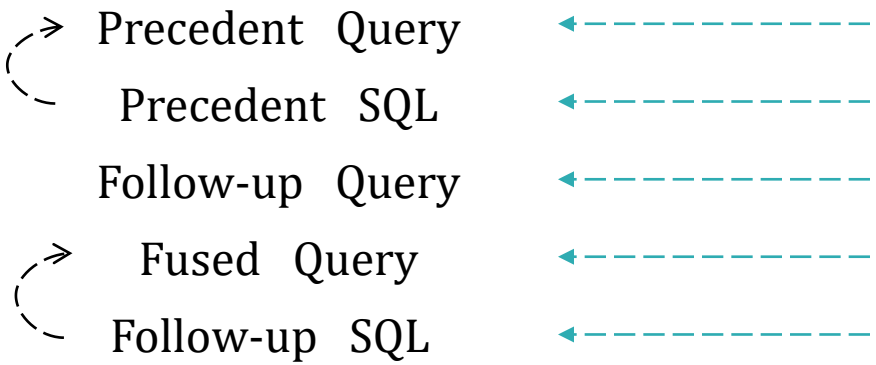
FANDA: A Novel Approach to Perform Follow-up Query Analysis

Qian Liu, Bei Chen, Jian-Guang Lou, Ge Jin, Dongmei Zhang

Microsoft Research
微软亚洲研究院

/01 Introduction



Brand	Sales	Profit	Year
BMW	31020	5000	2009
Ford	25220	3000	2009
Benz	47060	6000	2009

74.58% queries follow immediately after the question they are related to. (Bertomeu et al., 2006)

Interaction

User [1] : show the sales of BMW in 2009.

System : SELECT Sales WHERE Brand = BMW and Year = 2009

User [2] : what about profit?

show the profit of BMW in 2009.

System : SELECT Profit WHERE Brand = BMW and Year = 2009

User [3] : of Benz?

show the profit of Benz in 2009.

System : SELECT profit WHERE Brand = Benz and Year = 2009

User [4] : Compare it to Ford.

Compare the profit of Benz in 2009 to Ford.

System : SELECT profit WHERE (Brand = Benz or Brand = Ford) and Year = 2009

/01 | Introduction

ATIS3

- Dahl D A, Bates M, Brown M, et al. Expanding the scope of the ATIS task: The ATIS-3 corpus[C]//HLT-ACL 1994
- Miller S, Stallard D, Bobrow R, et al. A fully statistical approach to natural language interfaces[C]//ACL 1996
- Zettlemoyer L S, Collins M. Learning context-dependent mappings from sentences to logical form[C]//ACL 2009
- Suhr A, Iyer S, Artzi Y. Learning to Map Context-Dependent Sentences to Executable Formal Queries[C]//NAACL 2018

SequentialQA

- Iyyer M, Yih W, Chang M W. Search-based neural structured learning for sequential question answering[C]//ACL 2017

Non-sentential Question Resolution

- Kumar V, Joshi S. Non-sentential Question Resolution using Sequence to Sequence Learning[C]//COLING 2016
- Kumar V, Joshi S. Incomplete Follow-up Question Resolution using Retrieval based Sequence to Sequence Learning[C]//SIGIR 2017

/01 | Introduction

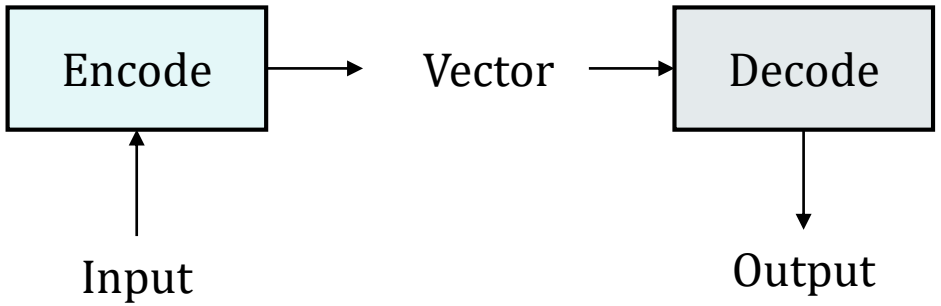
- Prior work in context-dependent parsing focuses on specific domain or simple scenarios.
- **Our goal:** language understanding in complex scenarios covering diverse domains in NLIDB.

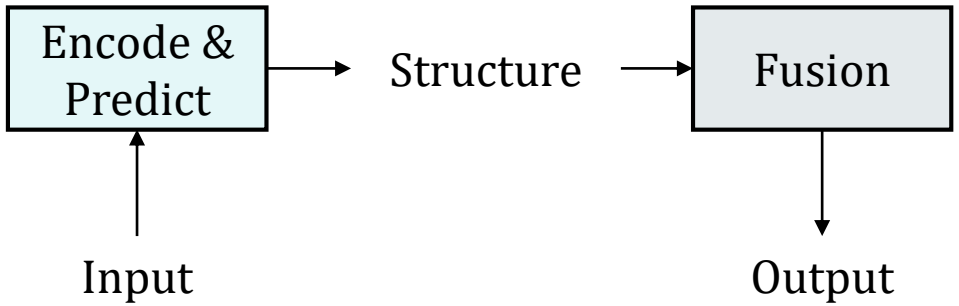
  - **Dataset**: A new dataset FollowUp is presented for research and evaluation.
  - **Method**: A novel approach is presented for taking account interaction history information into current sentence.

Dataset

Scenario	Example
Analytics	Precedent : In 1995, is there any network named CBC? Follow-up : Any TSN? Fused : In 1995, is there any network named TSN?
Compare	Precedent : How much money has Smith earned? Follow-up : Compare it with Bill Collins. Fused : Compare money Smith earned with Bill Collins.
Calc & Stats	Precedent : List all universities founded before 1855. Follow-up : Show their number. Fused : Show the number of all universities founded before 1855.
Extremum	Precedent : Which stadium has the most capacity? Follow-up : Which get the highest attendance? Fused : Which stadium get the highest attendance?
Filter	Precedent : How many roles are from studio paramount? Follow-up : List all titles produced by that studio. Fused : List all titles produced by studio paramount.
Group	Precedent : Show the industry which has the most companies? Follow-up : Show in different countries. Fused : Show the industry which has the most companies in different countries.
Sort	Precedent : Show all chassis produced after the year 1990. Follow-up : Sort them by year. Fused : Show all chassis produced after the year 1990 and sort by year.
Search	Precedent : What position did Sid O'Neill play? Follow-up : Which players else are in the same position? Fused : Which players play in the position of Sid O'Neill excluding Sid O'Neill?

