Coarse-to-Fine Decoding for Neural Semantic Parsing

July 16, 2018

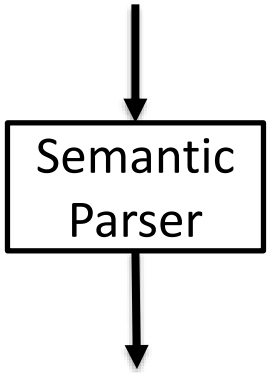
Li Dong and Mirella Lapata

Semantic Parsing

Mapping natural language to structured representations

Human-friendly -> Computer-friendly

all flights from dallas before 10am

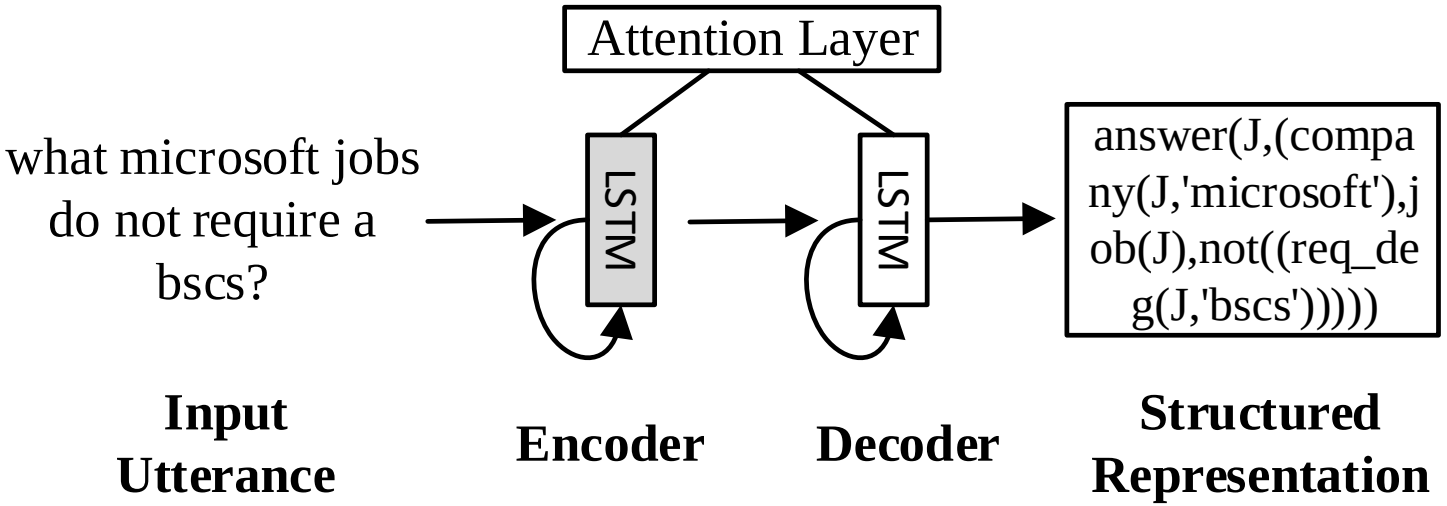


(lambda $0 e (and (flight $0) (from $0 dallas:ci) (< (departure_time $0) 1000:ti)))

Example from ATIS (Kwiatkowski et al., 2011)

Neural Semantic Parsing

- Sequence decoder (Jia and Liang, 2016; Dong and Lapata, 2016; Ling et al., 2016 ; Iyer et al., 2017)
- Syntactically-constrained decoder (Dong and Lapata, 2016; Xiao et al., 2016; Alvarez-Melis and Jaakkola, 2017; Yin and Neubig, 2017; Cheng et al., 2017; Krishnamurthy et al., 2017; Rabinovich et al., 2017; Xu et al., 2017)



This Work

all flights from dallas before 10am

Meaning Sketch

(lambda#2 (and flight@1 from@2 (< departure time@1 ?)))

&

Low-level Details

(e.g., arguments and variable names)

(lambda $0 e (and (flight $0) (from $0 dallas:ci) (< (departure_time $0) 1000:ti)))

Meaning Sketch

- Python code example

if length of bits is lesser than integer 3 or second element of bits is not equal to string ’as’ ,

```
if len ( NAME ) < NUMBER or NAME [ NUMBER ] != STRING :
```

```
if len(bits) < 3 or bits[1] != ’as’:
```

- SQL example

What record company did conductor Mikhail Snitko record for after 1996?

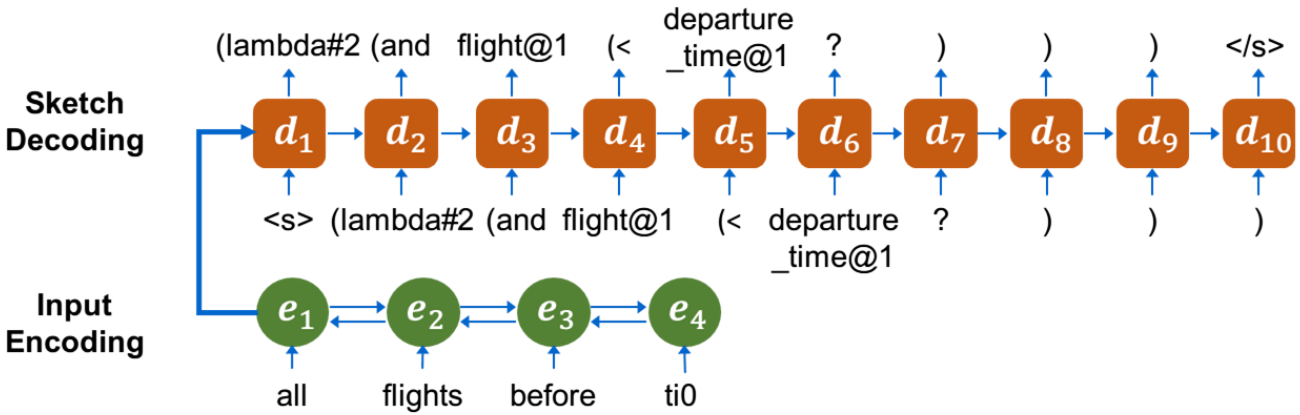
```
WHERE > AND =
```

```
SELECT Record Company WHERE (Year of Recording > 1996)
AND (Conductor = Mikhail Snitko)
```

