CS 839: Foundation Models
Prompting I

Fred Sala

University of Wisconsin-Madison

Sept. 21, 2023

Announcements

- **Logistics:**
  - Homework 1, info for presentations coming out today
  - Interesting talk: **Copyright's Latent Space: From Fair Use to Generative Art.** BJ Ard. **Wednesday, September 27th 10 AM,** Discovery Building, Orchard View Room

- Class roadmap:

Tuesday Sept. 26	Prompting I
Thursday Sept. 28	Prompting II
Tuesday Oct. 3	Reasoning & Chain-of-Thought
Thursday Oct. 5	In-Context Learning: Practice and Theory
Tuesday Oct. 10	Fine-Tuning, Specialization, Adaptation

Outline

- **Intro to Prompting**
  - Terminology: zero-shot, few-shot, in-context, etc, prompt characteristics: format, examples, orders
- **Hard and Soft Prompting**
  - Searching for good prompts, techniques for continuous/soft prompts
- **Prompt Ensembling and Other Methods**
  - Combinations, majority vote, chain-of-thought introduction, weighted ensembling

Outline

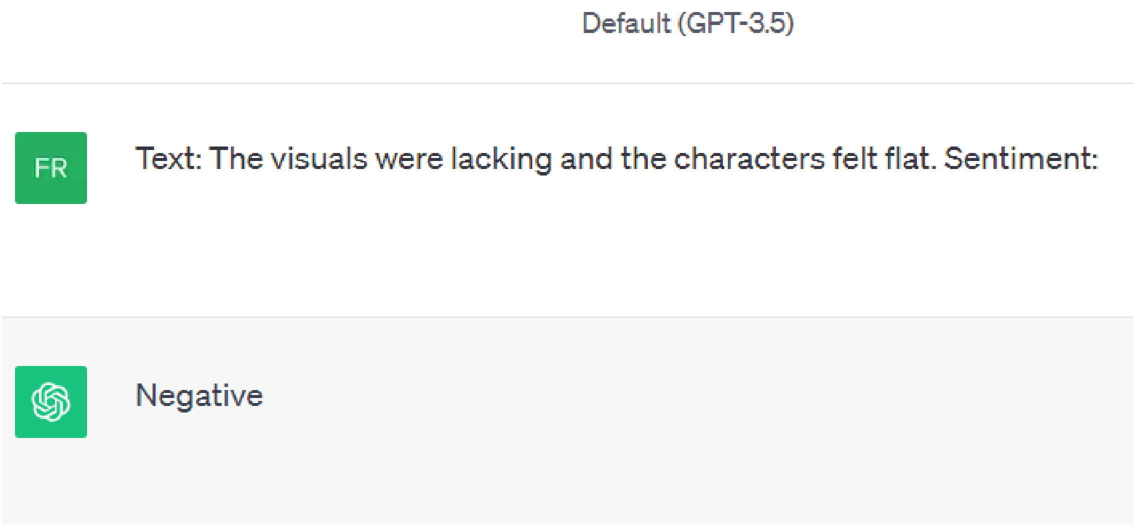
- **Intro to Prompting**
  - Terminology: zero-shot, few-shot, in-context, etc, prompt characteristics: format, examples, orders
- **Hard and Soft Prompting**
  - Searching for good prompts, techniques for continuous/soft prompts
- **Prompt Ensembling and Other Methods**
  - Combinations, majority vote, chain-of-thought introduction, weighted ensembling

Prompting: Ask Your Model

Essentially, ask your model to perform your goal task

Example: sentiment analysis task

- Prompt: "`Text: The visuals were lacking and the characters felt flat. Sentiment:`"
- Result: "`Negative`"



Prompting: Zero-shot vs Few-shot

Terminology:

- **Zero-shot:** No “examples” provided to the model.
- **Few-shot/in-context learning**: Provide “examples”

Input: Subpar acting. Sentiment: Negative
Input: Beautiful film. Sentiment: Positive
Input: Amazing. Sentiment:

Zhao et al '21

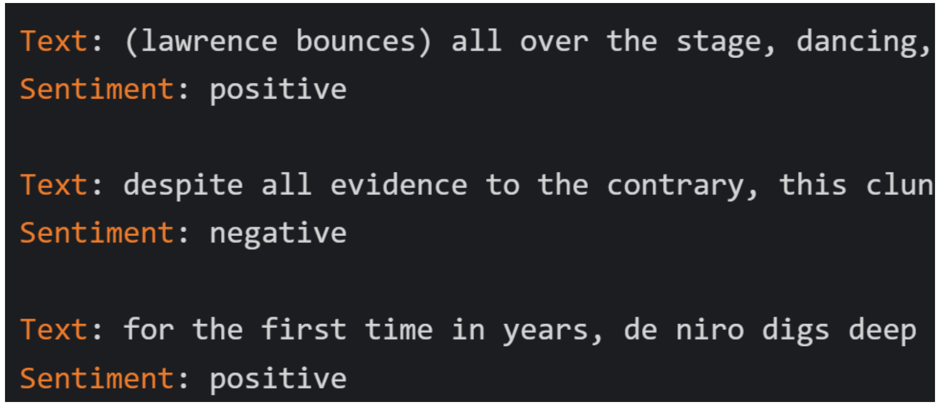
Positive

Prompting: Few-shot vs. In-context learning

Terminology conflicts! Note: we have a set of labeled examples. Could **fine-tune**!

Few-shot: *sometimes* means fine-tune on this dataset, then prompt

In-context learning: do not fine-tune. Model weights unchanged.

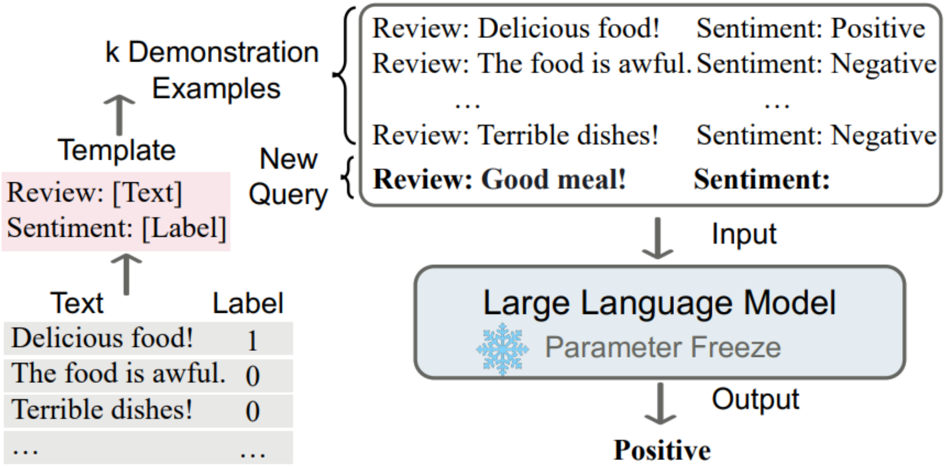


```
Text: (lawrence bounces) all over the stage, dancing,
Sentiment: positive

Text: despite all evidence to the contrary, this clun
Sentiment: negative

Text: for the first time in years, de niro digs deep
Sentiment: positive
```

Weng / SST

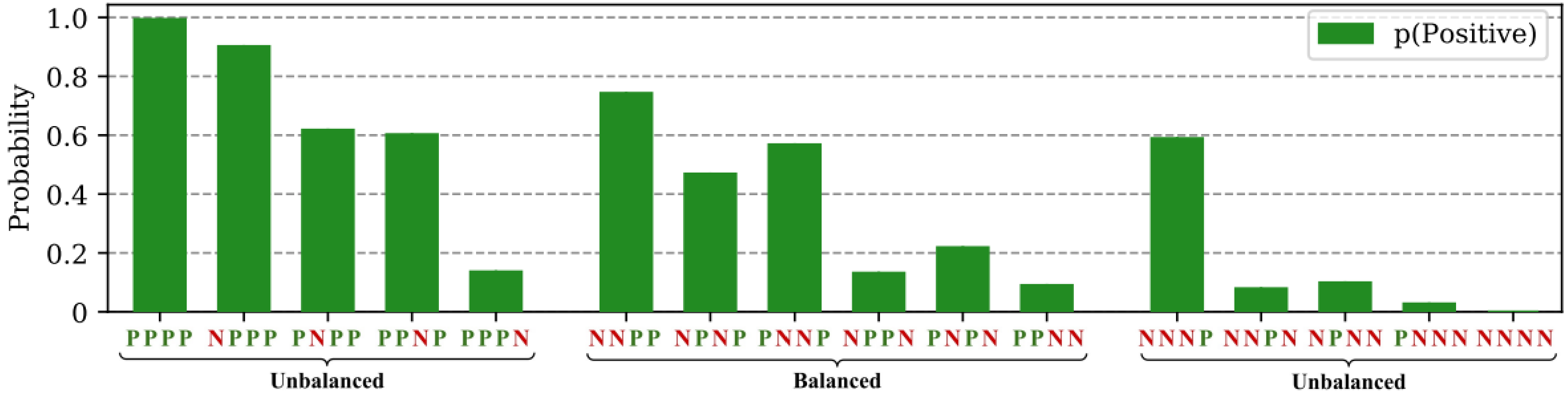


Dong et al, ‘23

Few-Shot Choices

Examples/structure affect performance:

1. Prompt **format** (affects everything)
2. **Choice** of examples
3. **Order** of examples (permutation)



Zhao et al '21

1. Prompt Formats

The choice of model affects the prompt format

Masked language model: “Cloze”-style prompt

- “`I love this movie, it is a [Z] movie:`”

Left-to-right language model: prefix prompt

- “`I love this movie. What is the sentiment of this review?`”

Note: eval datasets have pre-created prompts.

- LAMA (LAnguage Model Analysis): Cloze prompts

1. Prompt Formats: **Recent Models**

Modern instruction-tuned models have more complex instructions/formats

- **The good:** more natural way to tell the model what to do
- **The bad:** searching over formats/templates increasingly challenging

    - *Example: (White et al, '23):* `"From now on, I would like you to ask me questions to deploy a Python application to AWS. When you have enough information to deploy the application, create a Python script to automate the deployment."`

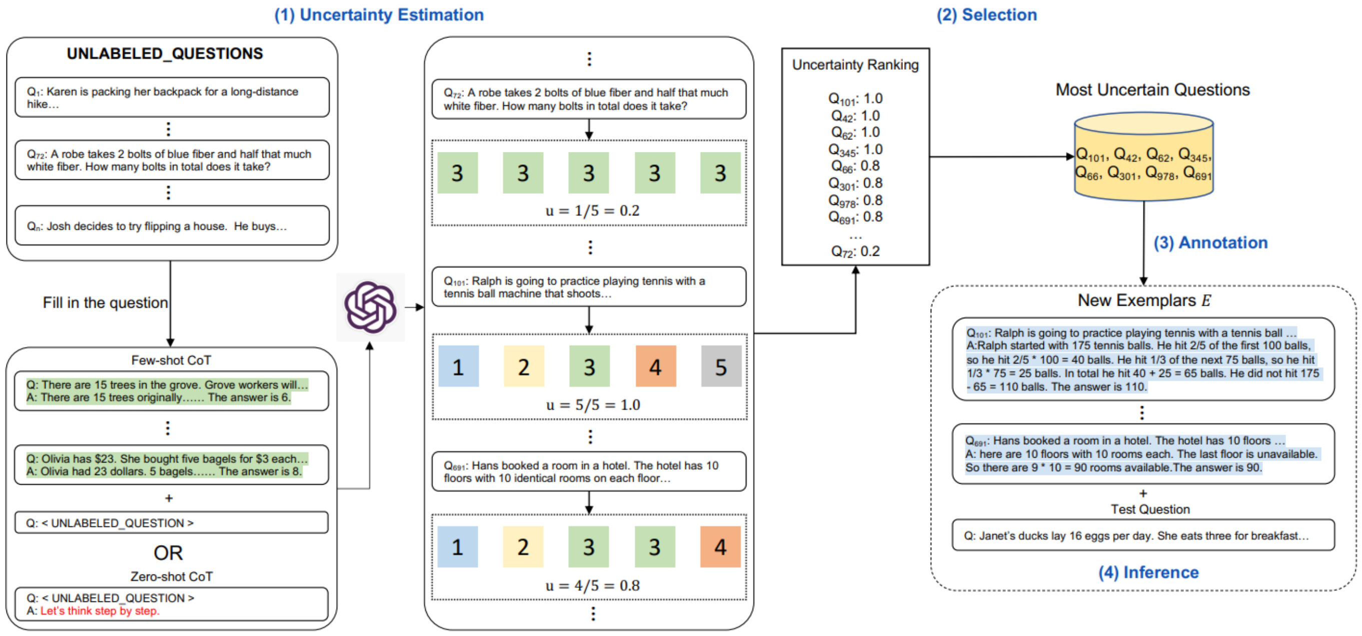
2. Choice of Examples

How to pick appropriate examples in few-shot?

- **Note:** only a “small’ number of examples can be shown, unlike in supervised learning.

Many options. Sampling:

- Liu et al, ‘21: kNN in embedding space (semantic similarity)
- Su et al, ‘22: Encourage diversity in embeddings
- Diao et al, ’23: “Active prompting”



Diao et al ‘23

3. Order of Examples

What order to show them to the model?