- 1000 queries in 120 different Tables inherited from WikiSQL
- Annotation with Query Triple : (Precedent, Followup, Fused)
- Train/Dev/Test : 640/160/200





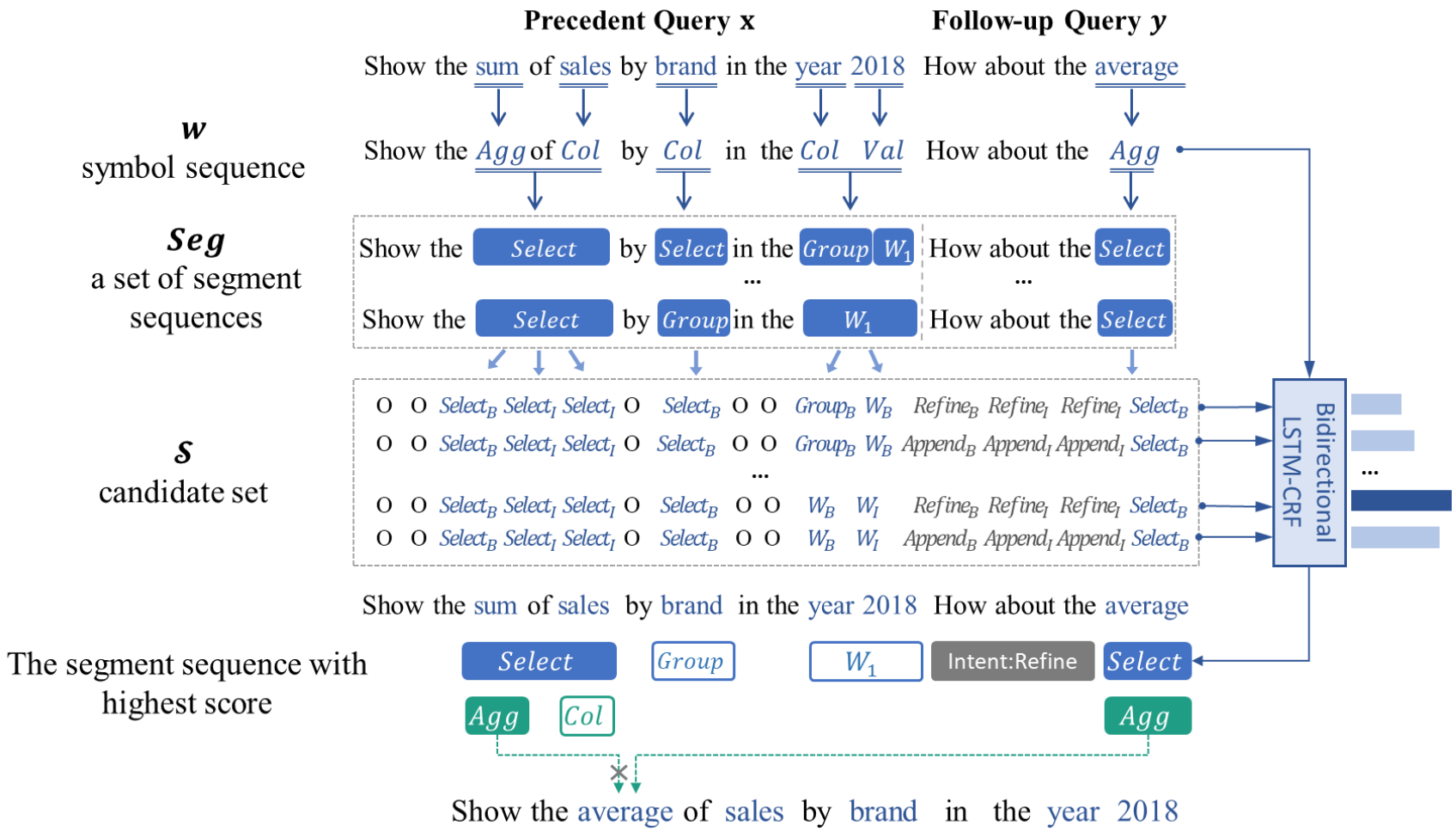
Sequence to Sequence

- Learning to encode and decode
- Non-interpretable
- Require lots of training data

Follow-up ANalysis for DAtabase

- Learning to encode, fusion with semantic rules
- Reason when fusion
- Cold start for little data

/02 | Follow-up Analysis for Database



Anonymization

Generate **symbol** sequence

Generation

Generate **segment** sequence

Fusion

Query fusion in structure level

Microsoft

Follow-up Analysis for Database

- Words in utterance are split into two types: **analysis-specific words** and **rhetorical words**.
- All numbers and dates belong to **Val**.
- One analysis-specific word could belong to different symbols, generating several symbol sequences.

Show the sum of sales by brand in the year 2018

↓ ↓ ↓ ↓ ↓

Show the *Agg* of *Col* by *Col* in the *Col Val*

Symbol	Meaning	
Col	Column Name	Table-Related Knowledge
Val	Cell Value	
Agg	Aggregation	
Com	Comparison	
Dir	Order Direction	
Per	Personal Pronoun	Language-Related Knowledge
Pos	Possessive Pronoun	
Dem	Demonstrative	

Follow-up Analysis for Database

- **symbol** does not consider the context around and **segment** structure is designed.
- **segment** is a combination of adjacent symbols, inspired by SQL parameter and common sense.