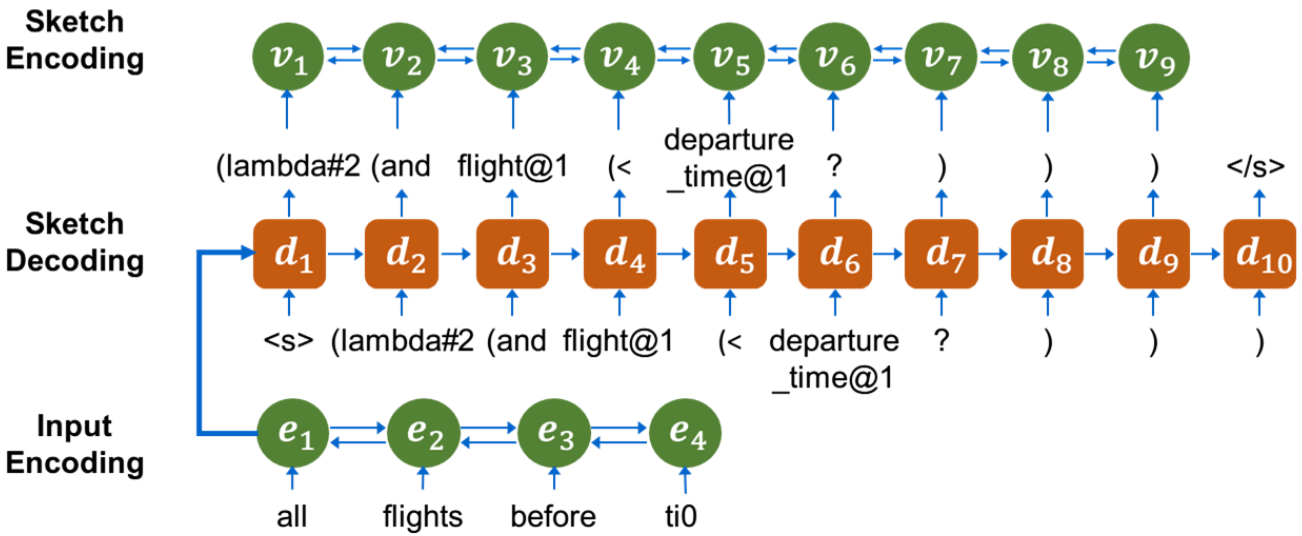
Meaning Sketch

- Disentangle high-level from low-level semantics
  - Model meaning at different levels of granularity
- More compact meaning representation
  - Length: $21.1 \rightarrow 9.2$ (on ATIS)
- Explicit sharing coarse structure
  - For examples that have the same basic meaning
- Provide global context to fine meaning decoder
  - Know what the basic meaning of input looks like

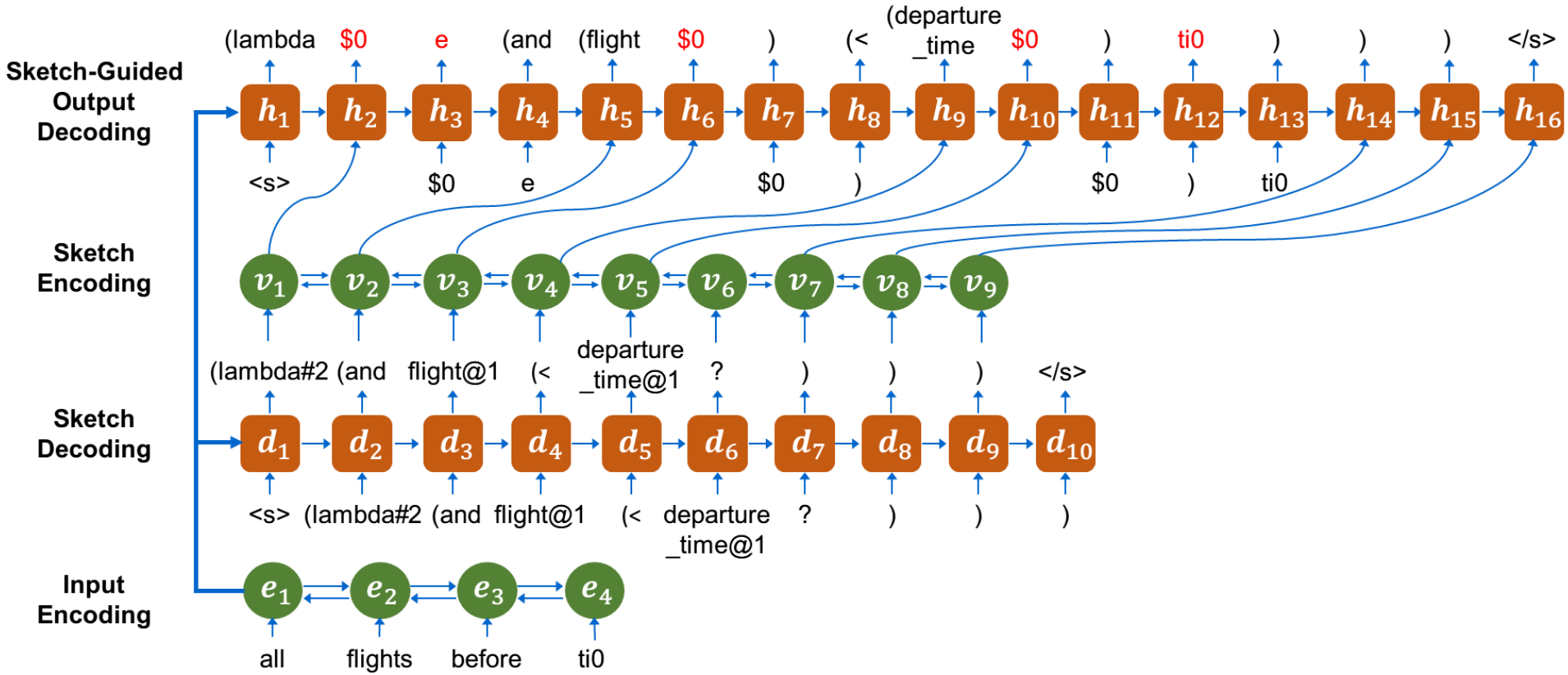
Method


Sketch Decoding
Input Encoding
(lambda#2 (and flight@1 (< departure _time@1 ?))) </s>
<s> (lambda#2 (and flight@1 (< departure _time@1 ?)))
all flights before ti0

Method


Sketch Encoding
v_1 v_2 v_3 v_4 v_5 v_6 v_7 v_8 v_9
(lambda#2 (and flight@1 (< departure _time@1 ?))) </s>
Sketch Decoding
d_1 d_2 d_3 d_4 d_5 d_6 d_7 d_8 d_9 d_{10}
<s> (lambda#2 (and flight@1 (< departure _time@1 ?)))
Input Encoding
e_1 e_2 e_3 e_4
all flights before ti0

Method


Sketch-Guided Output Decoding
Sketch Encoding
Sketch Decoding
Input Encoding
(lambda $0 e (and (flight $0) (< (departure _time $0) ti0))) </s>
<s> $0 e $0) $0) ti0
(lambda#2 (and flight@1 (< departure _time@1 ?))) </s>
<s> (lambda#2 (and flight@1 (< departure _time@1 ?)))
all flights before ti0

Method

Sketch constrains the decoding output

- Example 1: one augment is missing

- Example 2: type information

Training and Inference

- x: input, a: sketch, y: meaning representation
- Training: maximize the log likelihood

$$\max \sum_{(x,a,y)\in\mathcal{D}} \underbrace{\log p\left(y|x,a\right)}_{\text{Fine Meaning Decoder}} + \underbrace{\log p\left(a|x\right)}_{\text{Coarse Meaning Decoder}}$$

