Constrained Language Models Yield Few-Shot Semantic Parsers

Richard Shin, Christopher H. Lin, Sam Thomson, Charles Chen, Subhro Roy, Emmanouil Antonios Platanios, Adam Pauls, Dan Klein, Jason Eisner, Benjamin Van Durme

Semantic parsing

What time am I getting coffee with Megan? → Model → `findEvent( subject="coffee", attendees=includes( findPerson(name="Megan") ) ).start.time`

Utterance — Meaning representation

Goal: low-resource semantic parsing

Training data

"What time am I getting coffee with Megan?" → `findEvent( subject="coffee", attendees=includes( findPerson(name="Megan") ) ).start.time`

"Who's coming to the weekly standup?" → `findEvent( subject="weekly standup" ).attendees`

...

× ~~10^5–10^6~~ (tens of thousands)

× 10^2–10^3 (hundreds)

Motivation: prototyping domains, developing new features

Pre-trained language models

BART GPT-3

🏆 State-of-the-art performance on many tasks

What time am I getting coffee with Megan? → `findEvent( subject="coffee", attendees=includes( findPerson(name="Megan") ) ).start.time`

Utterance — Meaning representation

What is the best way to leverage pre-trained language models for low-resource semantic parsing?

Pitfalls:
1. Pre-trained on language, rather than on meaning representations (code)
2. LMs don't know what functionality is available in the domain

Approach

Learn to **map to canonical utterances**

"What time am I getting coffee with Megan?" (Utterance) → "What is the start time of the event titled "coffee" with attendee "Megan"?" (Canonical utterance) ↔ `findEvent( subject="coffee", attendees=includes( findPerson(name="Megan") ) ).start.time` (Meaning representation)

Deterministic mapping ↔ SCFG

Paraphrasing into a controlled sublanguage

SCFG:

S → What is the start time of **E**?	↔ **E**.start.time
E → the event **P***	↔ findEvent(**P***)
P → titled "**T**"	↔ subject="**T**"
P → with attendee "**N**"	↔ attendees=includes(findPerson(name="**N**"))

Dynamically select relevant examples for GPT-3 prompt

u_1 … u_n — *training examples*; u — *test utterance*

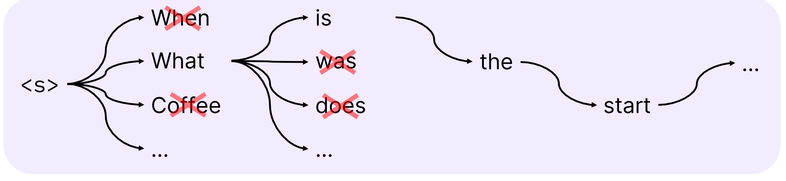
Find the top 20 examples by: $P_{\text{GPT-3}}(u \mid u_i)$

Prompt-based learning: $P_{\text{GPT-3}}(\ldots \mid$

```
Input: What time am I getting coffee with Megan?
Output: What is the start time of the event titled "Coffee" with attendee "Megan"?
Input: Who's coming to the weekly standup?
Output: Who is attending the event titled "weekly standup"?
...
Input: Who will join the event with Damon?
Output:
```

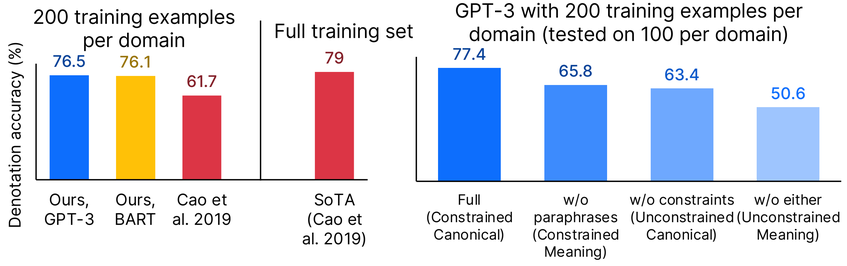
Top 20 training examples / *Test utterance* $)$

Constrained decoding to predict canonical utterance covered by SCFG

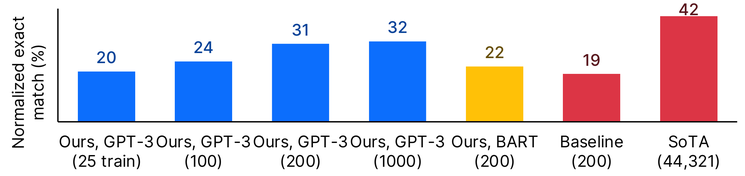


Case studies

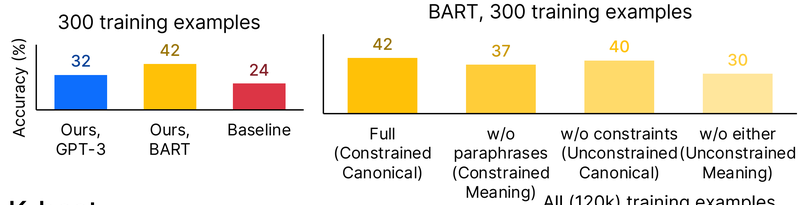
Overnight: semantic parsing, contains 8 different domains



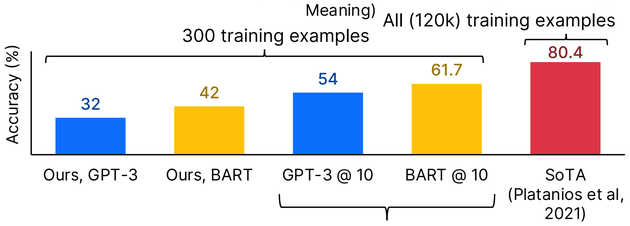
Break: question understanding, aggregates 10 benchmarks



SMCalFlow: task-oriented dialogue dataset



K-best lists for data collection



Annotators can find the correct output among 10 choices more than half the time