Segment	Compositional Deduction Rule
Select	[Agg + [Val]] + Col
Group	Col
Order	[Dir] + Col
W_1	[Col] + [Com] + Val
W_2	Col + Com + Col
P_1	Per
P_2	Pos
P_3	Dem + Col

Precedent Query:
Could you tell me the player whose score is larger than 67

Select W_1

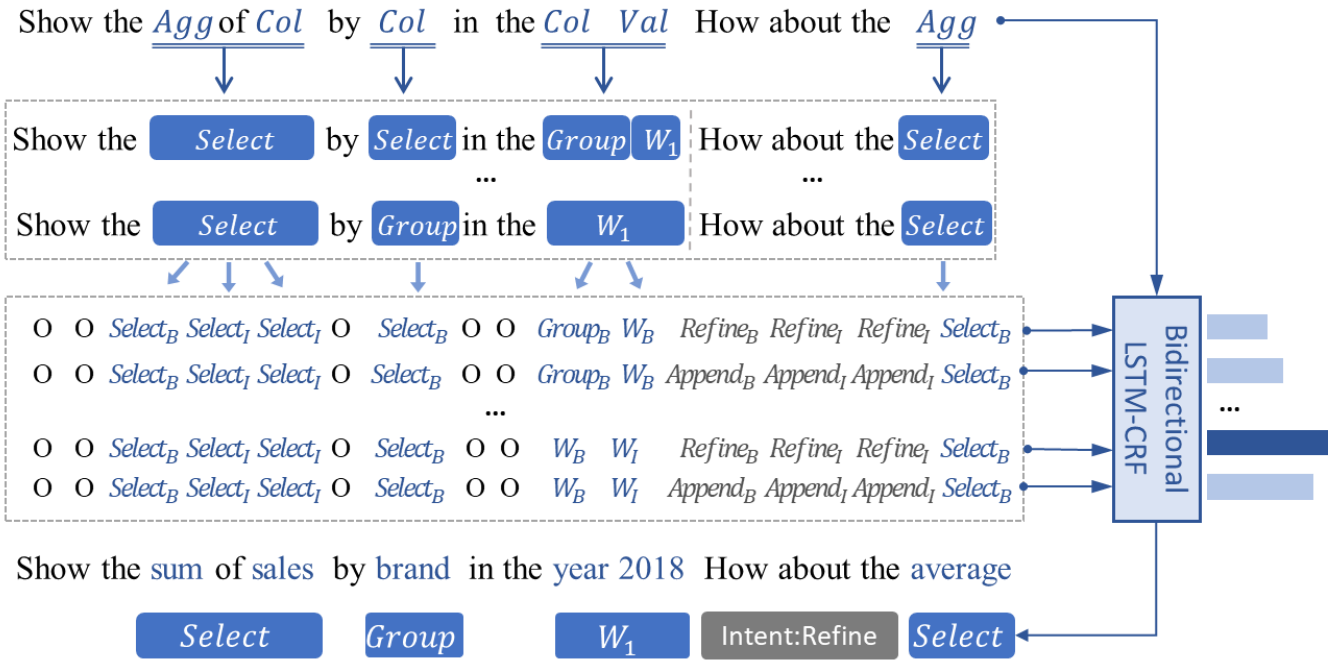
Follow Query (1):
Who play the same position as him ?

P_3 P_1

Follow Query (2):
sort them using their score in ascending order.

P_1 $Order$

Follow-up Analysis for Database



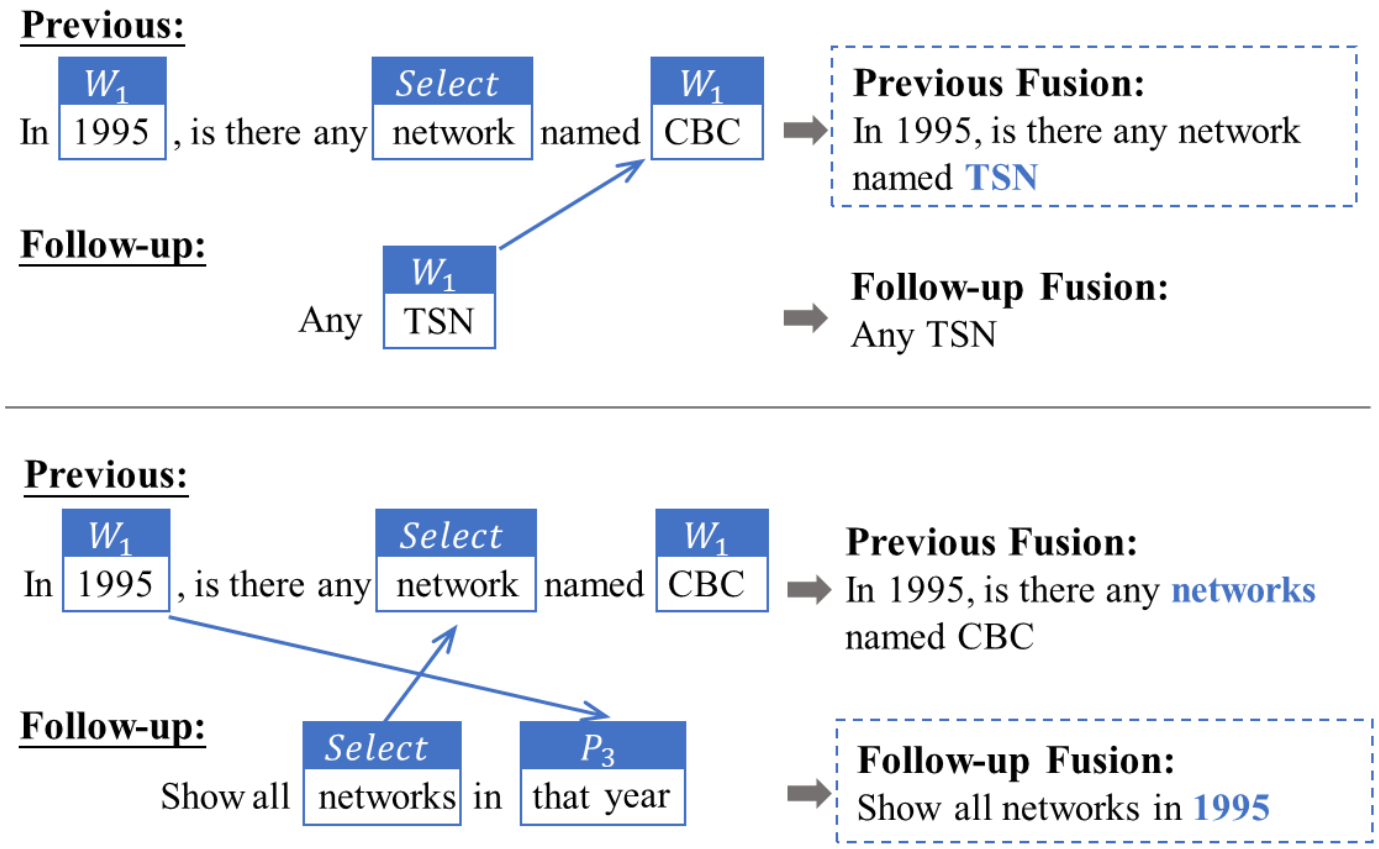
1. Symbols are combined to generate all possible segment sequences.
2. A ranking model is built to score these segment sequences and pick the best one as output.
3. Intent was introduced to distinguish two scenarios: Refine & Append.

