

Syntactic Question Abstraction and **Retrieval** for **Data-Scarce Semantic Parsing**

Wonseok Hwang

Jinyeong Yim

Seunghyun Park

Minjoon Seo

Clova AI, NAVER Corp.

WONSEOK.HWANG@NAVERCORP.COM

JINYEONG.YIM@NAVERCORP.COM

SEUNG.PARK@NAVERCORP.COM

MINJOON.SEO@NAVERCORP.COM

Wonseok Hwang

Task

Task: Question → SQL query

	#	Player	Country	Score	To par	Points	Winnings (\$)
0	1	Steve Stricker	United States	67-67-65-69=268	-16	9000	1260000
1	2	K.J. Choi	South Korea	64-66-70-70=270	-14	5400	756000

Q: What is the points of South Korea player?

L: select Points where Country = South Korea
with few labels as possible

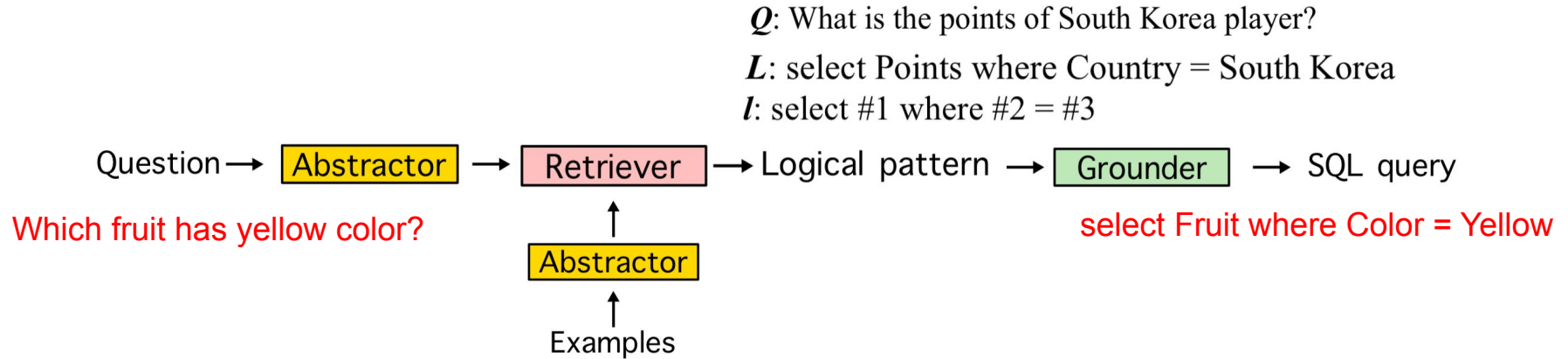
Model?

→ **Retrieve partial query** from "examples"

Which part should be retrieved?

Can we leverage low-cost datasets?

Model



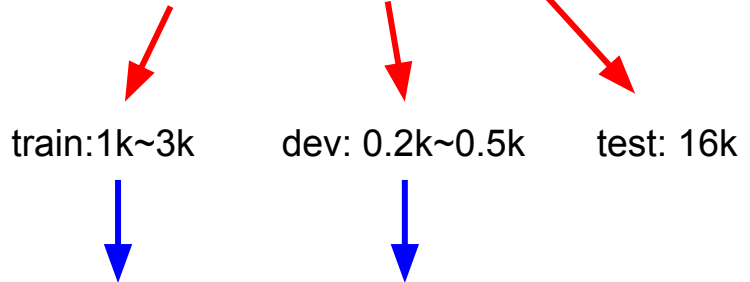
Dataset construction under low-resource condition

Q) Ideal data distribution in train set?

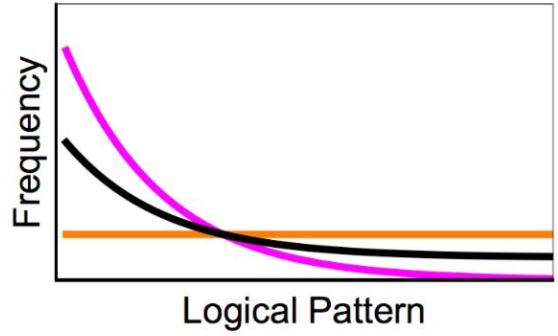
Q) Can we leverage other low-cost dataset?