Fantastically Ordered Prompts and Where to Find Them: Overcoming Few-Shot Prompt Order Sensitivity

Yao Lu[†] **Max Bartolo**[†] **Alastair Moore**[‡] **Sebastian Riedel**[†] **Pontus Stenetorp**[†]
[†]University College London [‡]Mishcon de Reya LLP
{yao.lu,m.bartolo,s.riedel,p.stenetorp}@cs.ucl.ac.uk
alastair.moore@mishcon.com

- **Findings:**
  - Model size doesn’t guarantee low-variance
  - Adding more examples doesn’t reduce variance
  - Good prompts don’t transfer from one model to another ☹
  - Good orders don’t transfer

Break & Questions

Outline

- **Intro to Prompting**
  - Terminology: zero-shot, few-shot, in-context, etc, prompt characteristics: format, examples, orders
- **Hard and Soft Prompting**
  - Searching for good prompts, techniques for continuous/soft prompts
- **Prompt Ensembling and Other Methods**
  - Combinations, majority vote, chain-of-thought introduction, weighted ensembling

Hard Prompting

Also called **zero-shot**.

- Note: terminology conflict with another area called zero-shot learning

"Hard prompt discovery is a specialized alchemy, with many good prompts being discovered by trial and error, or sheer intuition

(Wen et al '23)

- Note: not just for language models!

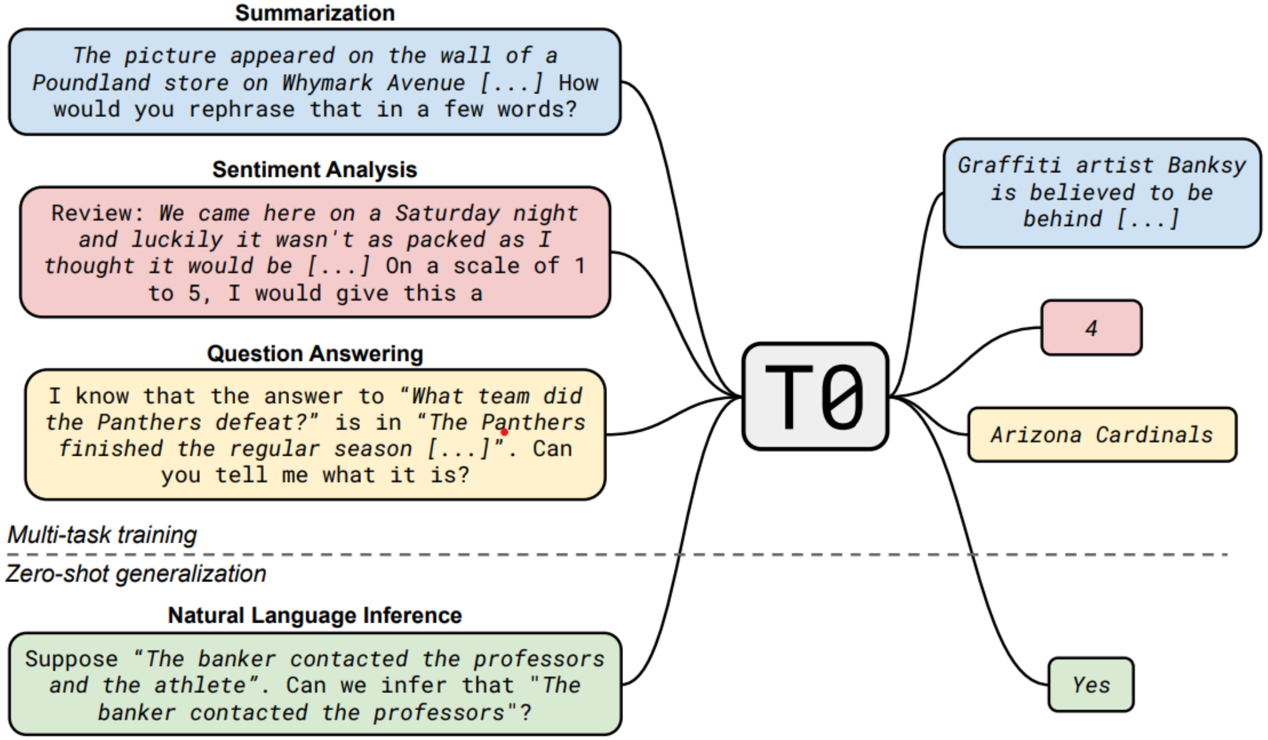
Zero-shot Generalization

Most exciting aspect of zero-shot: don't need to have been explicitly trained or fine-tuned.