Follow-up Analysis for Database

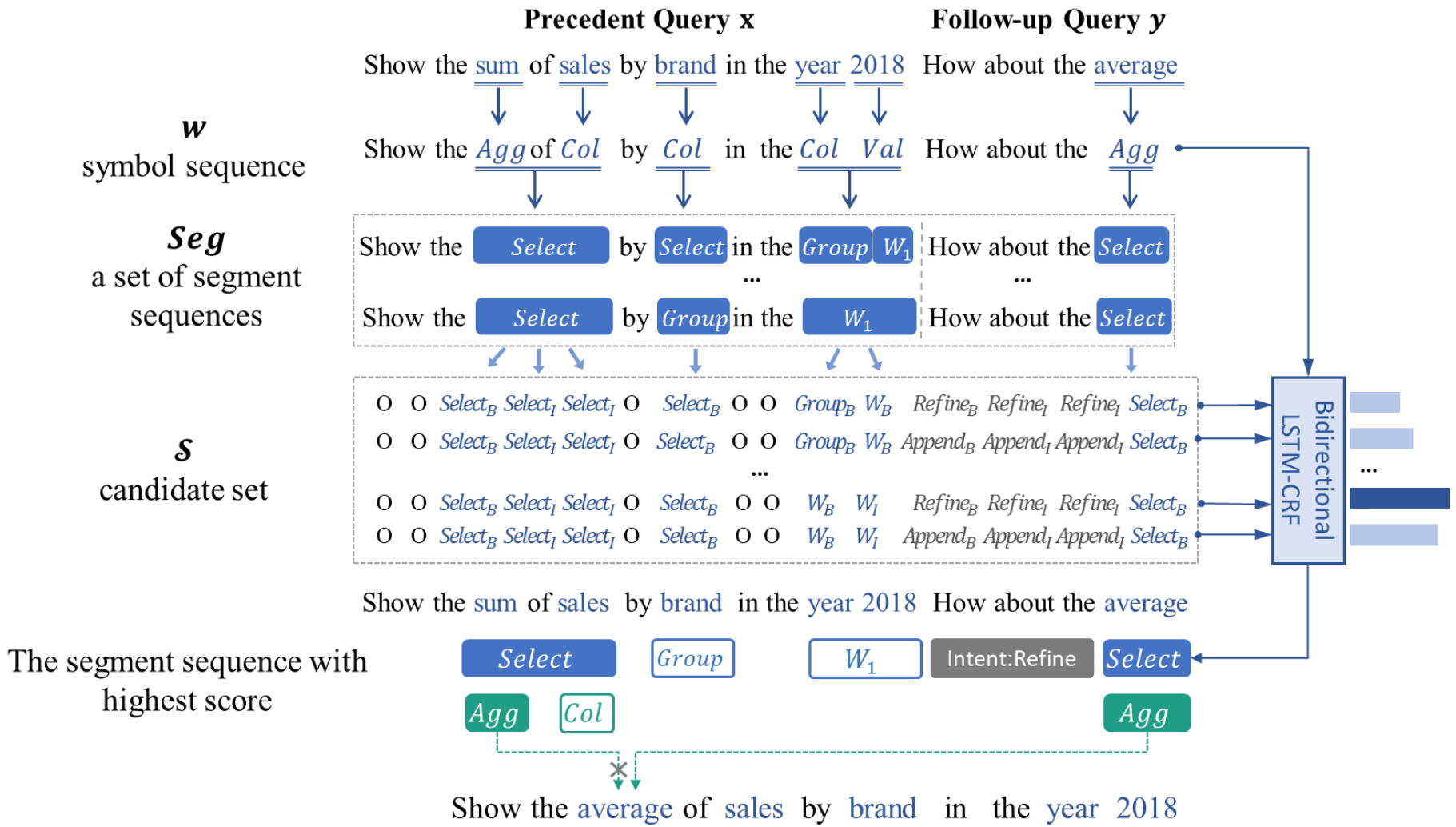
Network	Year
TSN	1995
CBC	1995
CFL	1996

1. Conflicting segment pairs will not happen at the same time.

2. Utilize one sentence to make up for the lack of the another sentence.



Follow-up Analysis for Database



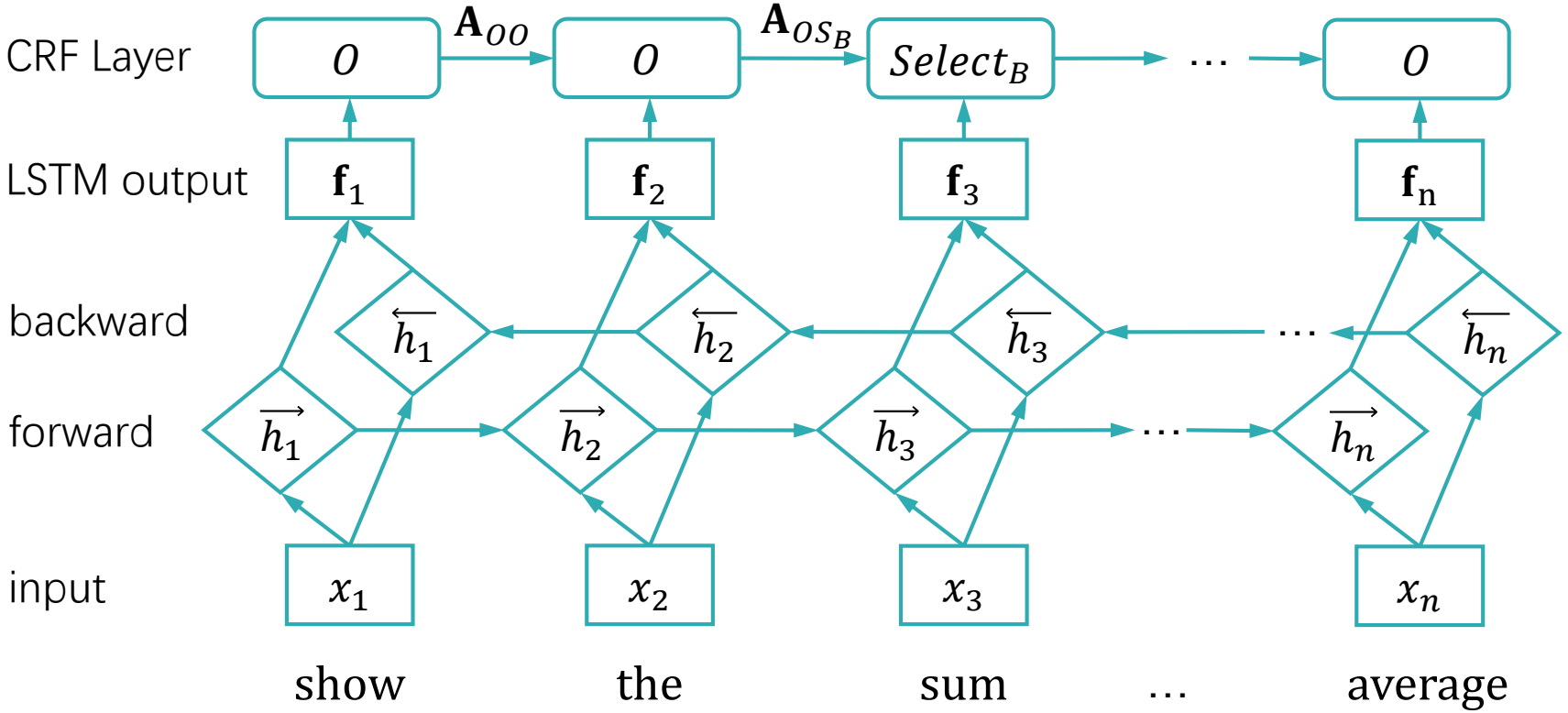
Follow-up Analysis for Database

$\mathcal{P}$

$O\ O\ Select_B\ Select_I \ldots Refine_I\ Select_B$
$O\ O\ Select_B\ Select_I \ldots Refine_I\ Select_B$

$\mathcal{N}$

$O\ O\ Select_B\ Select_I \ldots Append_I\ Select_B$
$O\ O\ Select_B\ Select_I \ldots Append_I\ Select_B$



$$g(\boldsymbol{s}_k|\Theta) = \sum_{i=1}^{N}(\mathbf{A}_{t_{i-1}^k t_i^k} + \mathbf{f}_i[t_i^k])$$

$$\widehat{\boldsymbol{s}}_p = \arg\max_{\boldsymbol{s}\in\mathcal{P}} g(\boldsymbol{s}|\Theta)$$

$$\widehat{\boldsymbol{s}}_n = \arg\max_{\boldsymbol{s}\in\mathcal{N}} g(\boldsymbol{s}|\Theta)$$

$$\max(0, \Delta - \frac{g(\widehat{\boldsymbol{s}}_p|\Theta)}{|\widehat{\boldsymbol{s}}_p|} + \frac{g(\widehat{\boldsymbol{s}}_n|\Theta)}{|\widehat{\boldsymbol{s}}_n|})$$

