

Multitask Prompted Training Enables Zero-Shot Task Generalization

Victor Sanh, Albert Webson, Colin Raffel, Stephen H. Bach
& the BigScience crew
ICLR 2022

Presented by Giuseppe Attanasio

Multitask Prompted Training Enables Zero-Shot Task Generalization

Multitask Prompted Training

Enables Zero-Shot Task Generalization

Solving tasks the model
was not trained for


Training on multiple
tasks using prompts



Multitask Prompted Training

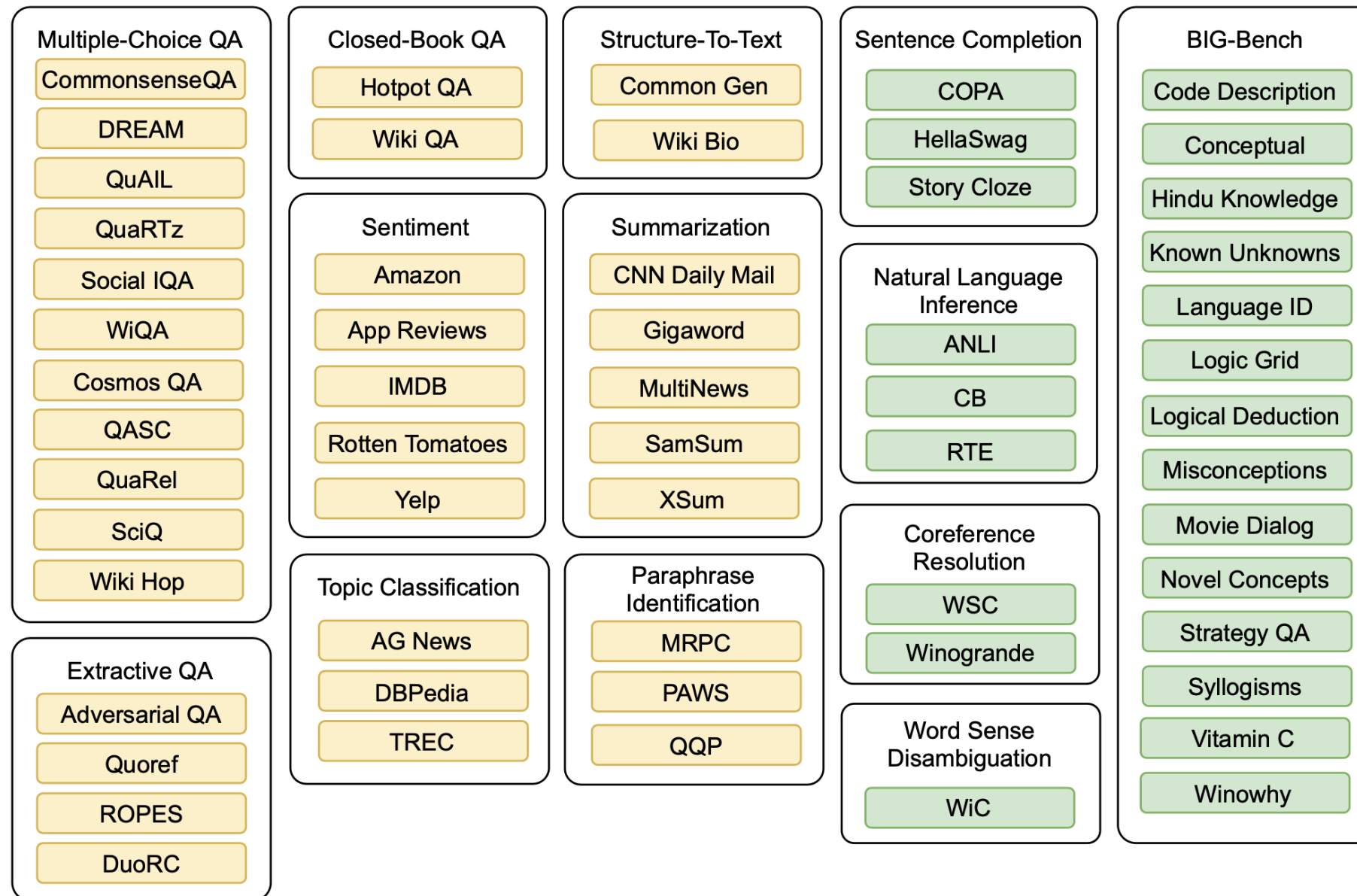
Enables Zero-Shot Task Generalization

Solving tasks the model
was not trained for





Multitask, Zero-Shot



Crazy, right?