- **Example: Multitask Prompted Training Enables Zero-Shot Task Generalization**

Recipe

- Pretrain
- Fine-tune
  - Multitask

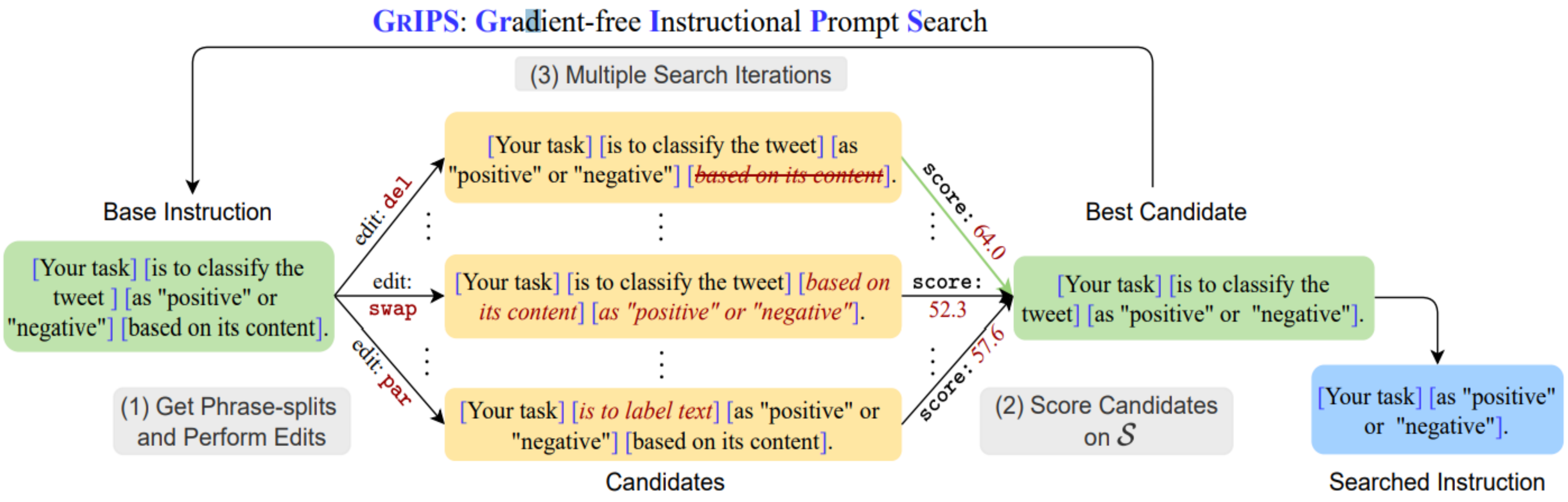


Sanh et al '22

Hard Prompting: **Discrete Optimization**

Sometimes, can avoid gradients

- Random search
- Greedy



Prasad et al '23

Soft Prompting

Also called **continuous prompting**

Basic idea: insert some (non-language) parameters into prompt

- Train these parameters
- Do not directly correspond to words in prompt

Prefix-Tuning: Optimizing Continuous Prompts for Generation

Xiang Lisa Li
Stanford University
xlisali@stanford.edu

Percy Liang
Stanford University
pliang@cs.stanford.edu

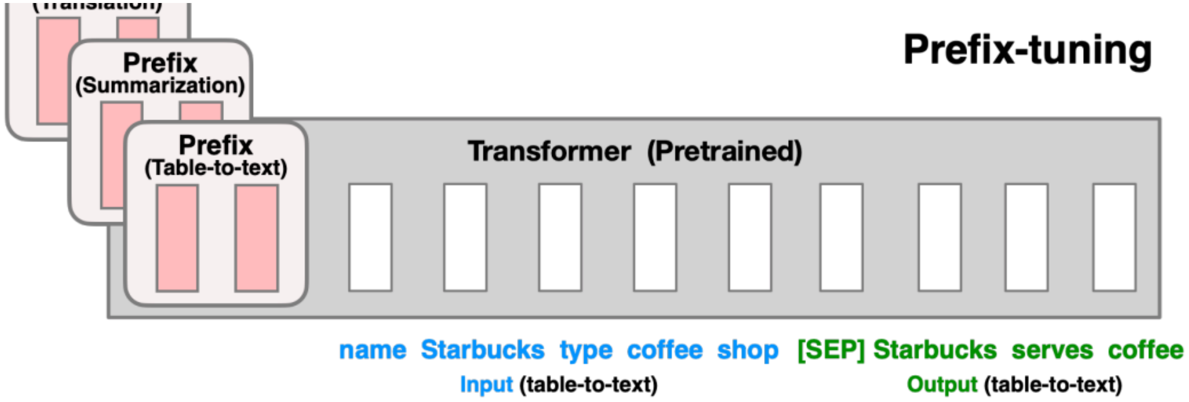
GPT Understands, Too

Xiao Liu [*1,2] **Yanan Zheng** [*1,2] **Zhengxiao Du** [1,2] **Ming Ding** [1,2] **Yujie Qian** [3] **Zhilin Yang** [4,2] **Jie Tang** [1,2]