Dataset

WikiSQL (<https://github.com/salesforce/WikiSQL>)
(train: 56k, dev: 8.4k, test: 16k)



— Uniform — Random — Hybrid



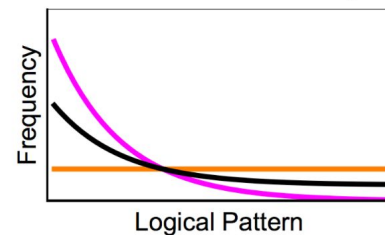
Quora Question Pairs for pretraining

Should I buy tiago?	What keeps children active and far from phone and video games?	0
How can I be a good geologist?	What should I do to be a great geologist?	1

Results

Model	Train set	Dev set	P (%)	LF (%)
COARSE2FINE ^a	Train-Rand-881	Dev-Rand-132	-	2.1 ± 0.0
SQLOVA-GLOVE ^b	Train-Rand-881	Dev-Rand-132	66.6 ± 0.4	17.6 ± 0.3
SQLOVA ^b	Train-Rand-881	Dev-Rand-132	75.3 ± 0.4	45.1 ± 0.7
SQAR w/o Quora	Train-Rand-881	Dev-Rand-132	74.1 ± 0.8	49.1 ± 0.9
SQAR	Train-Rand-881	Dev-Rand-132	75.5 ± 0.6	50.0 ± 0.6

— Uniform — Random — Hybrid



Uniform: +4.1%
Hybrid: +4.0%

	Train-Rand-881	Train-Uniform-85P-850	Train-Hybrid-85P-897
Test-Original	50.0 ± 0.6	37.9 ± 0.4	49.9 ± 1.1
Test-Uniform	17.2 ± 1.3	39.2 ± 1.2	37.6 ± 1.7

Results

Can SQAR parse unseen logical patterns?

Model	Train set	Set for retrieval	P (%)	LF (%)
SQAR	R-881	R-881	75.5 ± 0.6	50.0 ± 0.6
SQAR	R-881	R-881 + H-897	76.6 ± 0.4	50.6 ± 0.5

Summary

- SOTA under data-scarce environment (up to +4.9%).
- Natural language query similarity dataset can be used for semantic parsing (up to +5.9%).
- Careful design of the train set is important.
- Retrieval-based parser can handle unseen new logical pattern.

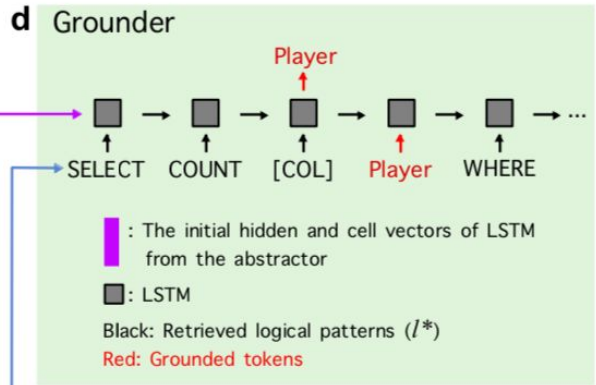
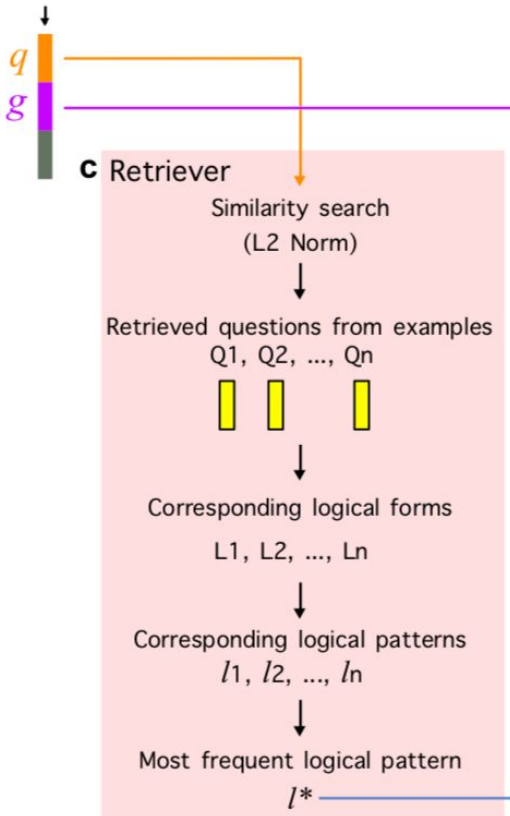
Appendix