- Inference: greedy search

$$\hat{a} = \arg\max_{a'} p\left(a'|x\right)$$

$$\hat{y} = \arg\max_{y'} p\left(y'|x,\hat{a}\right)$$

Semantic Parsing Tasks

- Natural language to logical form (Geo/ATIS)

what is the population of the state with the largest area?
(argmax $0 (and (mountain:t $0) (loc:t $0 alaska:s)) (elevation:i $0))

- Natural language to source code (Django)

if length of bits is lesser than integer 3 or second element of bits is not equal to string 'as' ,
if len(bits) < 3 or bits[1] != 'as':

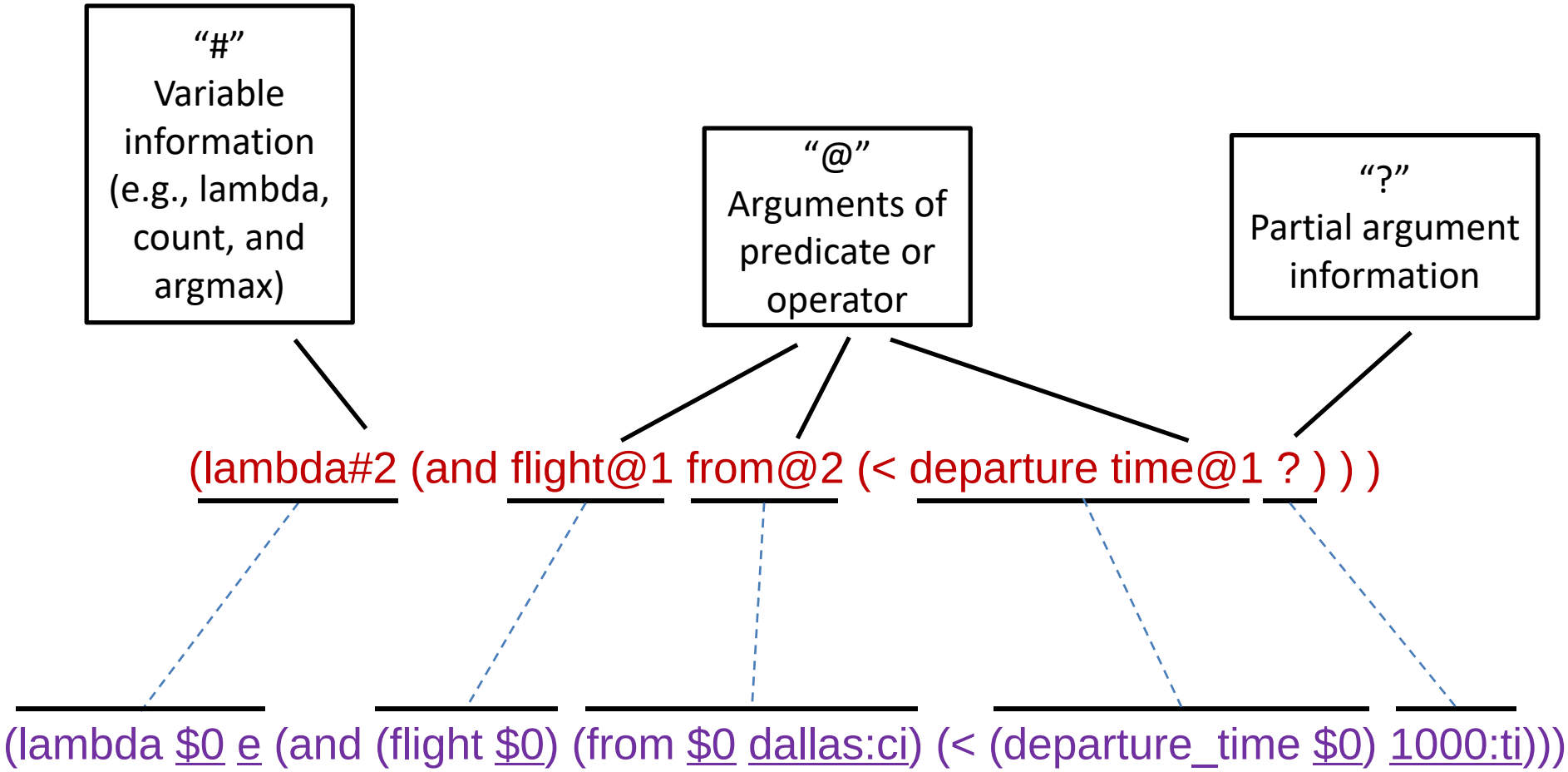
- Natural language to SQL (WikiSQL)

Pianist	Conductor	Record Company	Year of Recording	Format

What record company did conductor Mikhail Snitko record for after 1996?
SELECT *Record Company* WHERE (*Year of Recording* > *1996*) AND (*Conductor* = *Mikhail Snitko*)

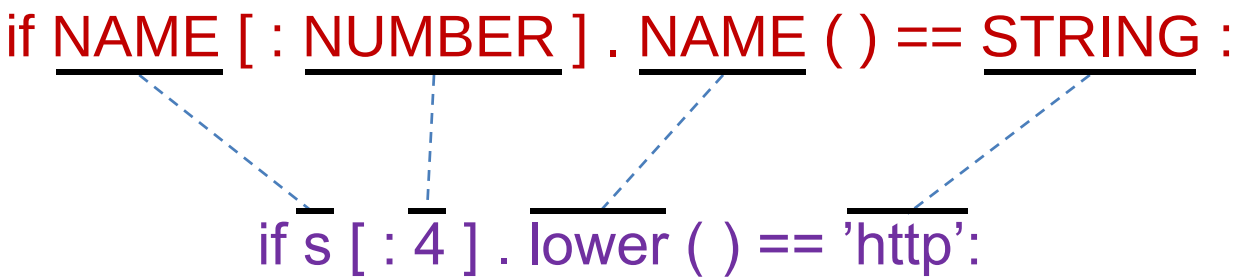
(Zettlemoyer and Collins, 2005; Kwiatkowski et al., 2011; Oda et al., 2015; Zhong et al., 2017)

Natural Language to Logical Form



Natural Language to Source Code

- Substitute tokens with their token types
- Except
  - Delimiters (e.g., “[”, and “:”)
  - Operators (e.g., “+”, and “*”)
  - Built-in keywords (e.g., “True”, and “while”)