Soft Prompting: **Prefix-Tuning**

Goal: create prefixes that *steer* models

- Prefixes are trainable parameters
- Train one for each goal task, only store these new parameters
- Enables cheap adaptation of frozen language model

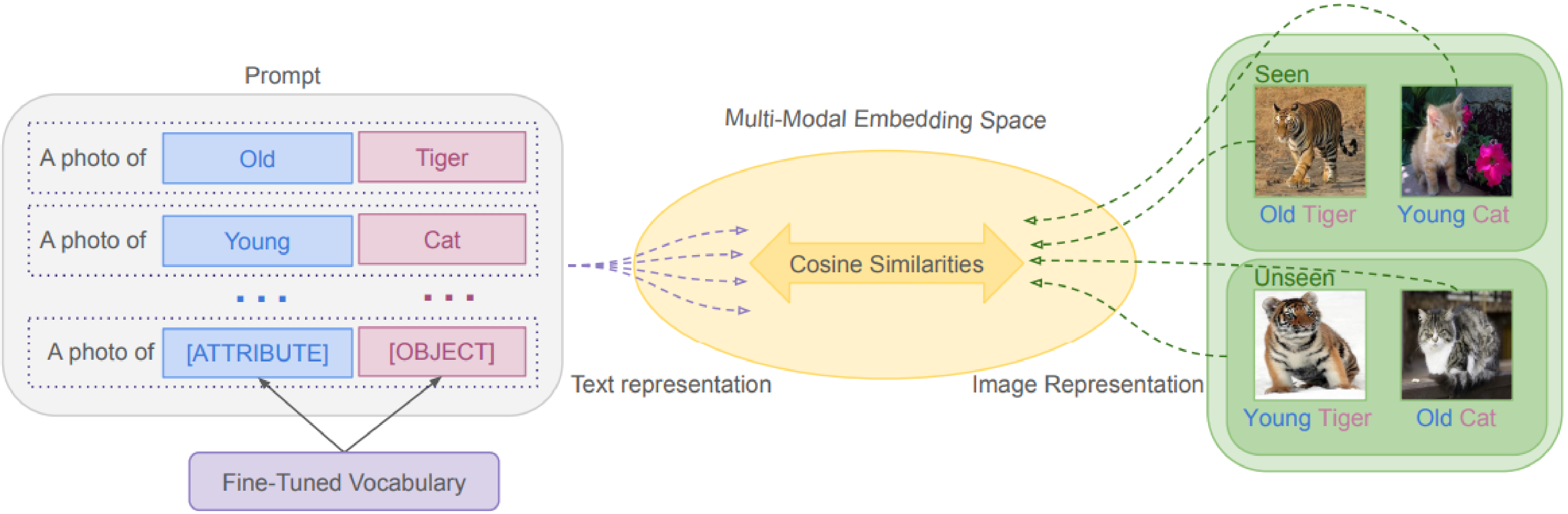


Li and Liang '21

Soft Prompting: **Composing**

What about multimodal models?

- Vision-language models like CLIP
- Not great composed concepts like *old tiger*
- Tune on [attribute] [object] pairs



