Language Model Prompting 3

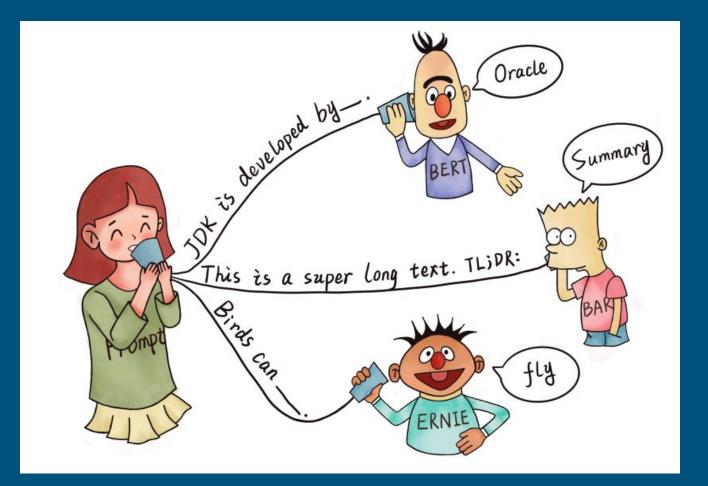
Wei-Rui Chen

Outlines

Language model prompting brief recap

The Power of Parameter-Efficient Prompt Tuning

SEQZERO: Few-shot Compositional Semantic Parsing with Sequential Prompts and Zero-shot Models



Formulation

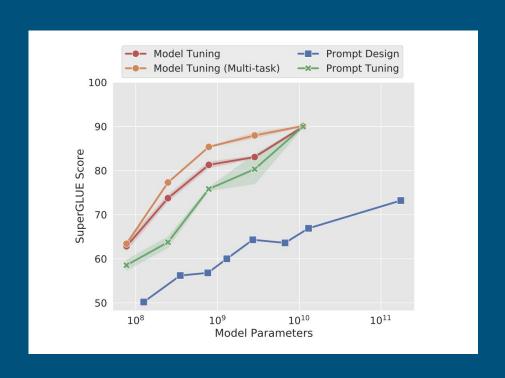
Name	Notation	Example	Description
Input	$oldsymbol{x}$	I love this movie.	One or multiple texts
Output	$oldsymbol{y}$	++ (very positive)	Output label or text
Prompting Function	$f_{ ext{prompt}}(oldsymbol{x})$	[X] Overall, it was a [Z] movie.	A function that converts the input into a specific form by inserting the input x and adding a slot $[Z]$ where answer z may be filled later.
Prompt	$oldsymbol{x}'$	I love this movie. Overall, it was a [Z] movie.	A text where $[X]$ is instantiated by input x but answer slot $[Z]$ is not.
Filled Prompt	$f_{\mathrm{fill}}(oldsymbol{x'},oldsymbol{z})$	I love this movie. Overall, it was a bad movie.	A prompt where slot [Z] is filled with any answer.
Answered Prompt	$f_{\mathrm{fill}}(oldsymbol{x'},oldsymbol{z}^*)$	I love this movie. Overall, it was a good movie.	A prompt where slot [Z] is filled with a true answer.
Answer	z	"good", "fantastic", "boring"	A token, phrase, or sentence that fills [Z]

Table 2: Terminology and notation of prompting methods. z^* represents answers that correspond to true output y^* .