Model

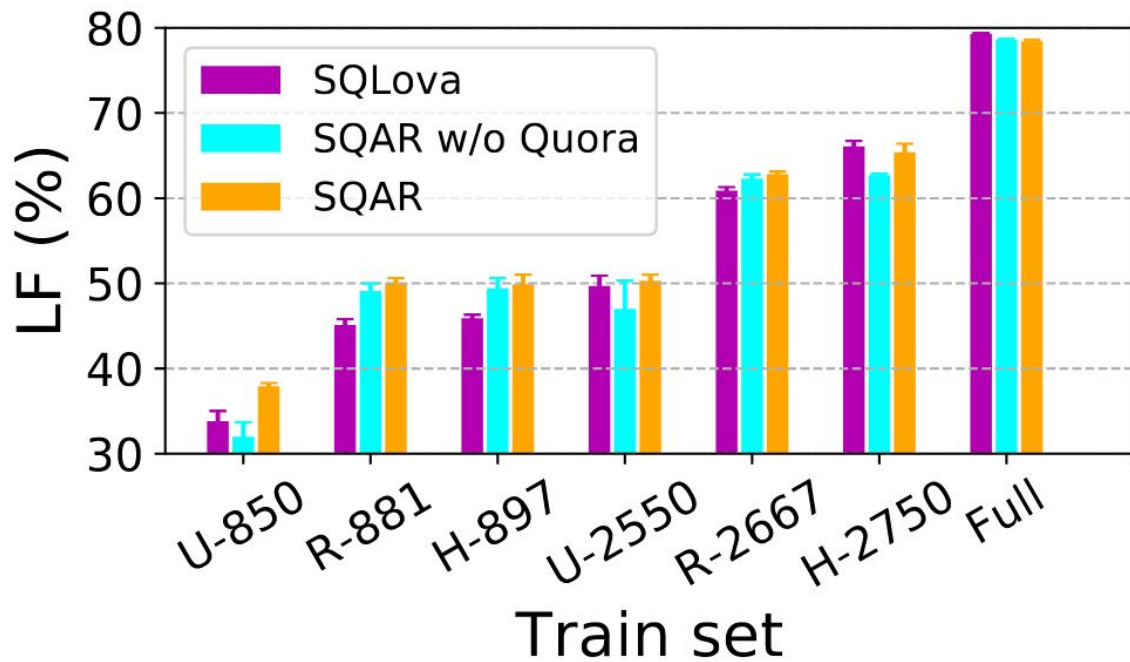
[CLS] E [SEP] Q [SEP] H [SEP]

↓ ↓ ↓ ↓ ↓ ↓ ↓ ↓

Table-aware BERT Encoder



Model	Train set	Dev set	P (%)	LF (%)
SQLOVA	Train-Rand-2677	Dev-Rand-527	81.2 ± 0.2	60.9 ± 0.4
SQAR w/o Quora	Train-Rand-2677	Dev-Rand-527	82.0 ± 0.2	62.3 ± 0.5
SQAR	Train-Rand-2677	Dev-Rand-527	81.4 ± 0.5	62.8 ± 0.3
SQLOVA	Train-Uniform-85P-2550	Dev-Uniform-80P-320	68.2 ± 1.6	49.7 ± 1.2
SQAR w/o Quora	Train-Uniform-85P-2550	Dev-Uniform-80P-320	66.2 ± 4.5	47.0 ± 3.3
SQAR	Train-Uniform-85P-2550	Dev-Uniform-80P-320	69.0 ± 1.2	50.3 ± 0.7
SQLOVA	Train-Hybrid-96P-2750	Dev-Hybrid-446	83.1 ± 0.2	66.1 ± 0.6
SQAR w/o Quora	Train-Hybrid-96P-2750	Dev-Hybrid-446	82.2 ± 0.2	62.7 ± 0.2
SQAR	Train-Hybrid-96P-2750	Dev-Hybrid-446	82.8 ± 0.4	65.4 ± 1.0