Nayak et al '21

Break & Questions

Outline

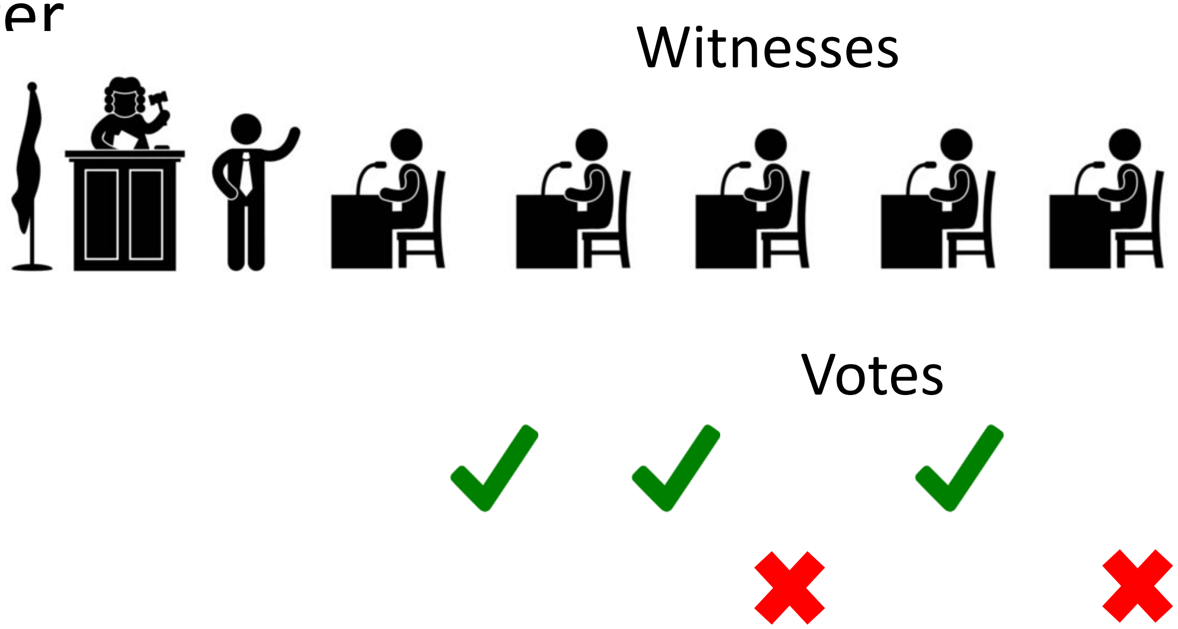
- **Intro to Prompting**
  - Terminology: zero-shot, few-shot, in-context, etc, prompt characteristics: format, examples, orders
- **Hard and Soft Prompting**
  - Searching for good prompts, techniques for continuous/soft prompts
- **Prompt Ensembling and Other Methods**
  - Combinations, majority vote, chain-of-thought introduction, weighted ensembling

Ensembling Prompts

One prompt can give you an answer… but might be wrong

- One simple approach: get multiple samples
- From?
  - Change temperature parameter
  - Vary your prompts