DMG

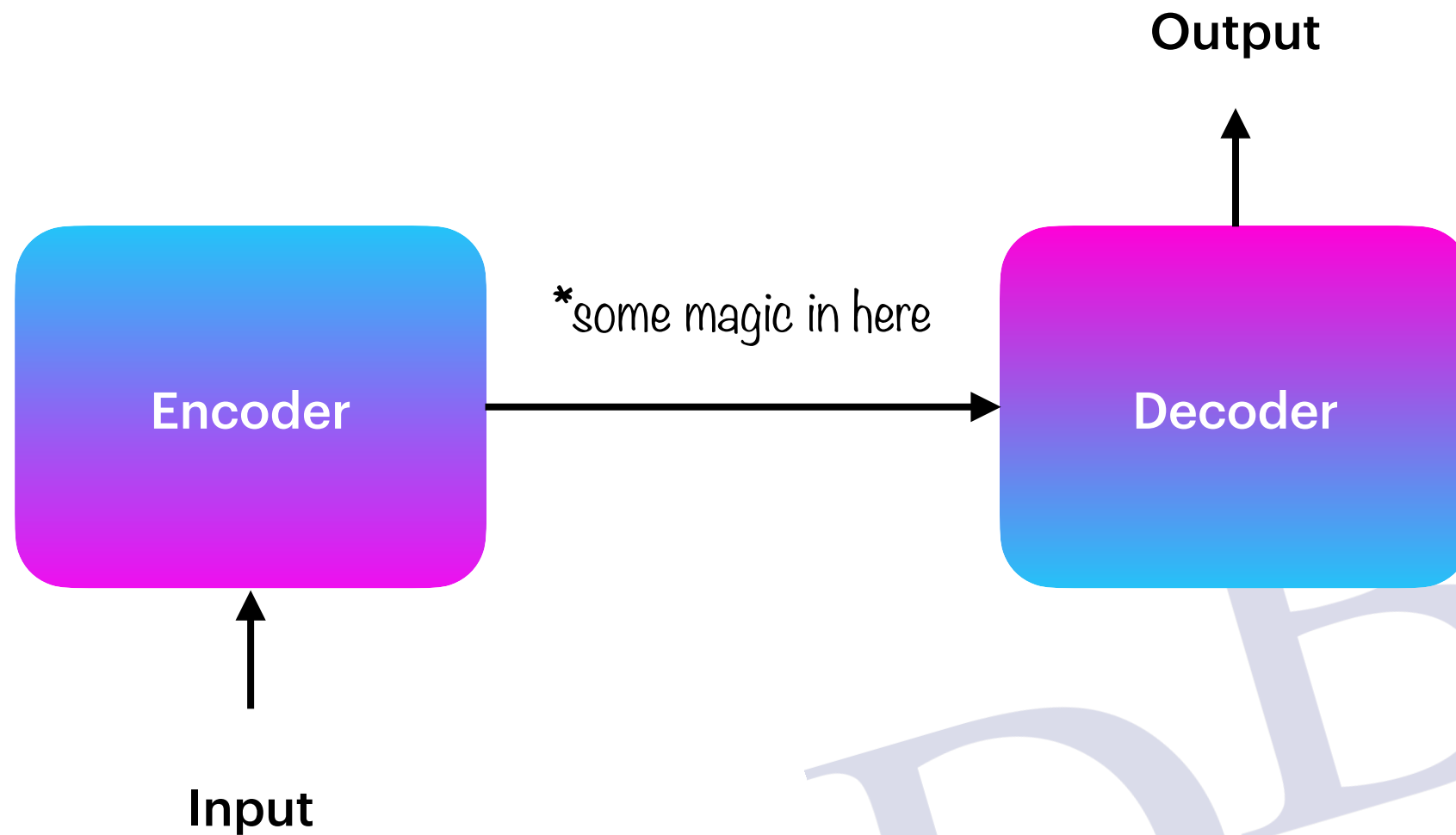
Prompting

(A clever NLP trick)



