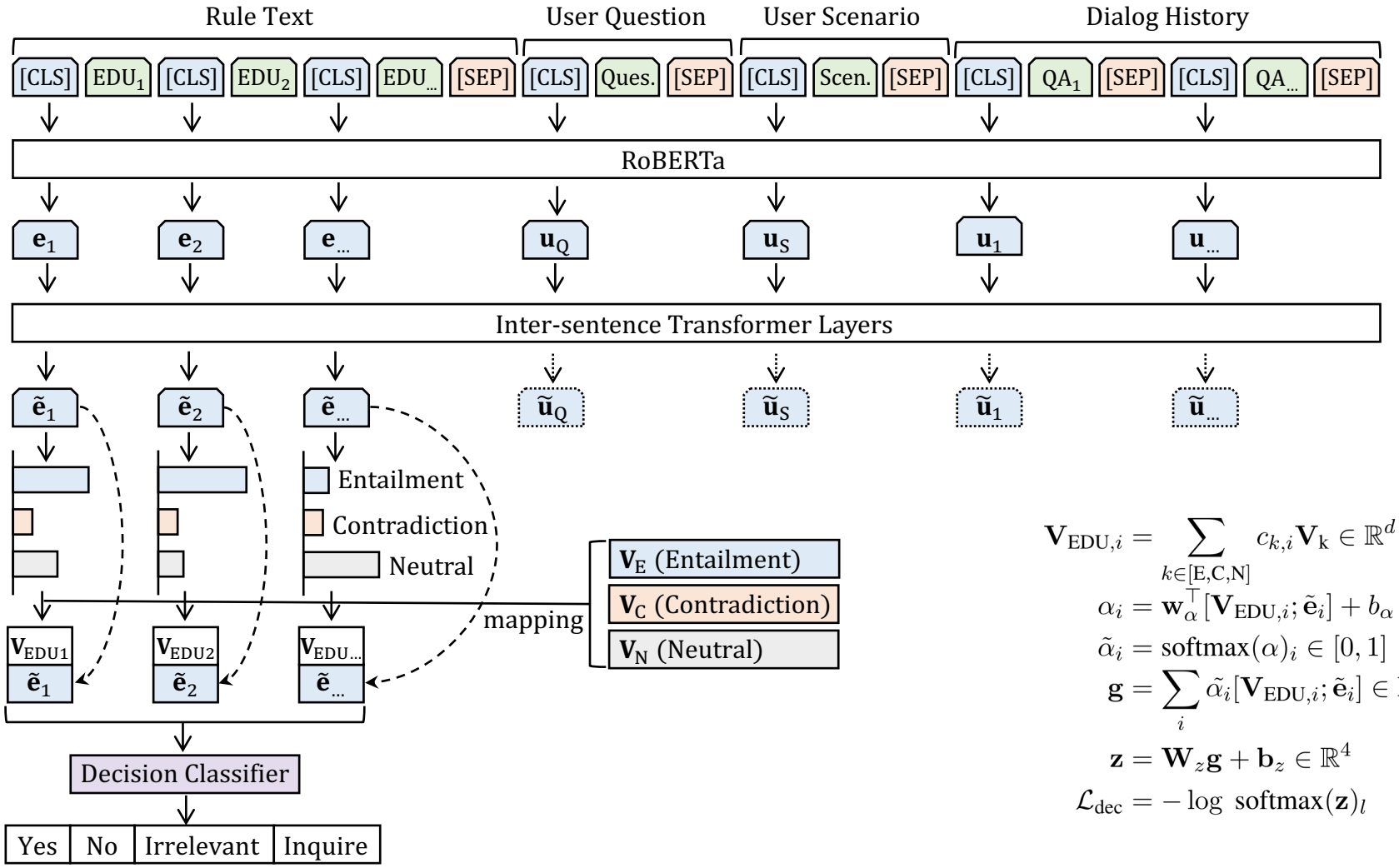
Discern: Discourse-Aware Entailment Reasoning Network for Conversational Machine Reading

Overview



Discern: Discourse-Aware Entailment Reasoning Network for Conversational Machine Reading