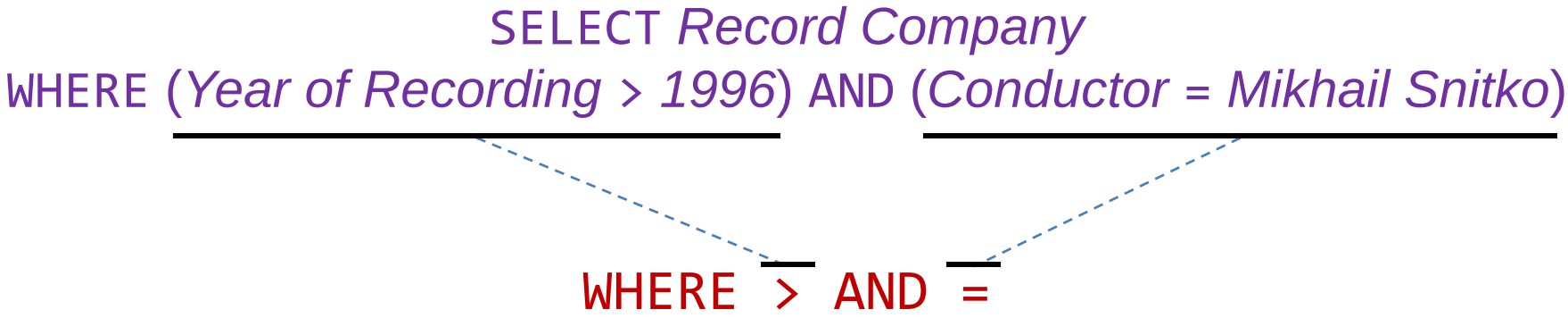


https://docs.python.org/3/library/tokenize.html

Natural Language to SQL

WikiSQL (Zhong et al., 2017)

```
SELECT agg_operator agg_column
WHERE (cond_column cond_operator cond_value)
AND ...
```



Natural Language to SQL

Decoding is table-aware

How many presidents are graduated from A?

President	College

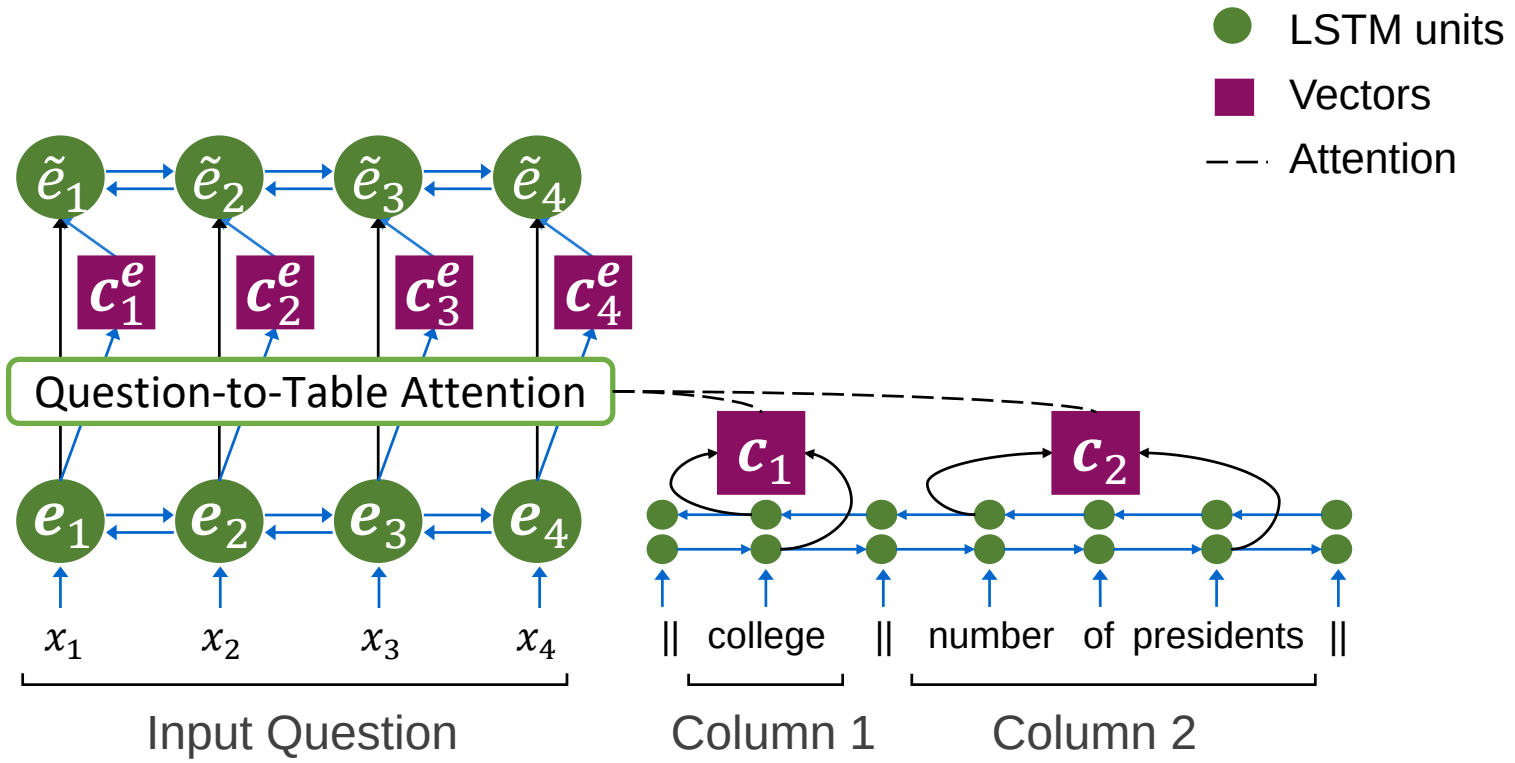
`SELECT COUNT(`*President*`) WHERE (`*College = A*`)`