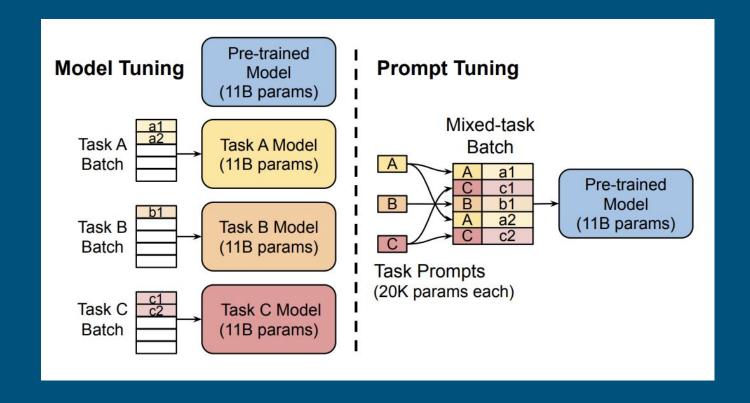
Source: Liu et al. 2021

The Power of Parameter-Efficient Prompt Tuning Lester et al. 2021

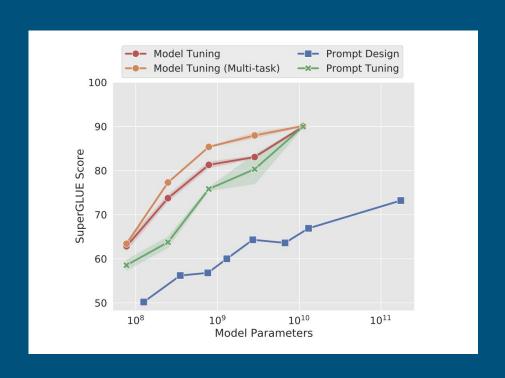
Results



Less computing, comparable performance to pretrain-finetune



Results

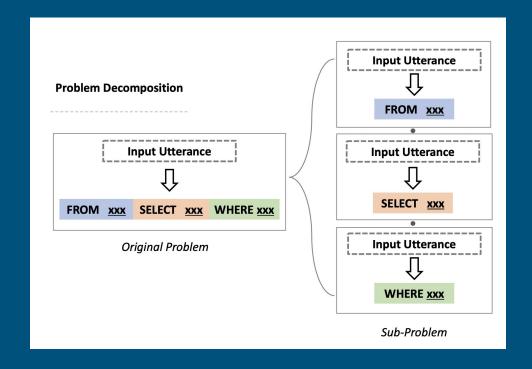


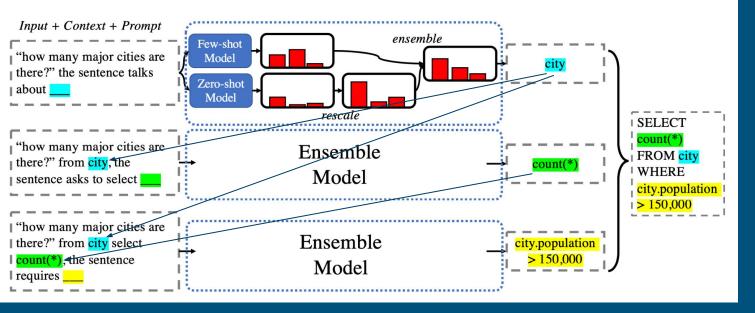
SEQZERO: Few-shot Compositional Semantic Parsing with Sequential Prompts and Zero-shot Models Yang et al. 2022

Goal

Question:	
how many peo	ple live in Utah ?
Gold SQL:	
SELECT state	. population FROM state
WHERE state	. state _name = "Utah"
Finetuned B	ART Predicted SQL:
SELECT city.	population FROM city
	city_name = "Utah"

Problem Decomposition





Ensemble of two models

Few-shot model + zero-shot model

Few-shot: a PLM fine-tuned with this utterance+prompt -> SQL query data

then used to infer given utterance+prompt

Zero-shot: directly infers SQL query given utterance+prompt

Why zero-shot?

Due to overfitting and Out-of-distribution (OOD) issue for few-shot

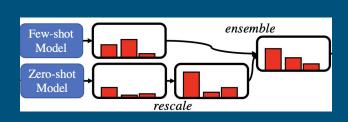


Ensemble and rescaling

Rescaling the zero-shot model probability distribution

V and V_i: V is Vocabulary, V_i is the constrained vocabulary for ith sub-clause

 $V_i \subseteq V$, $|V_i| \ll V$ where V_i is expert-built with SQL grammar, schema, and training data



$$P_{\theta_{i,z}}(w|x) = \frac{\mathbb{1}(w \in V_i(x))P_{\theta_0}(w|x)}{\sum_{w_j \in V_i(x)} P_{\theta_0}(w_j|x)}$$

$$P_{\theta_i} = \gamma_i P_{\theta_{i,f}} + (1 - \gamma_i) P_{\theta_{i,z}}$$

Gamma is determined by grid search on dev set

Results

Method	GeoQuery	EcoQuery
SEQZERO	74.7	46.2
-SEQ	74.2	44.5
-Zero	71.4	37.7

Predicted Examples

2	
Cases	Text
Question BART SEQZERO Ground Truth	<pre>what is the population of utah SELECT city . population FROM city WHERE city . city_name = "utah" SELECT state . population FROM state WHERE state . state_name = "utah" SELECT state . population FROM state WHERE state . state_name = "utah"</pre>
Question BART SEQZERO Ground Truth	petrol trimmer over 100 dollar SELECT * FROM ASINs WHERE Maching Algorithm("petrol trimmer") == True and Price < 100 SELECT * FROM ASINs WHERE Maching Algorithm("petrol trimmer") == True and Price > 100 SELECT * FROM ASINs WHERE Maching Algorithm("petrol trimmer") == True and Price > 100
Question BART SEQZERO Ground Truth	mi4 64 gb mobile phone SELECT * FROM ASINs WHERE Maching Algorithm("mi4 64 gb mobile phone") ORDER BY date SELECT * FROM ASINs WHERE Maching Algorithm("mi4 64 gb mobile phone") ORDER BY date SELECT * FROM ASINs WHERE Maching Algorithm("mi4 mobile phone") and Size = 64 gb

Table 8: Case study. The first example is from GeoQuery, and the last two examples are from EcoQuery.