Decision Making



$$\mathbf{V}_{\text{EDU},i} = \sum_{k\in[\text{E,C,N}]} c_{k,i}\mathbf{V}_{\text{k}} \in \mathbb{R}^d$$
$$\alpha_i = \mathbf{w}_\alpha^\top[\mathbf{V}_{\text{EDU},i};\tilde{\mathbf{e}}_i] + b_\alpha \in \mathbb{R}^1$$
$$\tilde{\alpha}_i = \text{softmax}(\alpha)_i \in [0,1]$$
$$\mathbf{g} = \sum_i \tilde{\alpha}_i[\mathbf{V}_{\text{EDU},i};\tilde{\mathbf{e}}_i] \in \mathbb{R}^{2d}$$
$$\mathbf{z} = \mathbf{W}_z\mathbf{g} + \mathbf{b}_z \in \mathbb{R}^4$$
$$\mathcal{L}_{\text{dec}} = -\log\ \text{softmax}(\mathbf{z})_l$$

Experiments

Main Results

Models	End-to-End Task (Leaderboard Performance)			
	Micro Acc.	Macro Acc.	BLEU1	BLEU4
Seq2Seq (Saeidi et al., 2018)	44.8	42.8	34.0	7.8
Pipeline (Saeidi et al., 2018)	61.9	68.9	54.4	34.4
BERTQA (Zhong and Zettlemoyer, 2019)	63.6	70.8	46.2	36.3
UrcaNet (Sharma et al., 2019)	65.1	71.2	60.5	46.1
BiSon (Lawrence et al., 2019)	66.9	71.6	58.8	44.3
E^3 (Zhong and Zettlemoyer, 2019)	67.6	73.3	54.1	38.7
EMT (Gao et al., 2020)	69.4	74.8	60.9	46.0
EMT+entailment (Gao et al., 2020)	69.1	74.6	63.9	**49.5**
DISCERN (our single model)	**73.2**	**78.3**	**64.0**	49.1

Experiments

Ablation Study

Models	Micro Acc.	Macro Acc.
DISCERN	74.97±0.27	79.55±0.35
DISCERN (BERT)	73.07±0.21	77.77±0.24
DISCERN (w/o EDU)	73.34±0.22	78.25±0.57
DISCERN (w/o Trans)	74.25±0.36	78.78±0.57
DISCERN (w/o $\tilde{\mathbf{e}}$)	73.55±0.26	78.19±0.30
DISCERN (w/o $\mathbf{V}_{\text{EDU}}$)	72.95±0.23	77.53±0.19

➡ RoBERTa > BERT

➡ Discourse Segmentation > Sentence Splitting

➡ Inter-Sentence Transformer **IS** Necessary!