College	Number of Presidents

`SELECT` *Number of Presidents* `WHERE (`*College = A*`)`

Natural Language to SQL

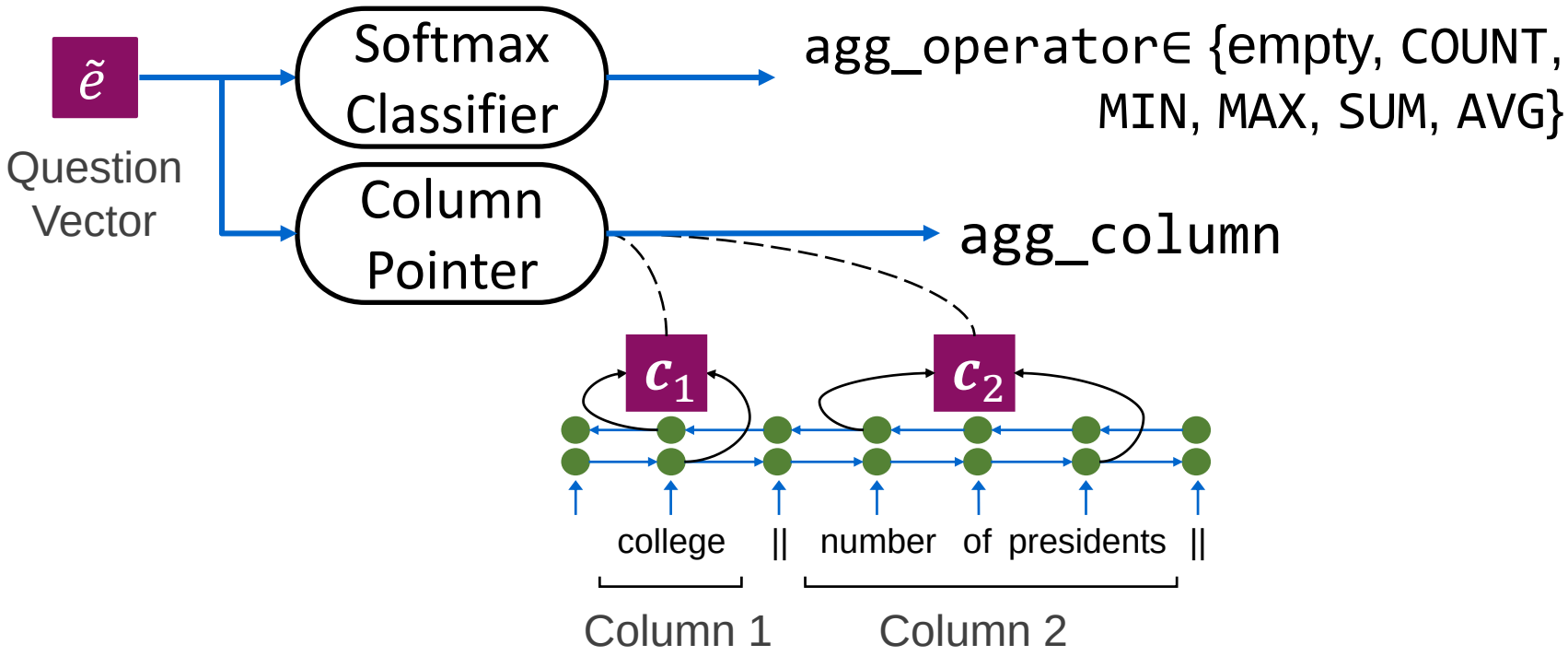
Table-aware input encoder



Natural Language to SQL

SELECT clause

```
SELECT agg_operator agg_column
WHERE (cond_column cond_operator cond_value)
AND ...
```



Natural Language to SQL

WHERE Clause

```
SELECT agg_operator agg_column
WHERE (cond_column cond_operator cond_value)
AND ...
```

Pianist	Conductor	Record Company	Year of Recording	Format

What record company did conductor Mikhail Snitko record for after 1996 ?

Sketch Classification

Natural Language to SQL

WHERE Clause

```
SELECT agg_operator agg_column
WHERE (cond_column cond_operator cond_value)
AND ...
```

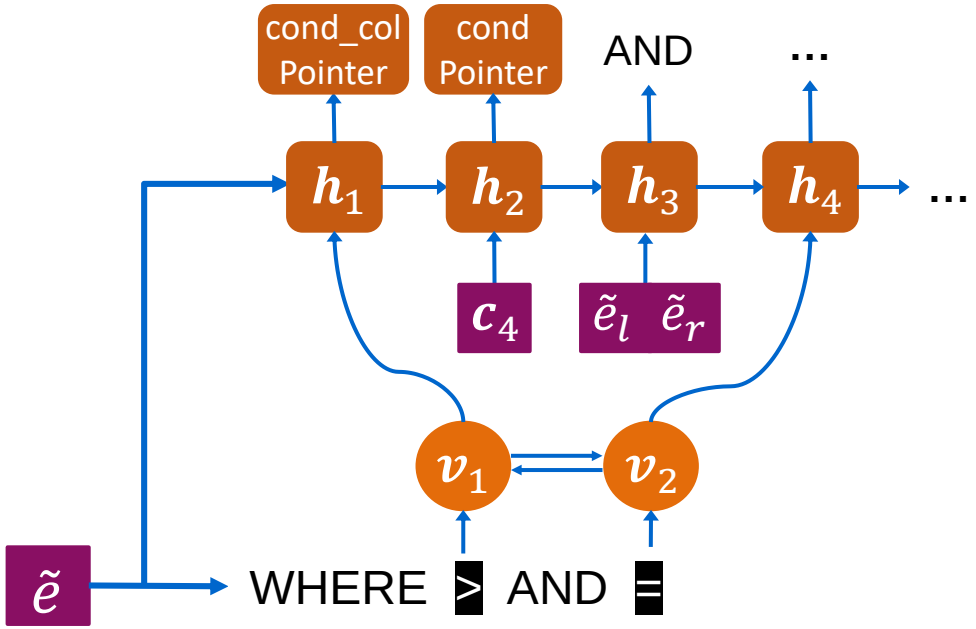
Pianist	Conductor	Record Company	Year of Recording	Format

What record company did conductor Mikhail Snitko record for after 1996 ?