Experiments

	Model	Symbol Acc (%)	BLEU (%)
Dev	SEQ2SEQ	0.63 ± 0.00	21.34 ± 1.14
	COPYNET	17.50 ± 0.87	43.36 ± 0.54
	S2S+ANON	18.75 ± 0.95	41.22 ± 0.33
	COPY+ANON	25.50 ± 2.47	51.45 ± 0.93
	FANDA	**49.00** ± 1.28	**60.14** ± 0.98
Test	CONCAT	22.00 ± –	52.02 ± –
	E2ECR	27.00 ± –	52.47 ± –
	SEQ2SEQ	0.50 ± 0.22	20.72 ± 1.31
	COPYNET	19.30 ± 0.93	43.34 ± 0.45
	S2S+ANON	18.80 ± 1.77	38.90 ± 2.45
	COPY+ANON	27.00 ± 4.32	49.43 ± 1.11
	FANDA	47.80 ± 1.14	59.02 ± 0.54

Model	Execution Accuracy (%)
CONCAT	25.24
E2ECR	27.18
COPY+ANON	40.77
FANDA	**60.19**

- ***Symbol Acc***: Symbol Consistent With Gold Fused Query
- ***BLEU***: Quality of Output Fused Query
- ***Execution Accuracy***: Output Query Execution Correctness (Parser using Coarse-to-Fine)

  - SEQ2SEQ*:* Attention SEQ2SEQ
  - COPYNET: + copy mechanism
  - S2S + ANON: SEQ2SEQ + anonymization
  - COPY + ANON: COPY+ anonymization
  - CONCAT: Concatenate Precedent Query and Follow-up
  - E2ECR: End to End Coreference Resolution System