} Both Condition Representations and Entailment Vectors Facilitate Decisions

Experiments

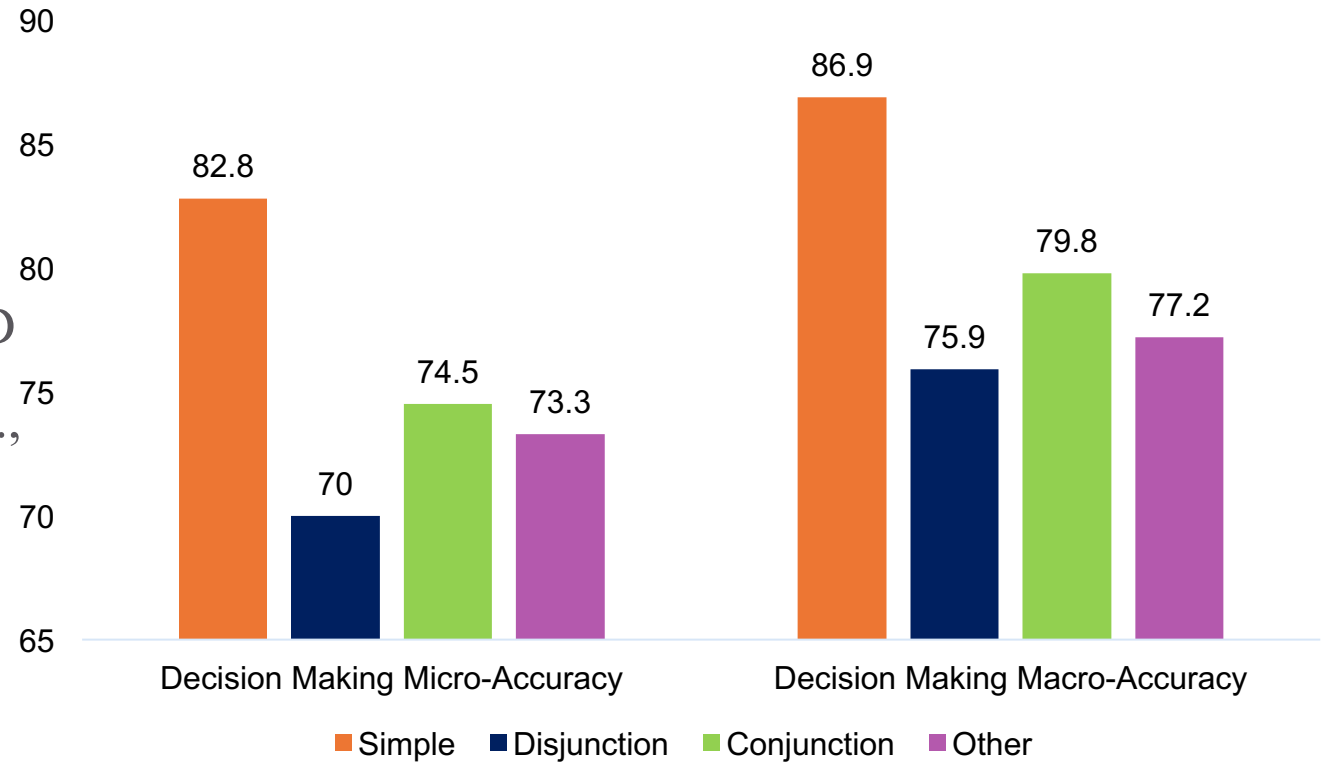
Analysis of Logical Structure of Rules

Simple: If A, then B

Disjunction: If A OR B OR C OR ..., then D

Conjunction: If A AND B AND C AND ..., then D

Other: Complicated AND/OR relationship

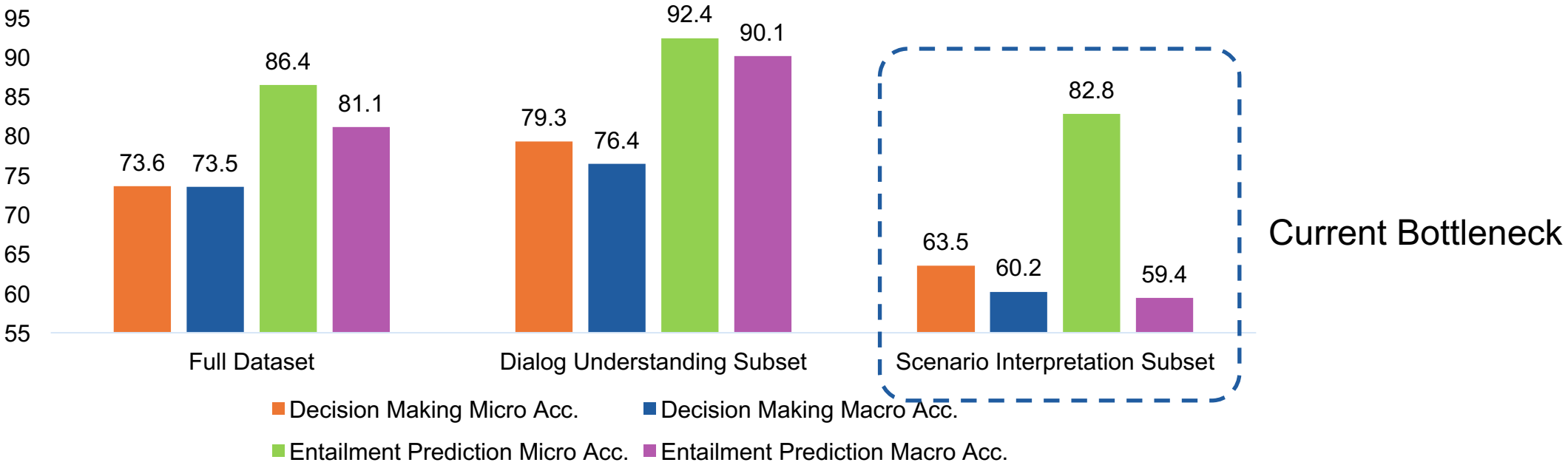


Experiments

How Far Has the Problem Been Solved?