Sketch-Guided WHERE Decoding

Sketch Encoding

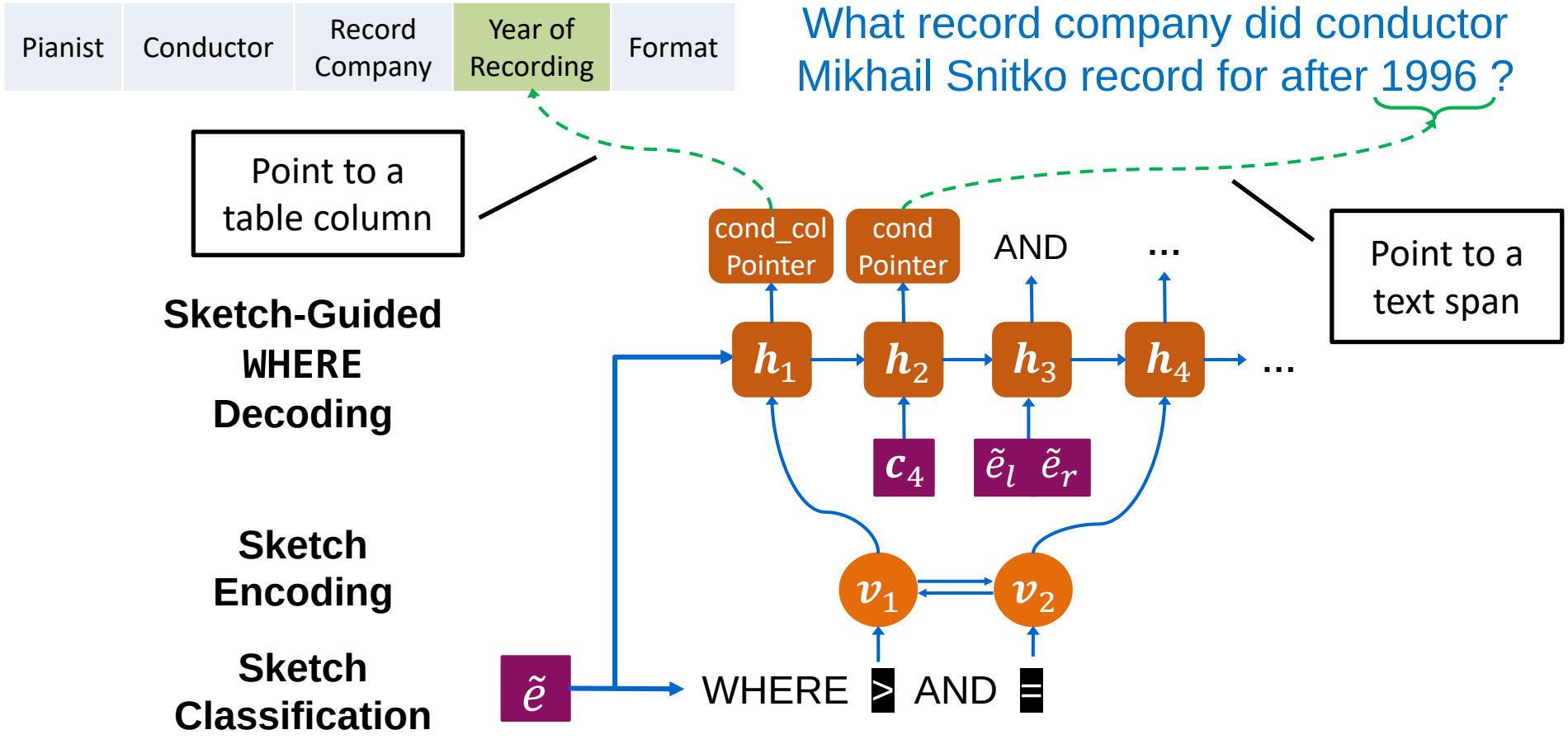
Sketch Classification



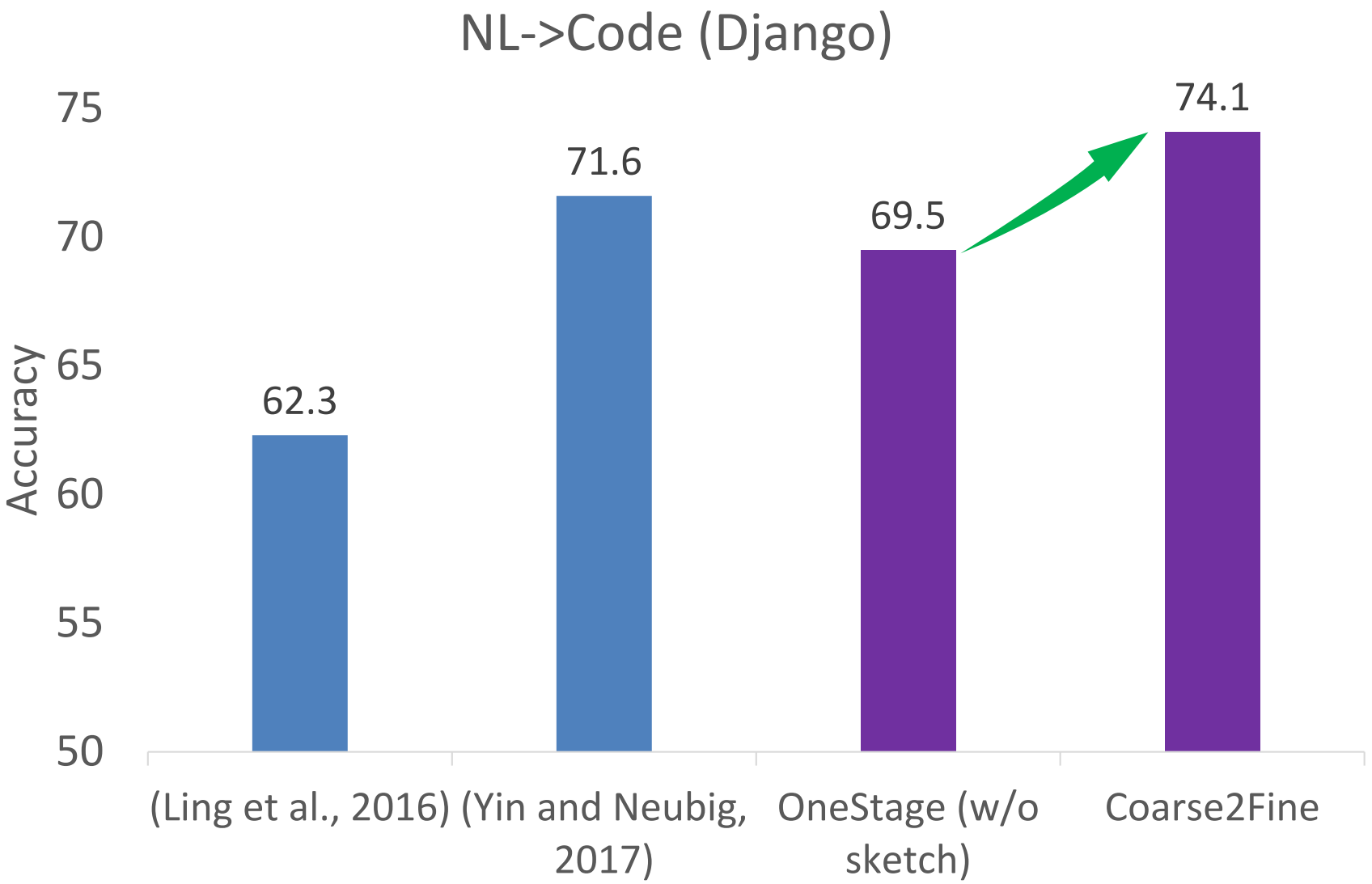
Natural Language to SQL

WHERE Clause

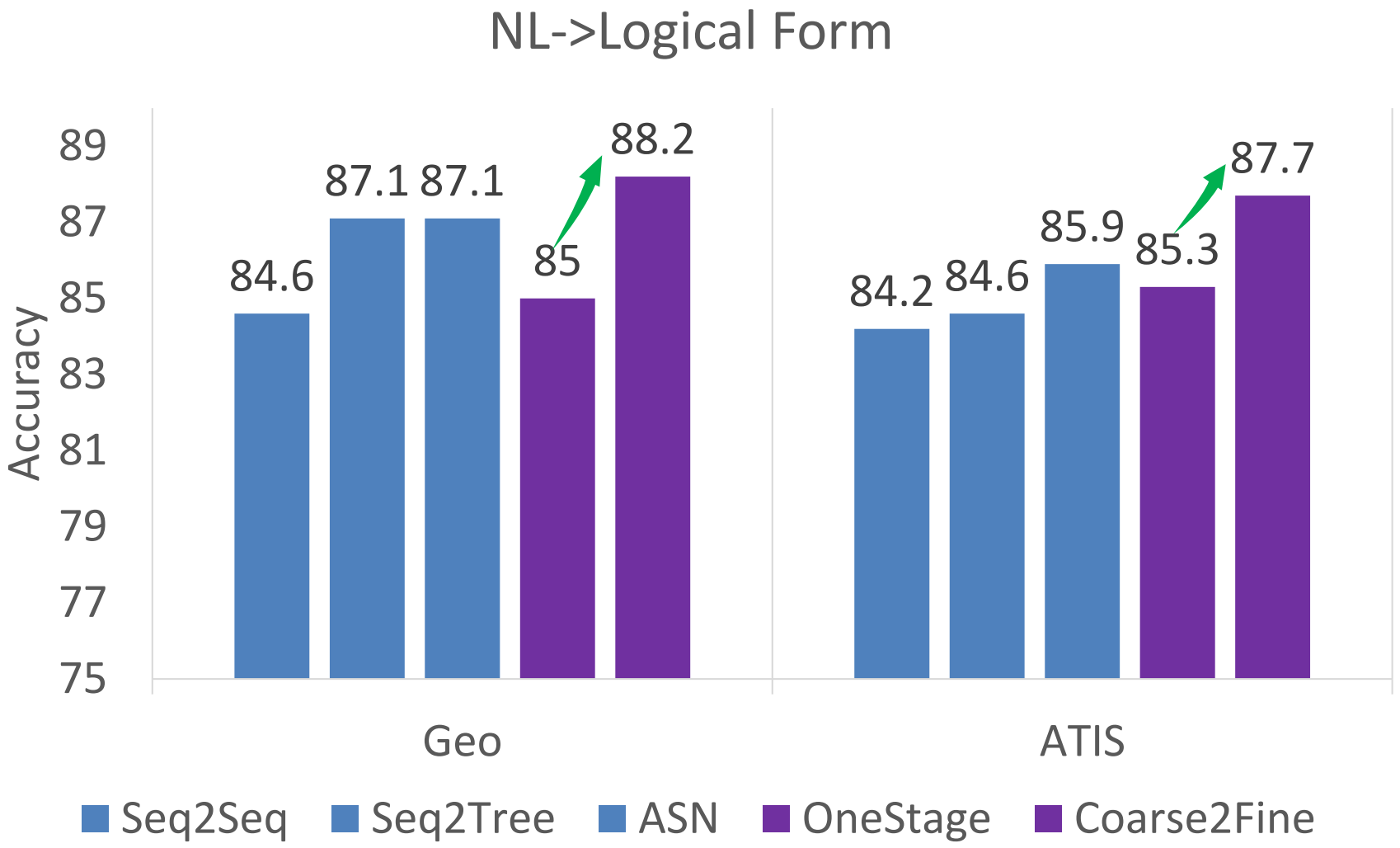
```
SELECT agg_operator agg_column
WHERE (cond_column cond_operator cond_value)
AND ...
```