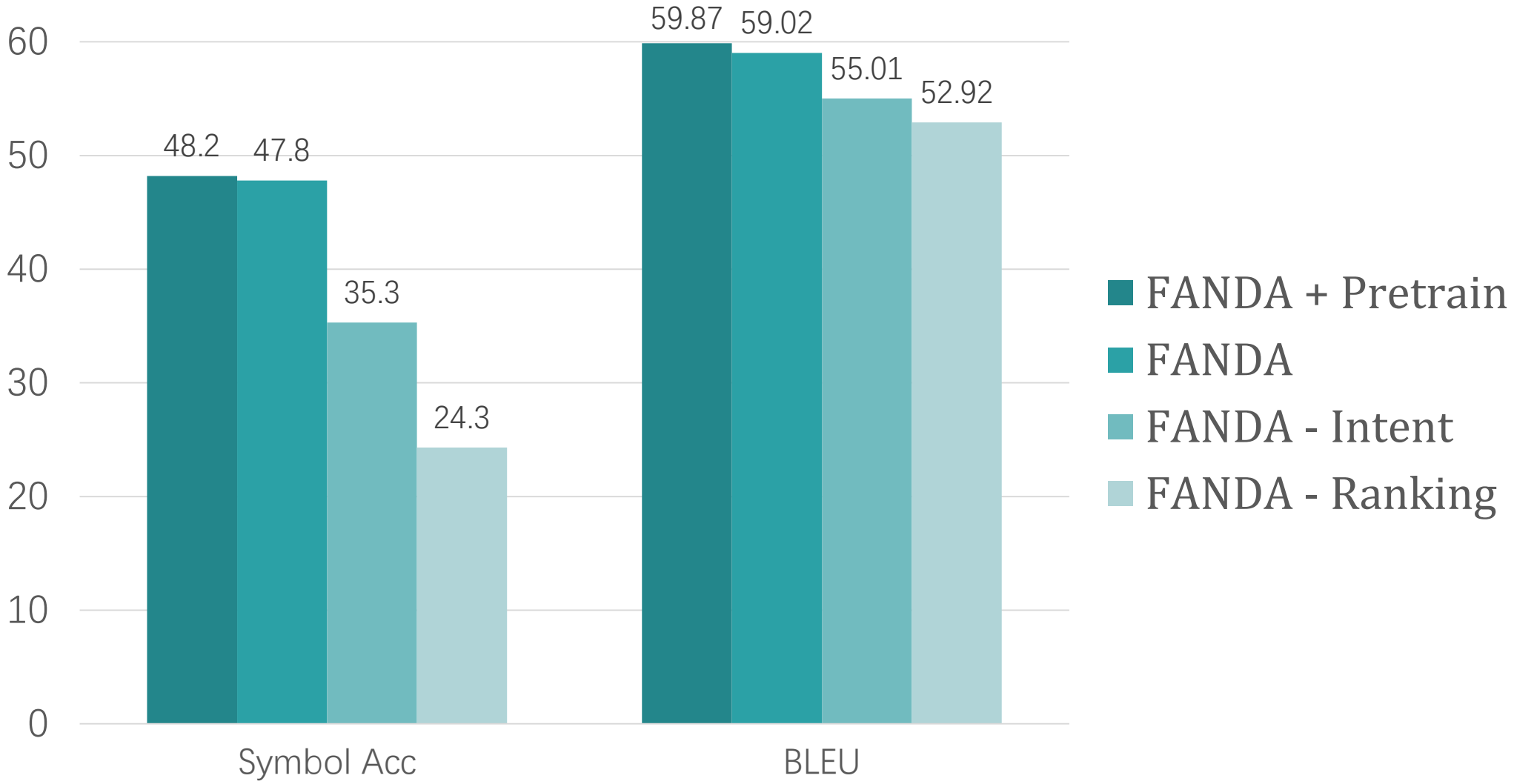
Anonymization

Origin : In 1995, is there any network named CBC? Any TSN?

Transform : In Val#1, is there any Col#1 named Val#2? Any Val#3?

Experiments

Ablation Results



Error Case Analysis

No	Case Analysis		
1	Precedent	: What is the result, when the home team score is 2.4.6?	Follow-up : What is the date?
	Gold Fusion	: What is the date, when the home team score is 2.4.6?	
	COPY+ANON	: What is the date, the home team score is 2.4.6?	
	FANDA	: What is the date, when the home team score is 2.4.6?	
2	Precedent	: Which is the draw number of Lowry?	Follow-up : How about Laura?
	Gold Fusion	: Which is the draw number of Laura?	
	COPY+ANON	: Which is the draw number of Lowry?	
	FANDA	: Which is the draw number of Laura?	
3	Precedent	: What are the names when elevation is feet?	Follow-up : Of those, whose GNIS feature is 1417308?
	Gold Fusion	: Of names when elevation is feet, whose GNIS feature is 1417308?	
	COPY+ANON	: What are the names when elevation is 1417308, whose GNIS feature is feet?	
	FANDA	: What are the names when elevation is feet whose GNIS feature is 1417308?	

	COPY+ANON	FANDA
Substantial Overlap	√	√ (Segment Type)
No Overlap		√ (Table Structure)
Ambiguous overlap		√ (Combination of Symbol)

