The Joint Conference of the 59th Annual Meeting of the Association for Computational Linguistics and the 11th International Joint Conference on Natural Language Processing (ACL-IJCNLP 2021)

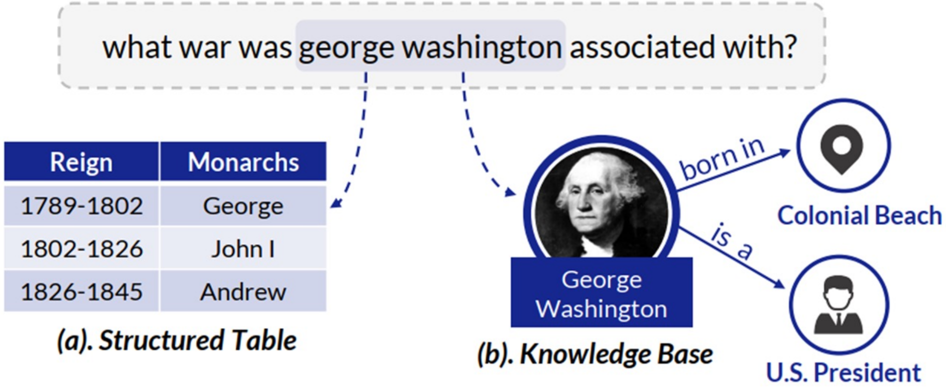
Awakening Latent Grounding from Pretrained Language Models for Semantic Parsing

Qian Liu[1], Dejian Yang[2], Jiahui Zhang[1],
Jiaqi Guo[3], Bin Zhou[1], Jian-Guang Lou[2]

[1] Beihang University; [2] Microsoft Research
[3] Xi'an Jiaotong University

Background

- Recent years pretrained language models (PLMs) hit a success on several downstream tasks, showing their power on modeling language.
- However, few efforts have been made to explore grounding capabilities of PLMs, which aims at connect linguistic symbols to real-world perception.



(a). Structured Table

(b). Knowledge Base

Background

- Existing methods are not satisfying:
  - ✓ Heuristic rules suffer from modeling flexibility => poor performance
  - ✓ Data driven models requires expensive annotations => not available in real

Can We Learn A Flexible Grounding Model Without Grounding Supervision?

Method

- YES! Our approach **Erasing-then-Awakening** (ETA) could automatically obtain grounding results from a pretrained language model with only weak supervision.

Illustration

- Given a question, we use a module to evaluate the confidence of a column being mentioned in a question, and the module is trained using text-to-SQL datasets.

Question	Column	Confidence
How many total games were at braly stadium	Venue	0.92

Illustration

- We erase "How" and feed the erased question into the module to obtain the confidence again and calculate the confidence difference after erasing "How".

Question	Column	Confidence
How many total games were at braly stadium	Venue	0.92

Erased Question	Column	Confidence
~~How~~ many total games were at braly stadium	Venue	0.91

Illustration

- We erase "How" and feed the erased question into the module to obtain the confidence again and calculate the confidence difference after erasing "How".

Question	Column	Confidence
How many total games were at braly stadium	Venue	0.92

Erased Question	Column	Confidence
~~How~~ many total games were at braly stadium	Venue	0.91

Difference

0.01

Illustration

- By erasing tokens in the question one by one, we can find which tokens are important for a model when predicting the column “Venue”.

Question	Column
How many total games were at braly stadium	Venue

0.01

Illustration

- By erasing tokens in the question one by one, we can find which tokens are important for a model when predicting the column “Venue”.