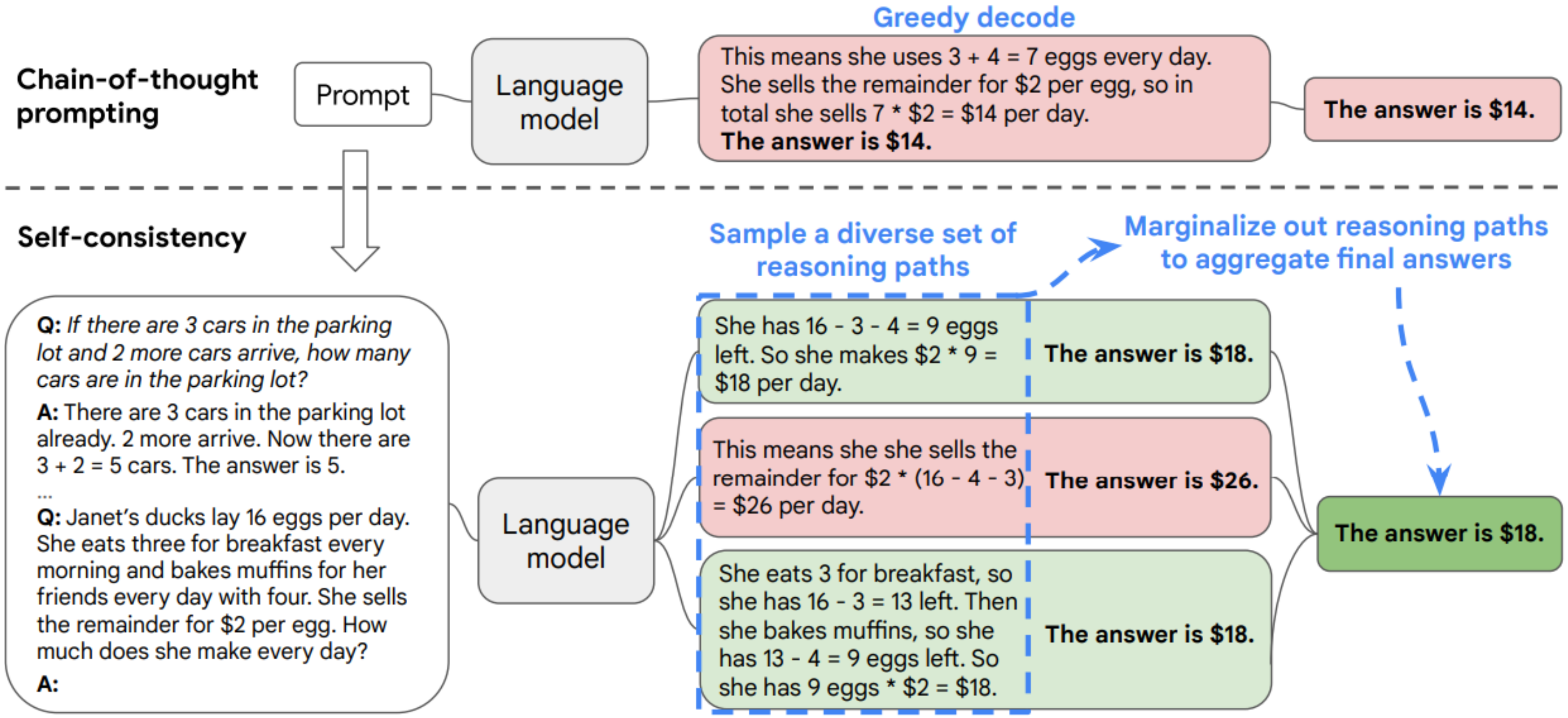
Then, run **majority vote**



Chain-of-Thought

A form of prompting that helps break down the problem (more in a week!)

- Produces more answers to run majority vote on

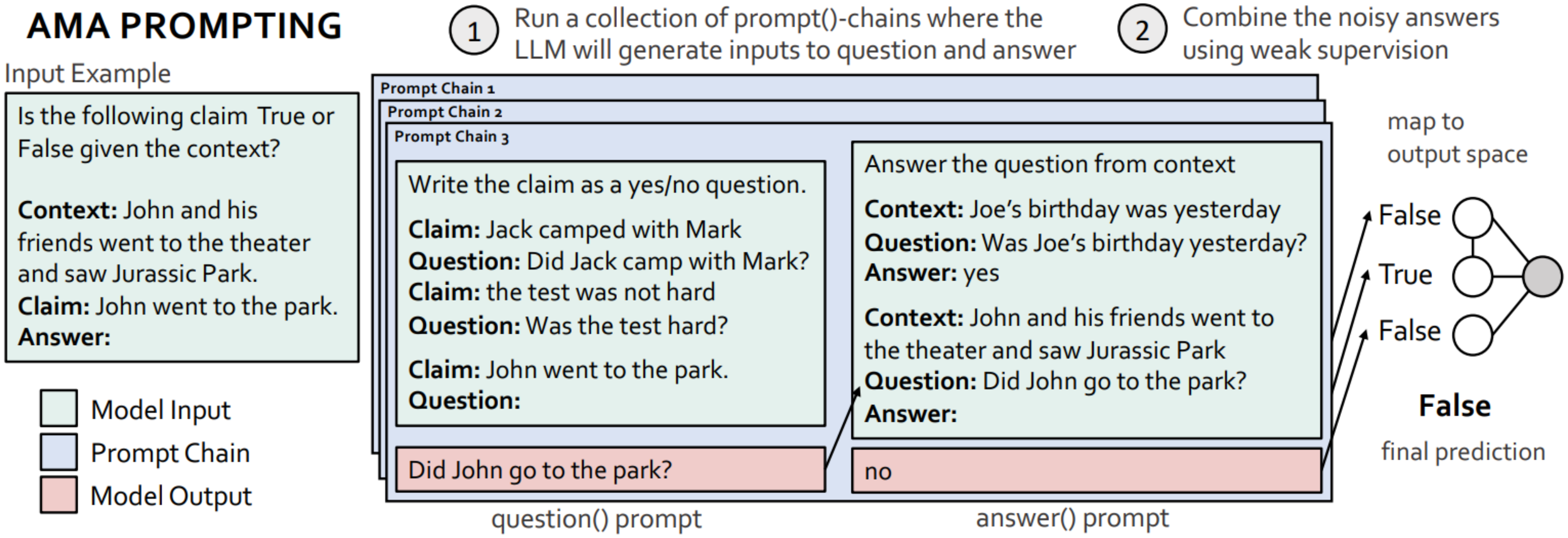


Wang et al '23

Ensembling Prompts: Weighted Version

Downside of majority vote… most responses might be wrong

- Should weight them by how accurate they are



Chen et al, '23

Thank You!