Future Work

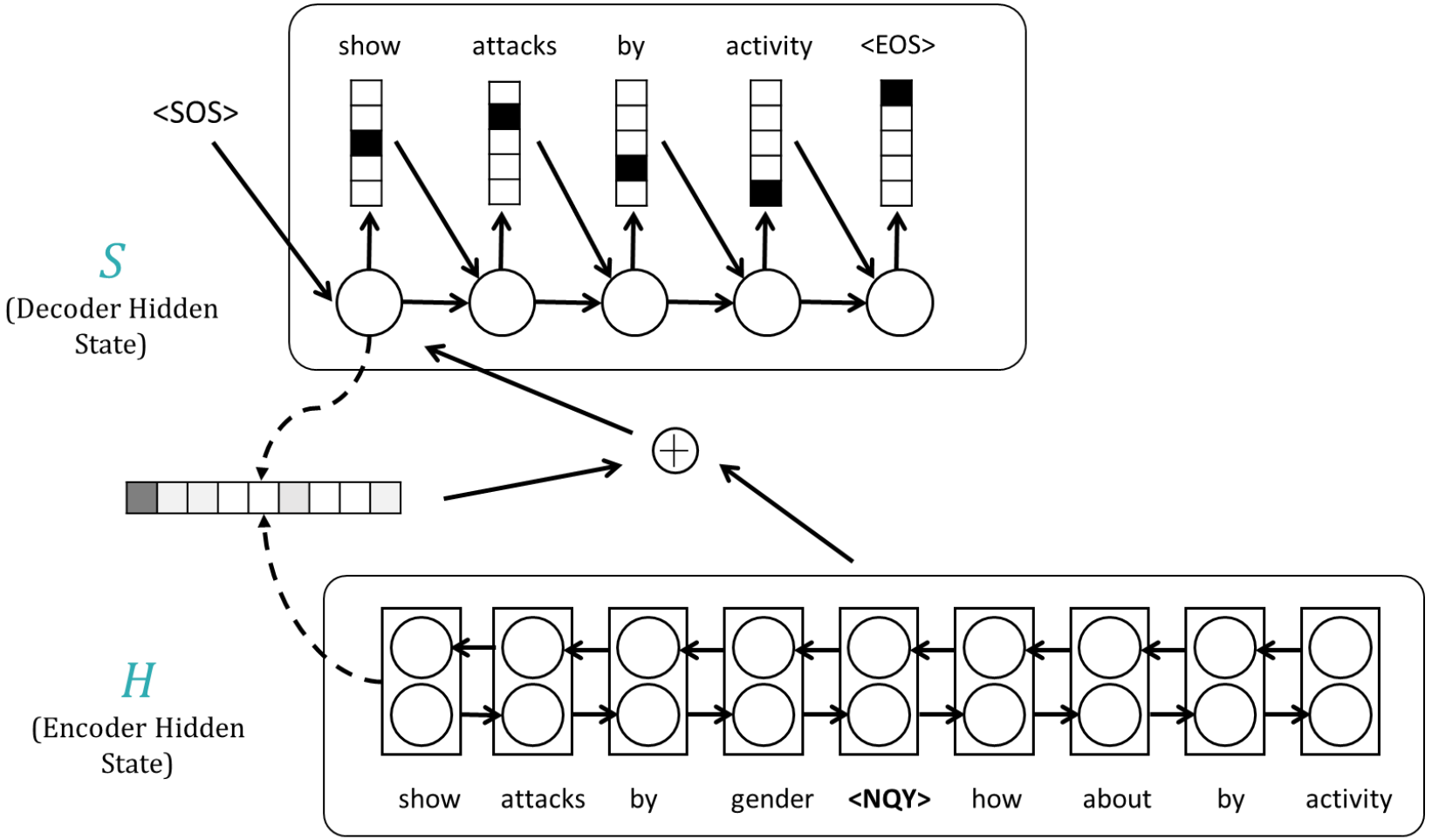
- Extending to multi-turns and multi-tables.
- Using reinforcement learning

Thank you!

Microsoft Research
微软亚洲研究院

Question & Answer

SEQ2SEQ



$$\boldsymbol{S}_1 = RNN(\boldsymbol{H}_n, \boldsymbol{E}_{sos})$$

$$\alpha_i = \boldsymbol{S}_1^T \boldsymbol{W} \boldsymbol{H}_i$$

$$a_i = \frac{e^{\alpha_i}}{\sum e^{\alpha_k}}$$

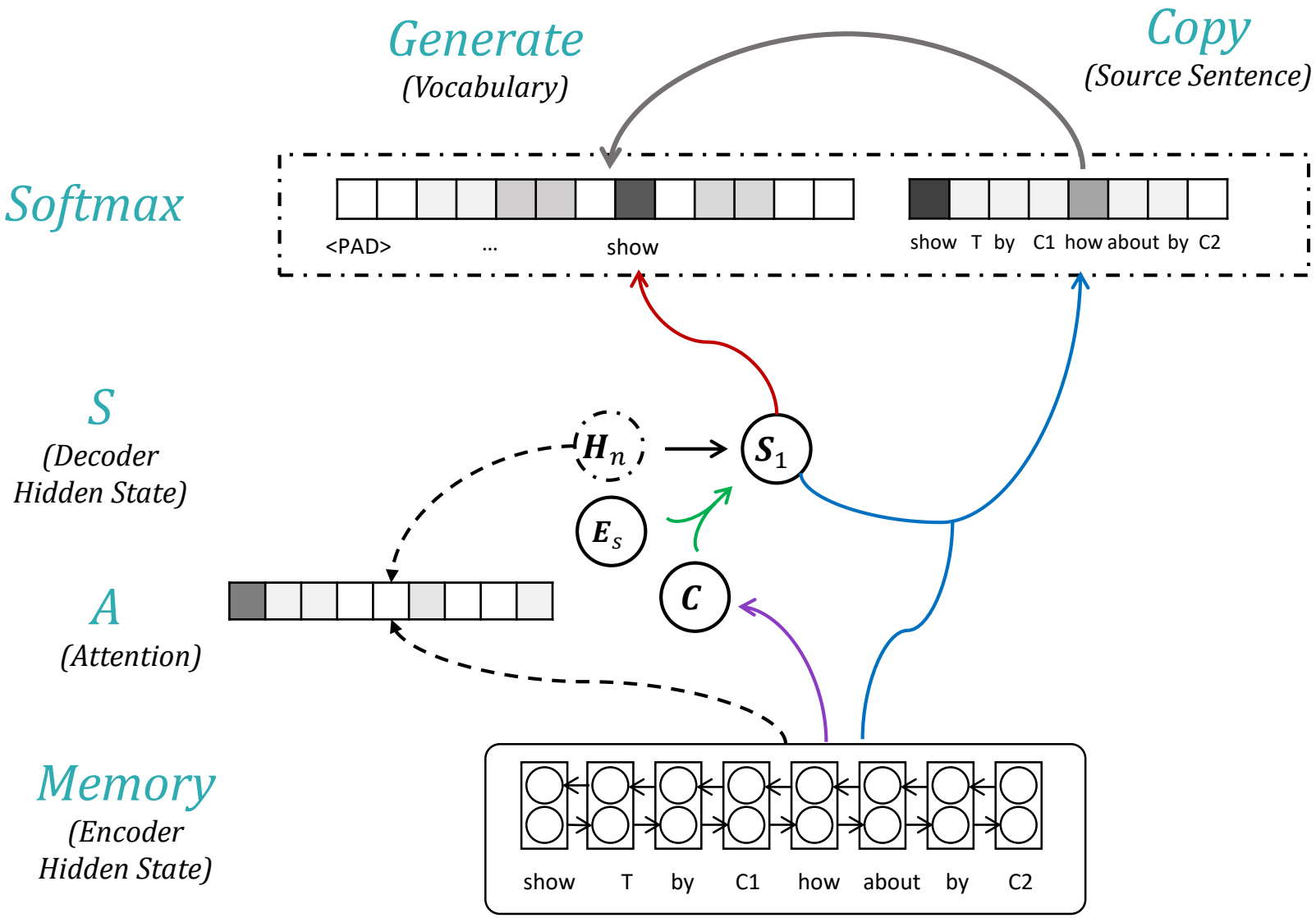
$$\boldsymbol{C} = \sum a_i \cdot \boldsymbol{H}_i$$