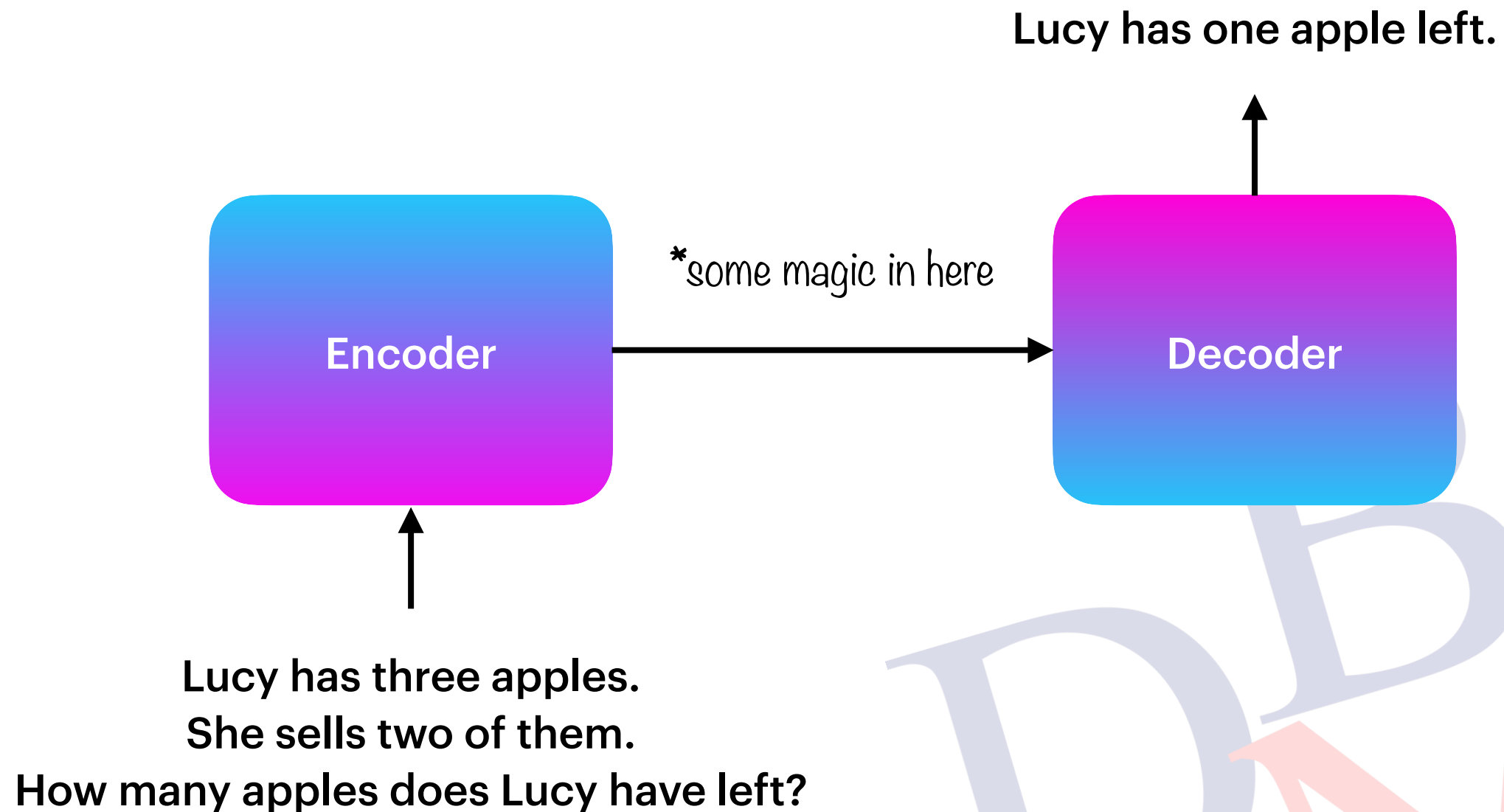
Prompting

Key ingredient: a S2S model



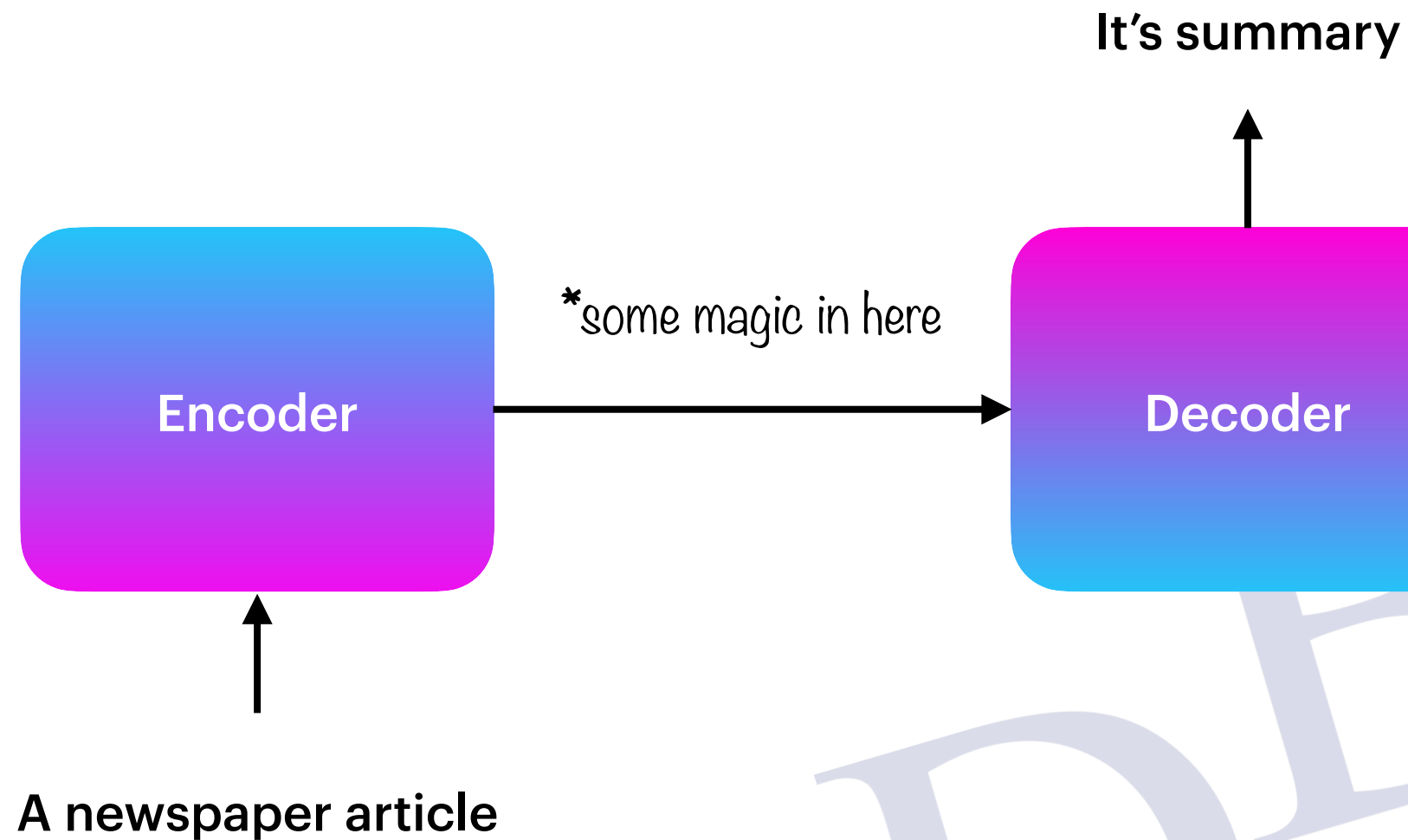
Prompting

Key ingredient: a S2S model