Idea: Disentangle the challenge between scenario interpretation and dialog understanding

- **Full Dataset**: Full development set of ShARC
- **Dialog Understanding Subset**: User information only contains dialog history
- **Scenario Interpretation Subset**: User information only contains user scenario



Conclusion

- We present Discern, a system that does discourse-aware entailment reasoning for conversational machine reading.
- Results on the ShARC benchmark shows that Discern outperforms existing methods by a large margin.
- We also conduct comprehensive analyses to unveil the limitations of Discern and challenges for ShARC.

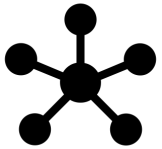
Thanks!

Keep CALM and Explore: Language Models for Action Generation in Text-based Games

Shunyu Yao, Rohan Rao,
Matthew Hausknecht,
Karthik Narasimhan

EMNLP 2020

RL view



- Discrete MDP
- Millions of states
- Billions of actions
- Very sparse reward

Observation: **West of House** You are standing in an open field west of a white house, with a boarded front door. There is a small mailbox here.

Action: **Open mailbox**

Observation: Opening the small mailbox reveals a leaflet.

Action: **Read leaflet**

Observation: (Taken) "WELCOME TO ZORK! ZORK is a game of adventure, danger, and low cunning. In it you will explore some of the most amazing territory ever seen by mortals. No computer should be without one!"

Action: **Go north**

Observation: **North of House** You are facing the north side of a white house. There is no door here, and all the windows are boarded up. To the north a narrow path winds through the trees.

Human view

- Text adventure
- Places and rooms
- Some valid actions
- Commonsense

Difference: Language understanding!

Text-based Game

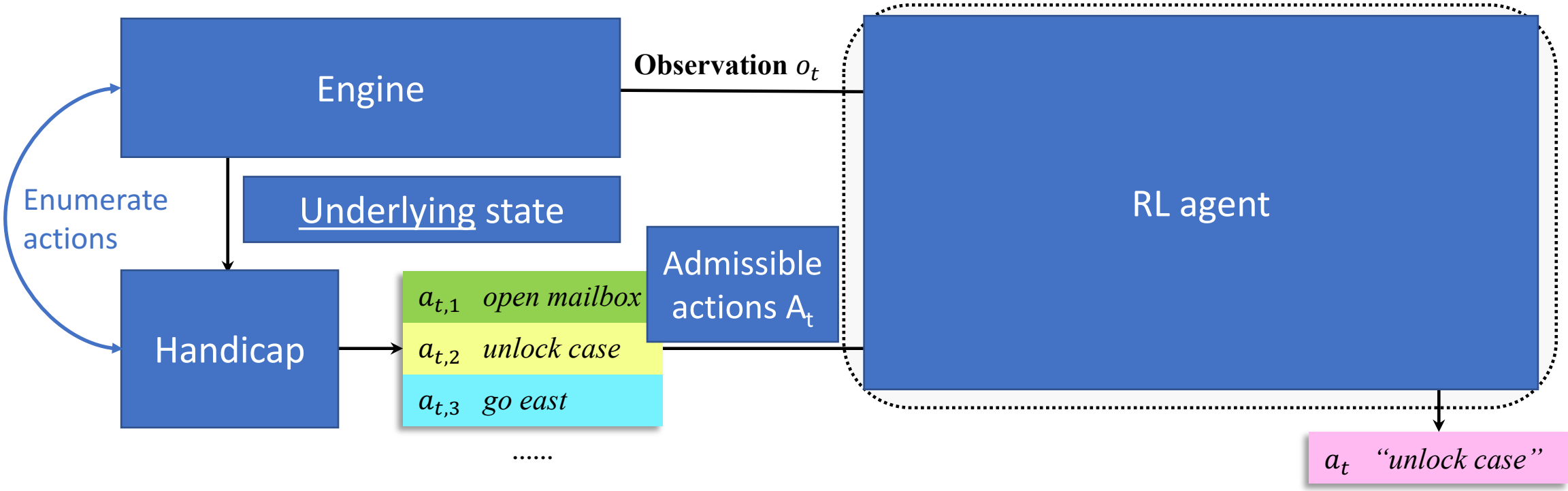
- A partially observable Markov decision process (POMDP) where a player
  - Issues text actions
  - Receives text observations and scalar rewards
- Immediate challenge: **action space**
  - Random actions: billions, mostly not admissible
  - Previous models: heuristic rules or action *handicap*
    - *Problem: ad-hoc, hard to generalize*
  - Our approach: train language models for action generation
    - Scalable, generalizable