$$\boldsymbol{O} = F([\boldsymbol{C}, \boldsymbol{S}_1])$$

$$\boldsymbol{S}_2 = RNN(\boldsymbol{S}_1, \boldsymbol{E}_o)$$

Question & Answer

COPYNET



$$\alpha_i = V_a^T \tanh(\boldsymbol{W}[\boldsymbol{H}_n \boldsymbol{H}_i])$$

$$a_i = \frac{e^{\alpha_i}}{\sum e^{\alpha_k}}$$

$$\boldsymbol{C} = \sum a_i \cdot \boldsymbol{H}_i$$

$$\boldsymbol{S}_1 = RNN(\boldsymbol{H}_n, [\boldsymbol{E}_s, \boldsymbol{C}])$$

$$\boldsymbol{O}_G = \boldsymbol{W}_o \boldsymbol{S}_1$$

$$\beta_j = \tanh(\boldsymbol{W}_C \boldsymbol{H}_j)\, \boldsymbol{S}_1$$

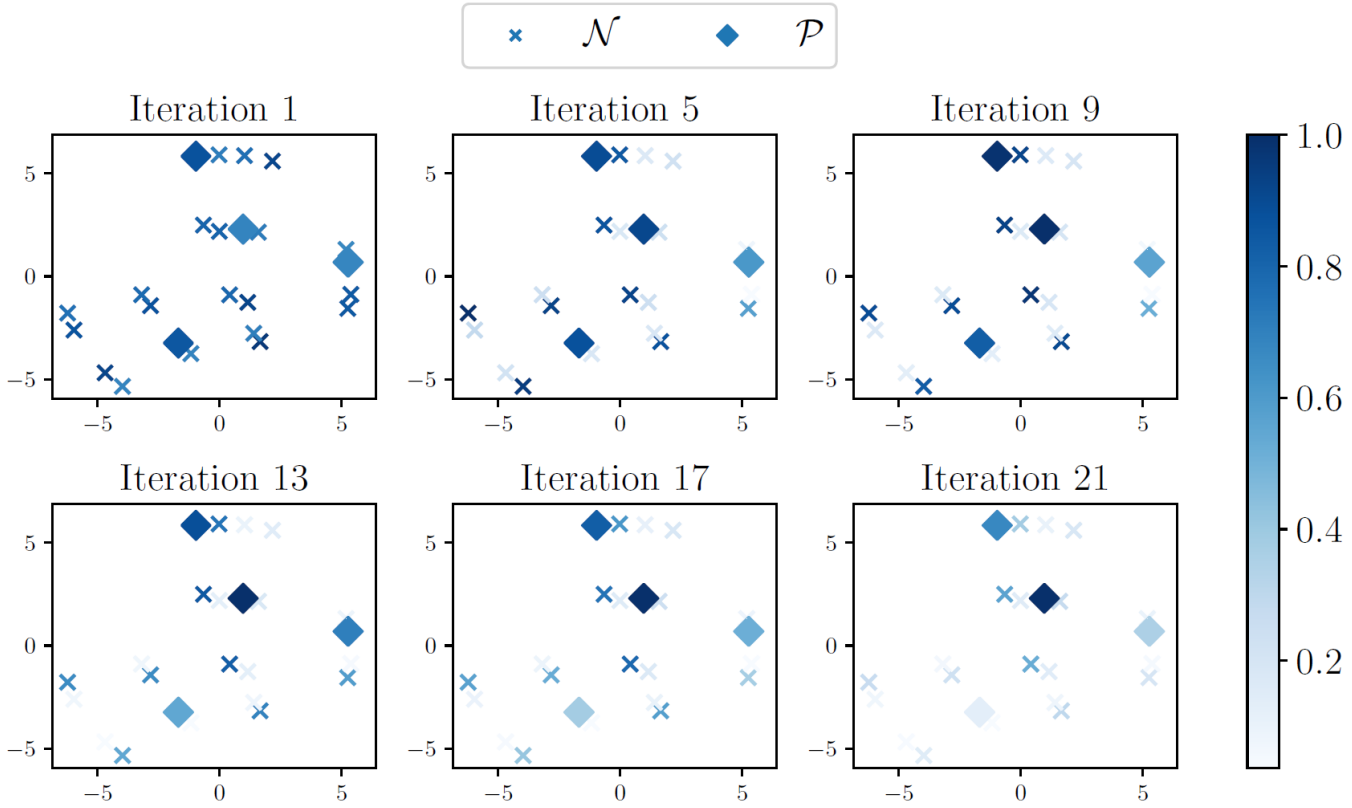
$$\boldsymbol{O}_C = [\beta_1, \ldots, \beta_s]$$

$$\boldsymbol{P} = Softmax([\boldsymbol{O}_G, \boldsymbol{O}_C])$$

$$p(y_t) = p(y_t, g|\cdot) + p(y_t, c|\cdot)$$

Question & Answer

Score Evolution



- In **Iteration 1**, different but similar scores are assigned to all candidates in P and N with random initialization.
- From **Iteration 5 to 21**, the score distribution becomes increasingly skewed.
- From **Iteration 13 to 21**, the candidate with the highest score remains unchanged, indicating the stability of weakly supervised learning.