Experimental Results

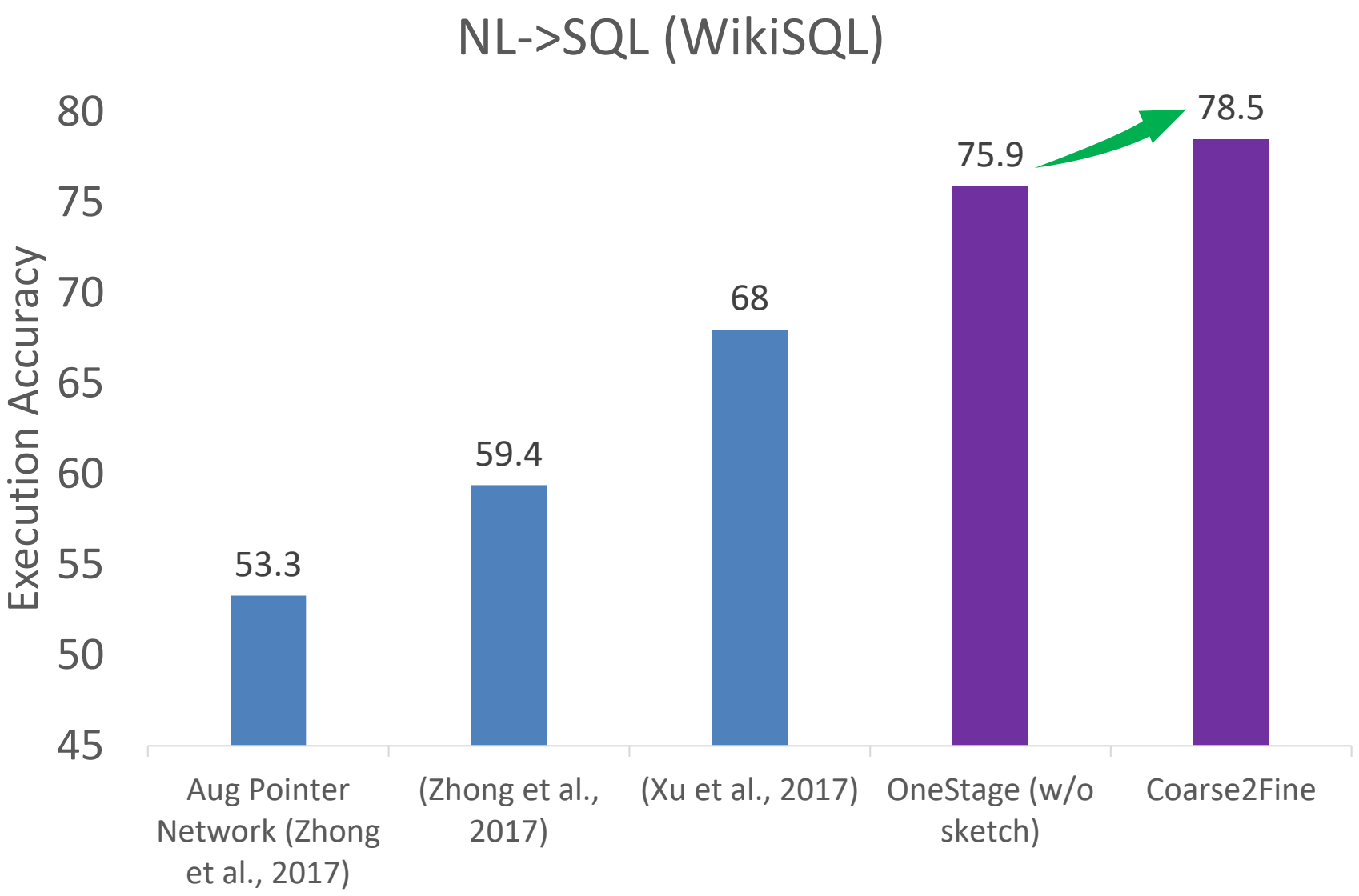


Experimental Results

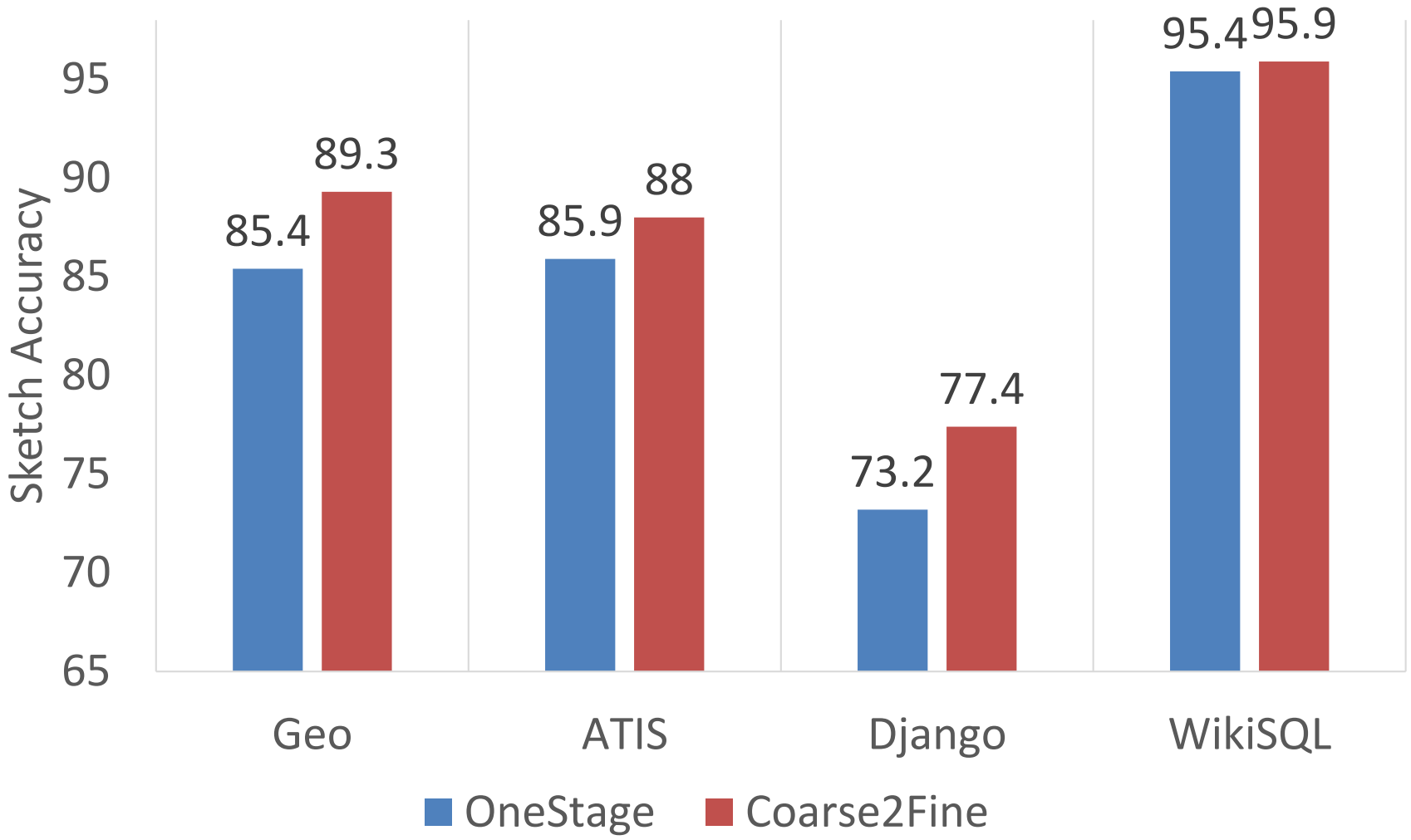


Baseline: (Dong and Lapata, 2016; Rabinovich et al., 2017)

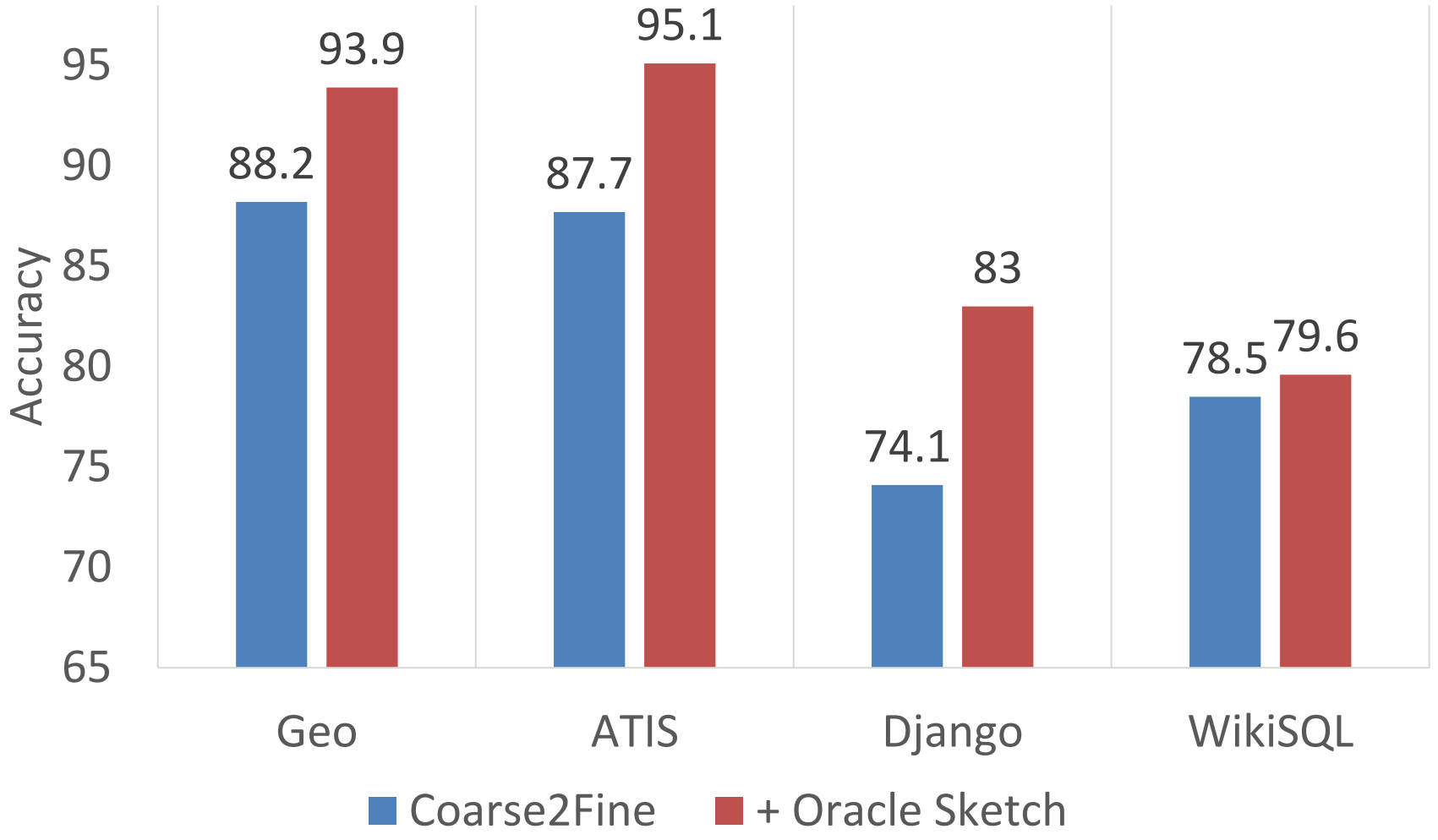
Experimental Results



Sketch Accuracy



Oracle Meaning Sketch



Future Work

- Alternative ways of defining meaning sketches
  - Different levels of granularity
- Weakly supervised setting
  - Meaning sketch reduces search space
- Partial annotation
  - Only annotate meaning sketches for some examples

Thanks!

Q&A

Code Available:
http://homepages.inf.ed.ac.uk/s1478528