Observation: You are in the living room. There is a doorway to the east, a wooden door with strange gothic lettering to the west, which appears to be nailed shut, a trophy case, and a large oriental rug in the center of the room. You are carrying: A brass lantern . . .

Random Actions:
close door, north a, eat troll with egg, . . .
CALM (n-gram) Actions:
enter room, leave room, lock room,
open door, close door, knock on door, . . .
CALM (GPT-2) Actions:
east, open case, get rug, turn on lantern,
move rug, unlock case with key, . . .

Next Observation: With a great effort, the rug is moved to one side of the room, revealing the dusty cover of a closed trap door...

Previous Approach: Language Model (CALM)



Action candidates: handicap → language model

1. **Fair** (no forbidden functions used)
2. **Generalize** (to unsupported games)

DRRN: He, et al. "Deep Reinforcement Learning with a Natural Language Action Space."

How to train CALM

o_t : observation at time t
a_t : action at time t

- Train CALM on human gameplay trajectories of form $(o_1,a_1,o_2,a_2,\ldots)$, predict action a_t based on context $c_t=(o_{t-1},a_{t-1},o_t)$
- Two language models:
  1. **GPT-2**: standard cross entropy loss on action tokens

$$\mathcal{L}_{\mathrm{LM}}(\theta) = -\mathbb{E}_{(a,c)\sim D} \log p_\theta(a|c) \qquad p_\theta(a|c) = \prod_{i=1}^{m} p_\theta(a^i|a^{<i}, c)$$

  2. **N-gram**:
     - **Probability**: only base on action n-grams, independent of context
     - **Generation**: only consider actions with objects seen in observation

How to use CALM in RL

- Given context c, use **trained, frozen** CALM to produce top-k candidates (k=30)
- RL as action re-ranker to maximize game performance
- **Rationale**: combine CALM's generic action priors and RL's ability to optimize gameplay performance

Observation: You are in the living room. There is a doorway to the east, a wooden door with strange gothic lettering to the west, which appears to be nailed shut, a trophy case, and a large oriental rug in the center of the room. You are carrying: A brass lantern . . .

Random Actions:
close door, north a, eat troll with egg, . . .
CALM (n-gram) Actions:
enter room, leave room, lock room,
open door, close door, knock on door, . . .
CALM (GPT-2) Actions:
east, open case, get rug, turn on lantern,
move rug, unlock case with key, . . .

Next Observation: With a great effort, the rug is moved to one side of the room, revealing the dusty cover of a closed trap door...

Data

	ClubFloyd Dataset	Jericho Walkthroughs
# unique games	590	28
Vocab size	39,670	9,623
Vocab size (game avg.)	2,363	1,037
Avg. trajectory length	360	98
Action Quality	Non-optimal	Optimal

ClubFloyd Dataset

- **Used for CALM training**
- 590 games, 223,527 samples
- Human actions: noisy, non-optimal, sometimes invalid
- A few very common actions (*north*, *take all*, *examine*, …)