Question

How many total games were at braly stadium

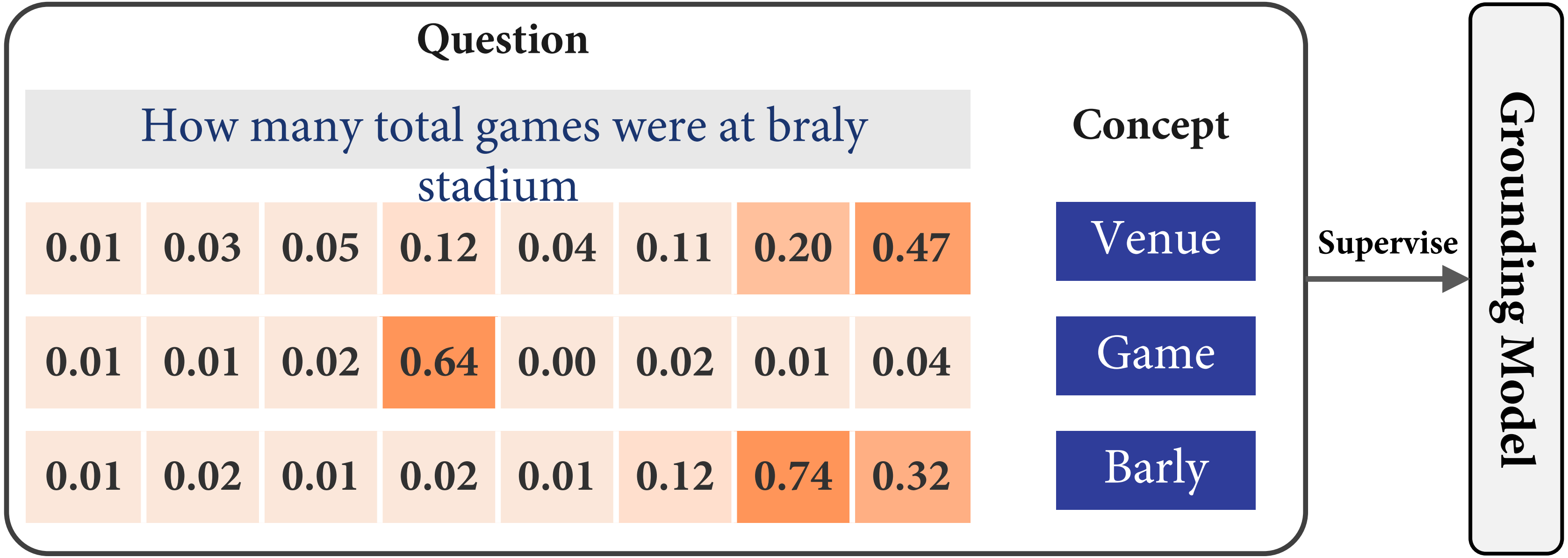
Column

Venue

0.01	**0.03**	**0.05**	**0.12**	**0.04**	**0.11**	**0.20**	**0.47**

Illustration

- Repeating the above steps on the input question tokens for all concepts. And then we could employ these differences to awaken the latent grounding.



Experiments

- Taking text-to-SQL as a case study, we couple our approach with two off-the-shelf parsers and improve their performance by a large margin.
- Finally, our approach could achieve similar or better performance on two challenging benchmarks WikiTableQuestions & Spider.

Model	Dev		Test
	Ex.Match	Ex.Acc	Ex.Acc
ALIGN_P	37.8 ± 0.6	56.9 ± 0.7	46.6 ± 0.5
ALIGN_P + BERT	44.7 ± 2.1	63.8 ± 1.1	51.8 ± 0.4
ETA + BERT	$\mathbf{47.6} \pm 2.5$	$\mathbf{66.6} \pm 1.7$	$\mathbf{53.8} \pm 0.3$
ALIGN♡	42.2 ± 1.5	61.3 ± 0.8	49.7 ± 0.4
ALIGN + BERT♡	47.2 ± 1.2	66.5 ± 1.2	54.1 ± 0.2

Model	Dev	Test
IRNet + BERT (Guo et al., 2019)	61.9	54.7
IRNet v2 + BERT (Guo et al., 2019)	63.9	55.0
BRIDGE + BERT_L (Lin et al., 2020)	70.0	65.0
RATSQL + BERT_L (Wang et al., 2020a)	69.7	**65.6**
SLSQL_P + BERT	57.4	-
SLSQL_P + BERT_L	61.0	-
ETA + BERT	64.5	59.5
ETA + BERT_L	**70.8**	65.3