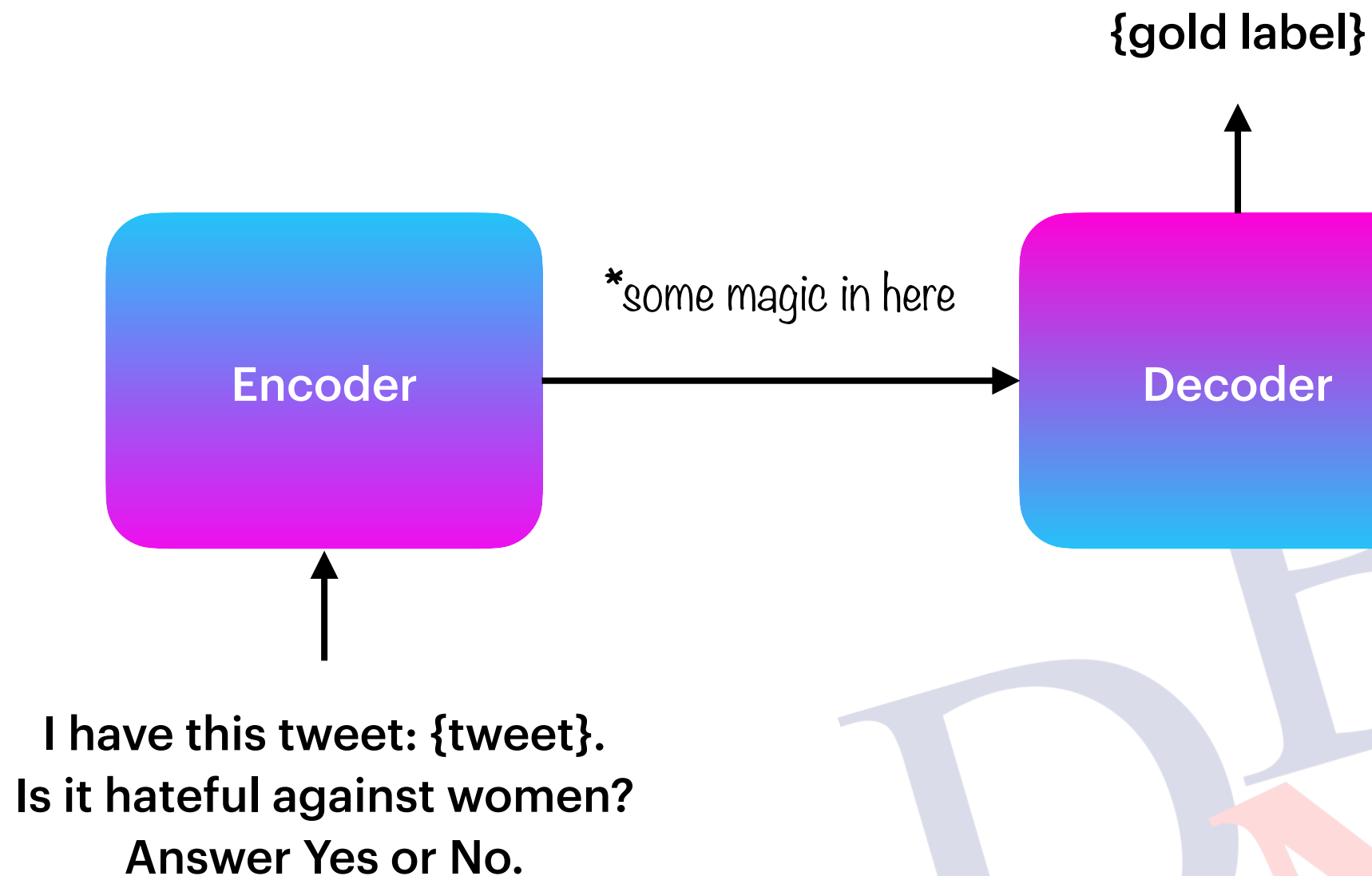
Prompting

Key ingredient: a S2S model

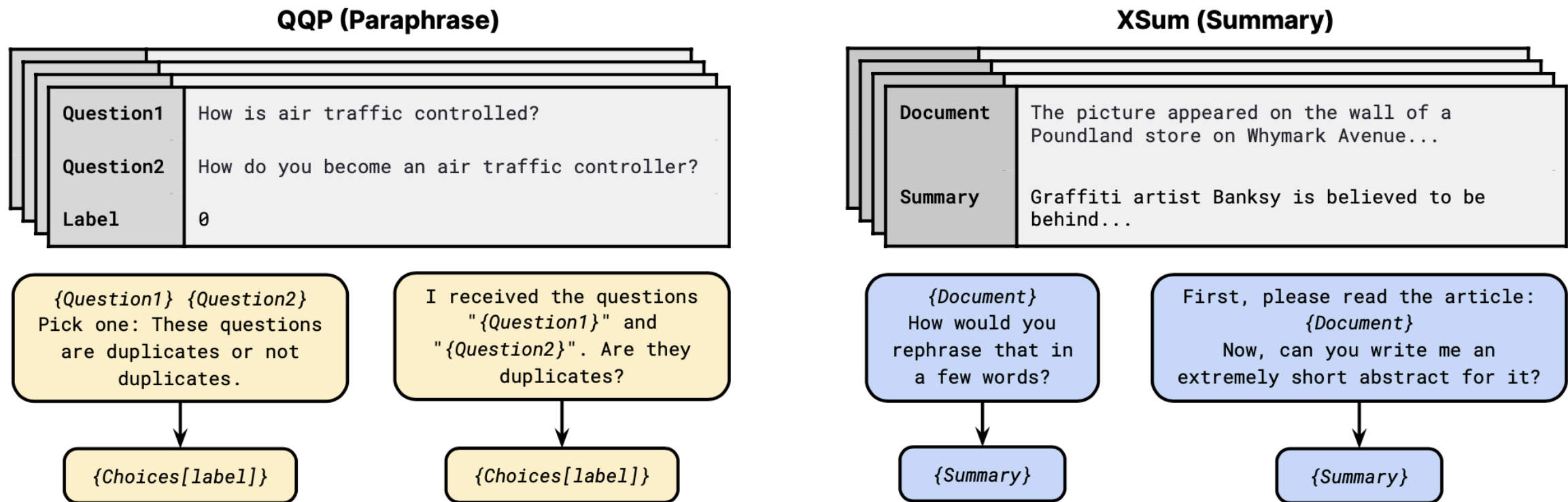


Prompting

Key ingredient: a S2S model



Unified Prompt Format



- Prompt: a input, output template
- Both formal and **creative** prompts
- ~8 prompts per dataset