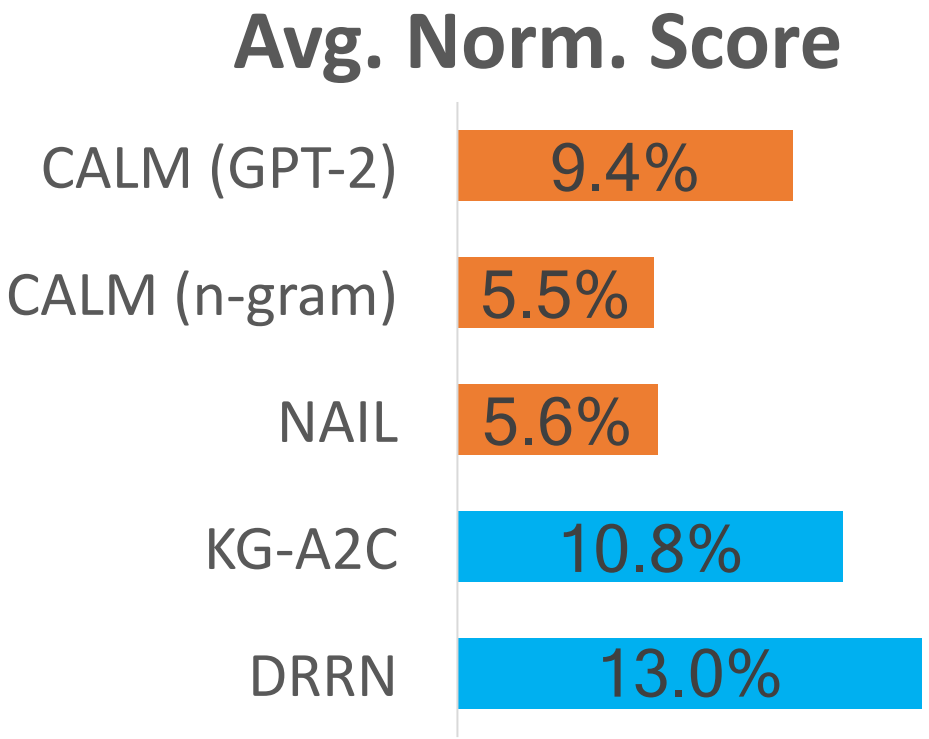
Jericho Benchmark

- **Used to evaluate CALM + RL**
- 28 games with simulator access
- Admissible action handicap at each step
- Optimal walkthrough for each game

No overlapping games!

Evaluate CALM with RL

- Baselines:
  - **NAIL**: ad-hoc rules, no RL or handicap
  - **DRRN**: need handicap for action space
  - **KG-A2C**: need handicap for action generation supervision, use KG
- **No handicap**: GPT-2 > N-gram ~ NAIL
- **Handicap models**: GPT-2 outperforms both DRRN and KG-A2C on 8/28 games
  - No “stupid” admissible actions
  - Handicap might miss some valid actions



NAIL: Hausknecht, et al. ”Nail: A general interactive fiction agent.”

KG-A2C: Ammanabrolu, Hausknecht. “Graph constrained reinforcement learning for natural language action spaces.”