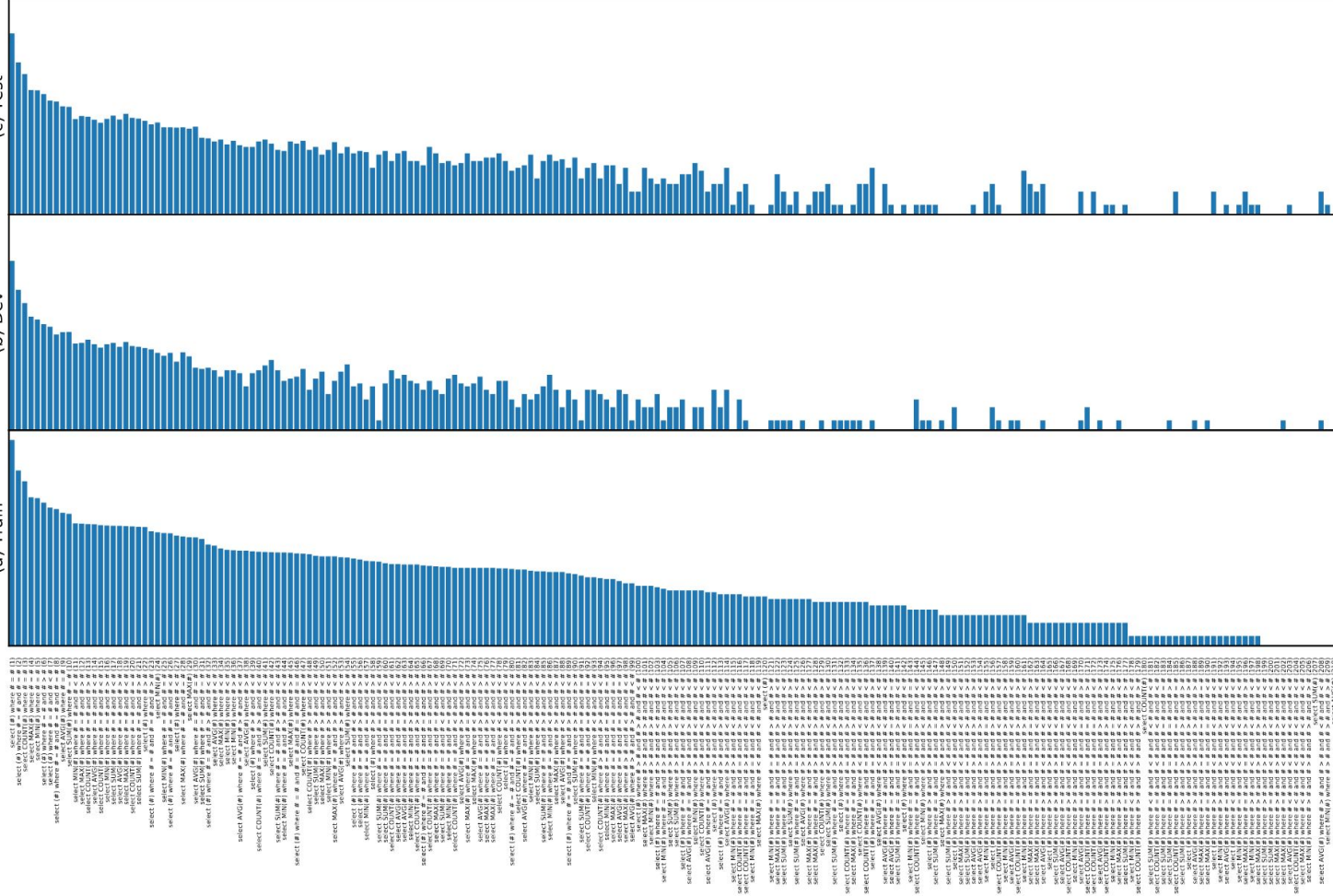


Model	Train set	Set for retrieval	P (%)	LF (%)	R-capacity	RG-capacity
SQAR	R-881	R-881	75.5 ± 0.6	50.0 ± 0.6	57.5 ± 2.2	47.3 ± 0.4
SQAR	R-881	R-881 + H-897	76.6 ± 0.4	50.6 ± 0.5	79.3 ± 1.9	58.8 ± 3.9
SQAR	R-881	R-881 + H-2750	77.5 ± 0.4	50.7 ± 0.6	91.0 ± 2.3	67.0 ± 1.7
SQAR	R-881	Full	79.6 ± 0.4	51.7 ± 0.5	102 ± 2	67.3 ± 2.5
SQAR	R-881	H-2750	77.2 ± 0.5	50.5 ± 0.5	92.0 ± 2.1	67.5 ± 2.2

(a) Train

(b) Dev

(c) Test



Patterns

Counts