Amazon Review Polarity

{title}, {content}, {label}

Input Template:

Input Template:

```
You are considering whether to buy a product. You look at the reviews.  
Would the following review {{answer_choices[0]}} or  
{{answer_choices[1]}} the chances of you buying the product?  
Review title: {{title}}  
Product review: {{content}}
```

Target Template:

```
{{answer_choices[label]}}
```

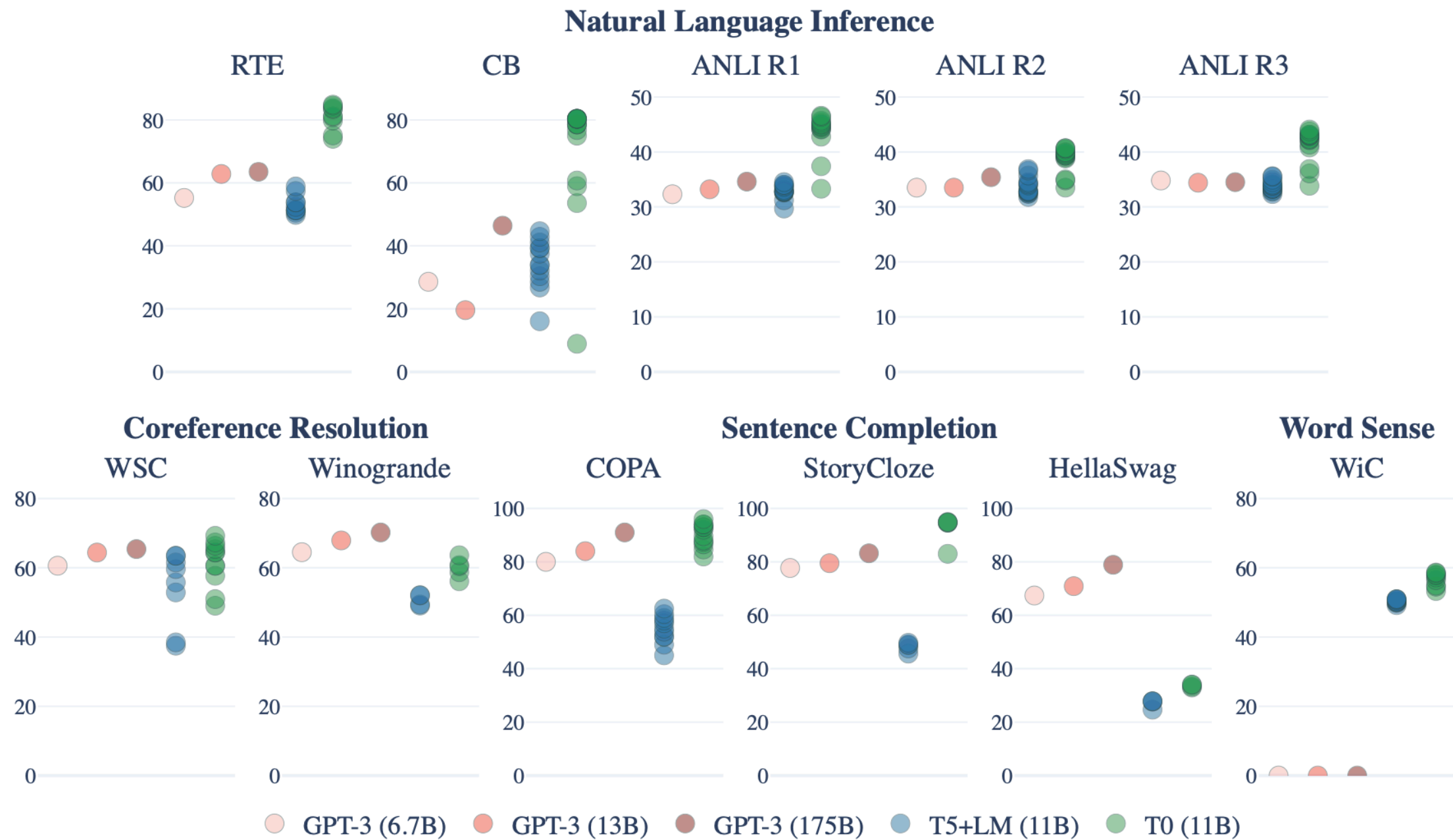
Answer Choices Template: decrease ||| increase

A photograph of two Great Horned Owls in a dark forest at night. The owl on the left is perched on a weathered log, looking directly at the camera. The owl on the right is perched on a branch, also looking forward. The background is filled with dark evergreen trees, and some light filters through the foliage, creating a moody atmosphere. The word "Results" is overlaid in the center in a large, white, sans-serif font.