Example *Zork I* walkthrough state

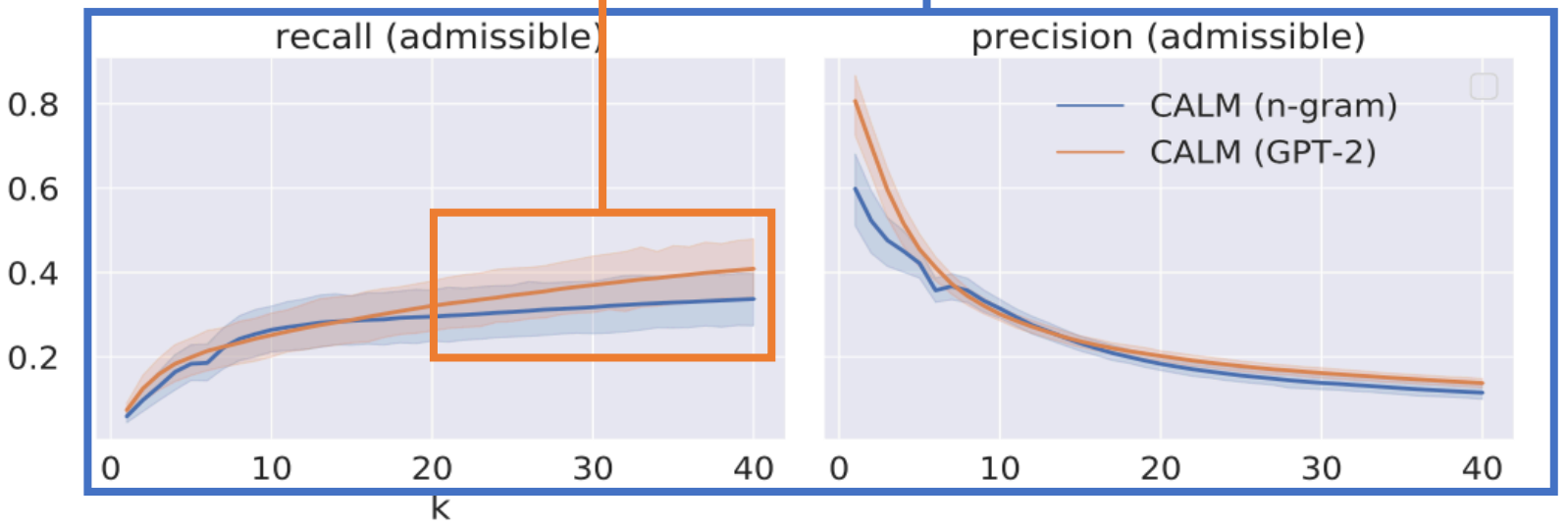
- **Context:** [CLS] with a great effort, the rug is moved to one side of the room, revealing the dusty cover of a closed trap door. [SEP] open trapdoor [SEP] the door reluctantly opens to reveal a rickety staircase descending into darkness. you are carrying: a sword a nasty knife a rope a brass lantern a clove of garlic a jewel-encrusted egg living room you are in the living room. there is a doorway to the east, a wooden door with strange gothic lettering to the west, which appears to be nailed shut, a trophy case, and a rug lying beside an open trap door. [SEP]
- **GPT-2:** [east, west, down, up, north, south, open trophy case, wait, knock on door, take rug, southeast, enter trapdoor, out, drop sword, take rope, in, southwest, northwest, get rope, open case, get rug, search rug, enter trap, climb rope, northeast, take sword, move rug, take all, put sword in trapdoor, close trapdoor]
- **n-gram:** [north, east, south, west, up, down, open door, examine door, take all, unlock door, get all, close door, drop all, put all, tie rope, examine knife, take knife, examine case, examine sword, open case, examine rope, examine west, take rope, take sword, examine lantern, put knife, pull rope, take lantern, examine egg, put sword, get sword, put egg]
- **Admissible actions (generated by handicap enumeration):** [east, open egg with lantern, throw rope at egg, throw egg at knife, throw sword at egg, throw garlic at egg, throw lantern at egg, throw knife at egg, throw knife at lantern, put down all, put down rope, put down egg, put down sword, put down garlic, put down lantern, put down knife, close trap, take on egg, open case, turn on lantern, down]
- **Walkthrough action:** down

CALM Analysis: Admissibility of CALM Actions

- Evaluate admissibility of CALM actions on Jericho walkthroughs
- Precision and recall of admissible actions among top-k CALM actions (k=1…40)

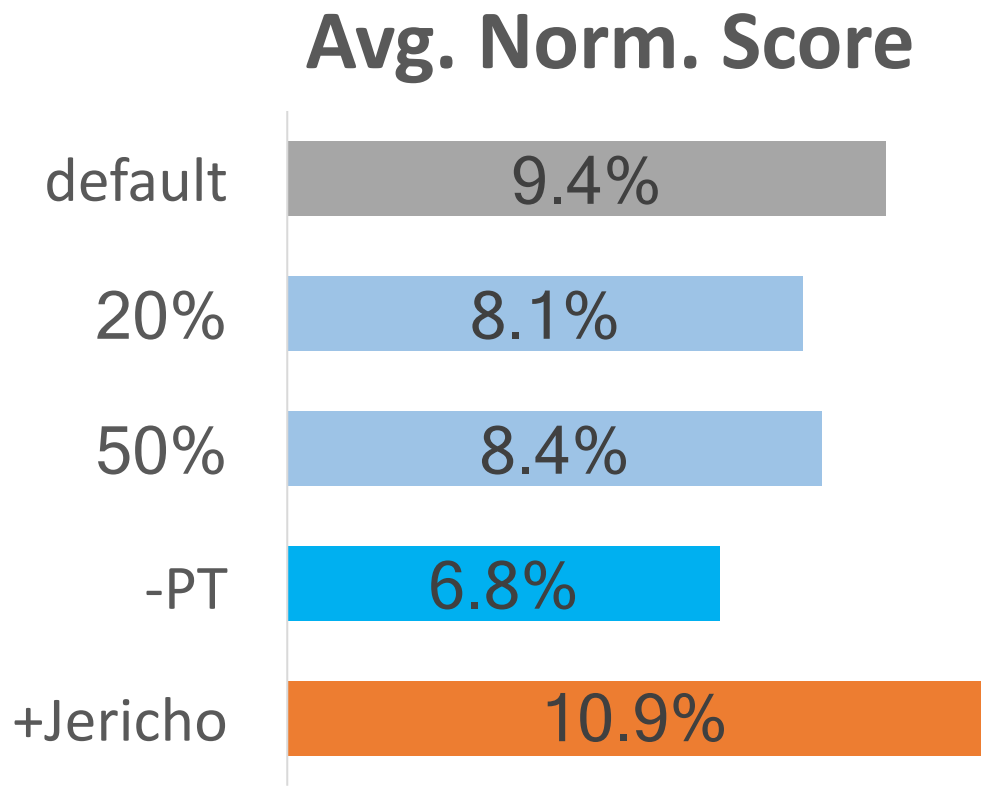
k: tradeoff action diversity v. quality!

After k=20, GPT-2 captures some complex/uncommon actions n-gram cannot



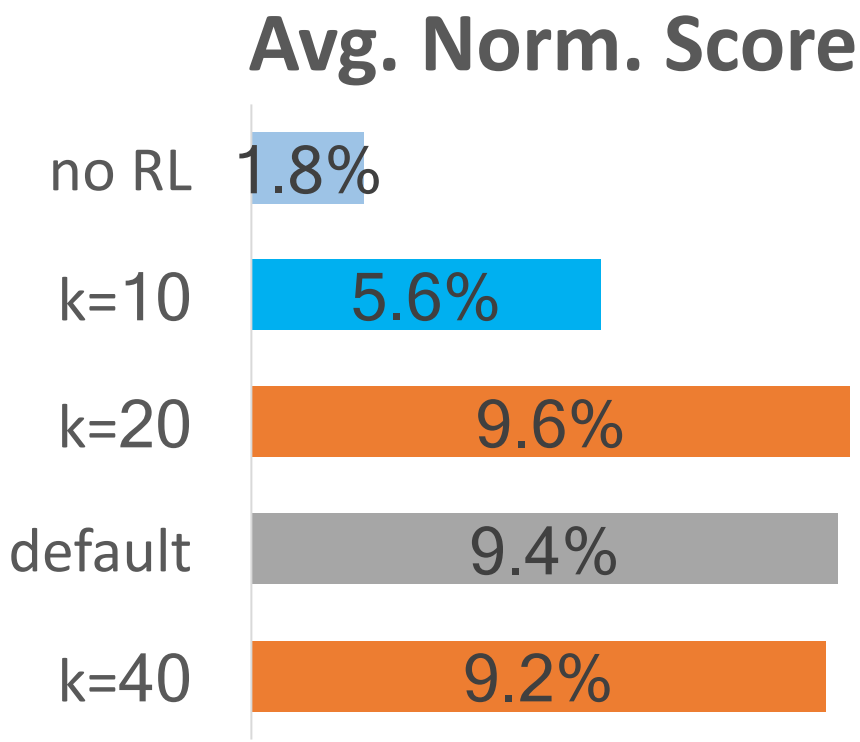
CALM (GPT-2) Analysis: Value of Data

- Two sources of data:
  - GPT-2 pretraining
  - ClubFloyd training
- **20%/50%:** less ClubFloyd data
- **-PT**: GPT-2 randomly initialized
- **+Jericho**: add back 8 Jericho game scripts in ClubFloyd training



CALM (GPT-2) Analysis: RL

- **No RL**: cannot work
- Number of actions k:
  - **k=30** is default
  - **k=20/40**: overall similar with k=30, different scores across games
  - **k=10**: significantly worse, ~ CALM (n-gram)



Summary

- A language model approach to text game action generation
- RL perspective: action space reduction via LM pre-training
- Language perspective: generalized functional use across games
- Code and dataset: https://github.com/princeton-nlp/calm-textgame