Results

Held-out tasks

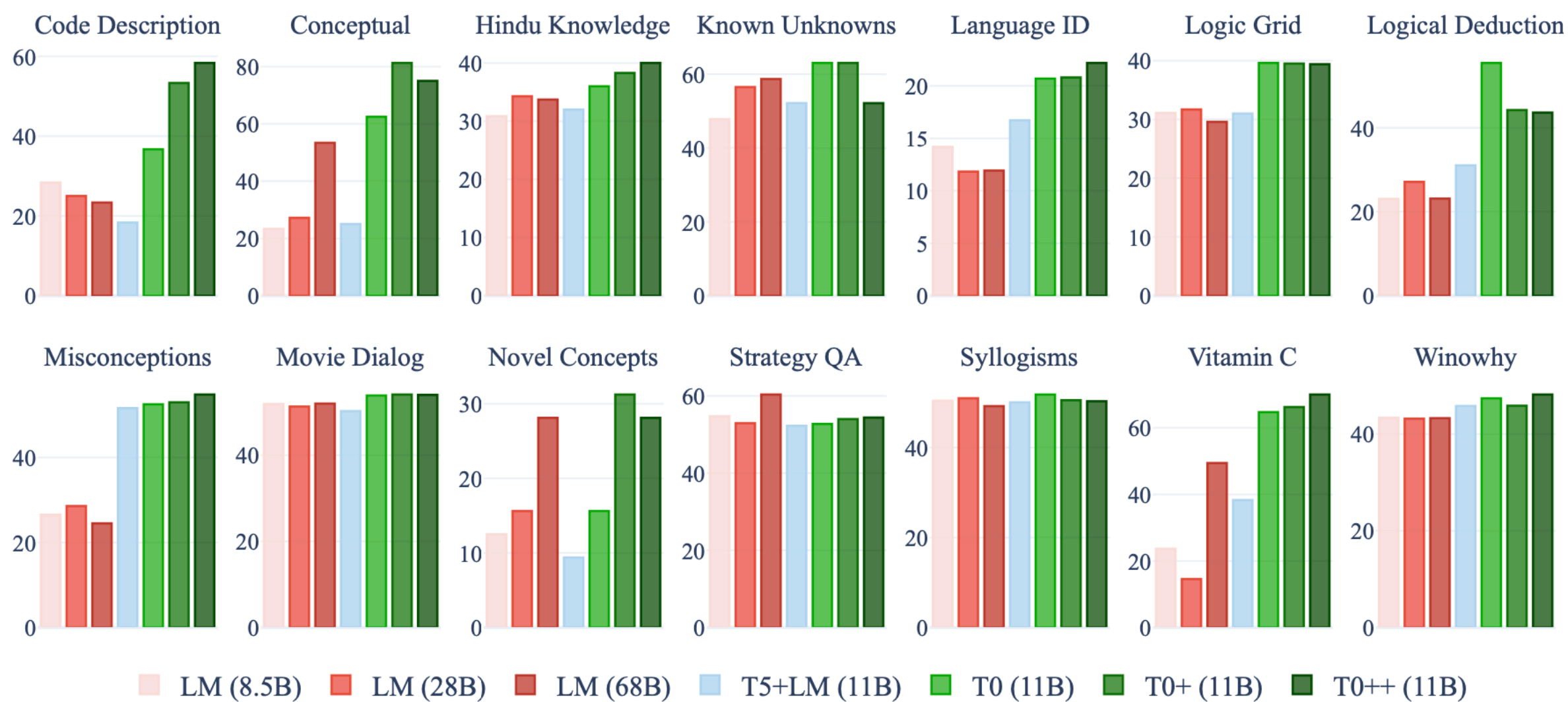
Classic tasks



Held-out tasks

BIG-bench suite

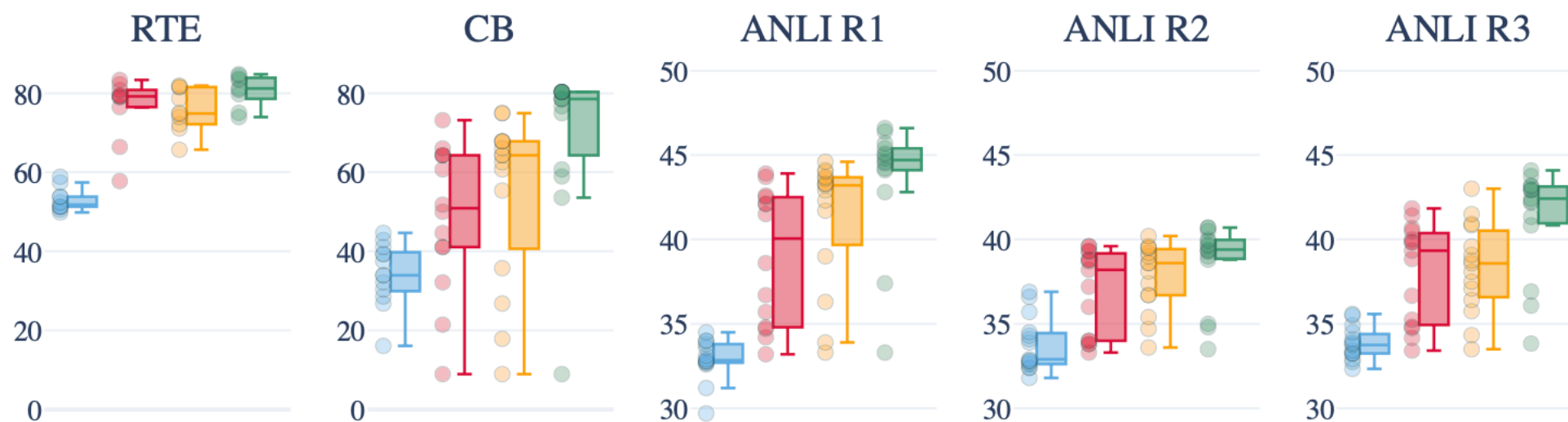
arXiv-ed yesterday!



On the number of prompts

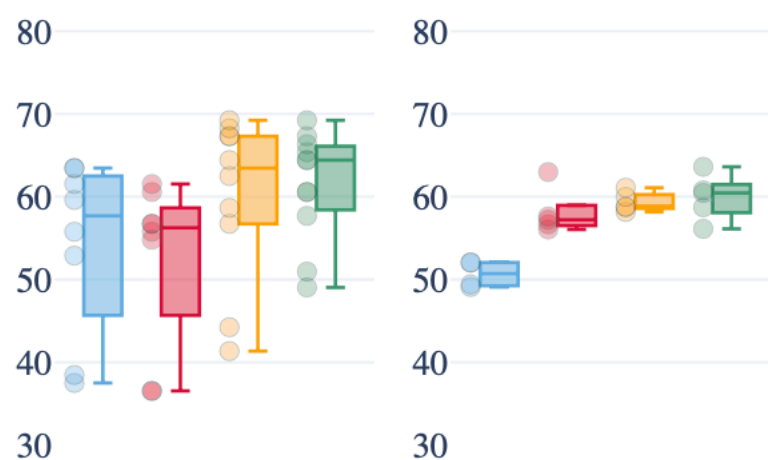
More is better

Natural Language Inference



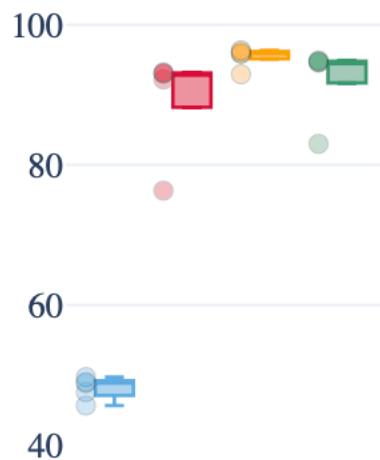
Coreference Resolution

WSC Winogrande



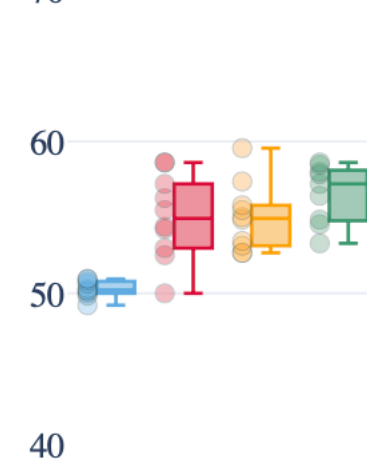
Sentence Completion

StoryCloze



Word Sense

WiC



$p = 0$ (T5+LM) $p = 1$ $p = 5.7$ $p = 8.03$ (T0)

My takes

- If you do NLP, **prompting is a now thing**
- Crazy, beautifully formulated, effective idea
- **Community effort**, one-year long project
- Gap between builders and users
- **Is it true generalization?**

And yours?

