Fine-Tuning / Instruction Tuning

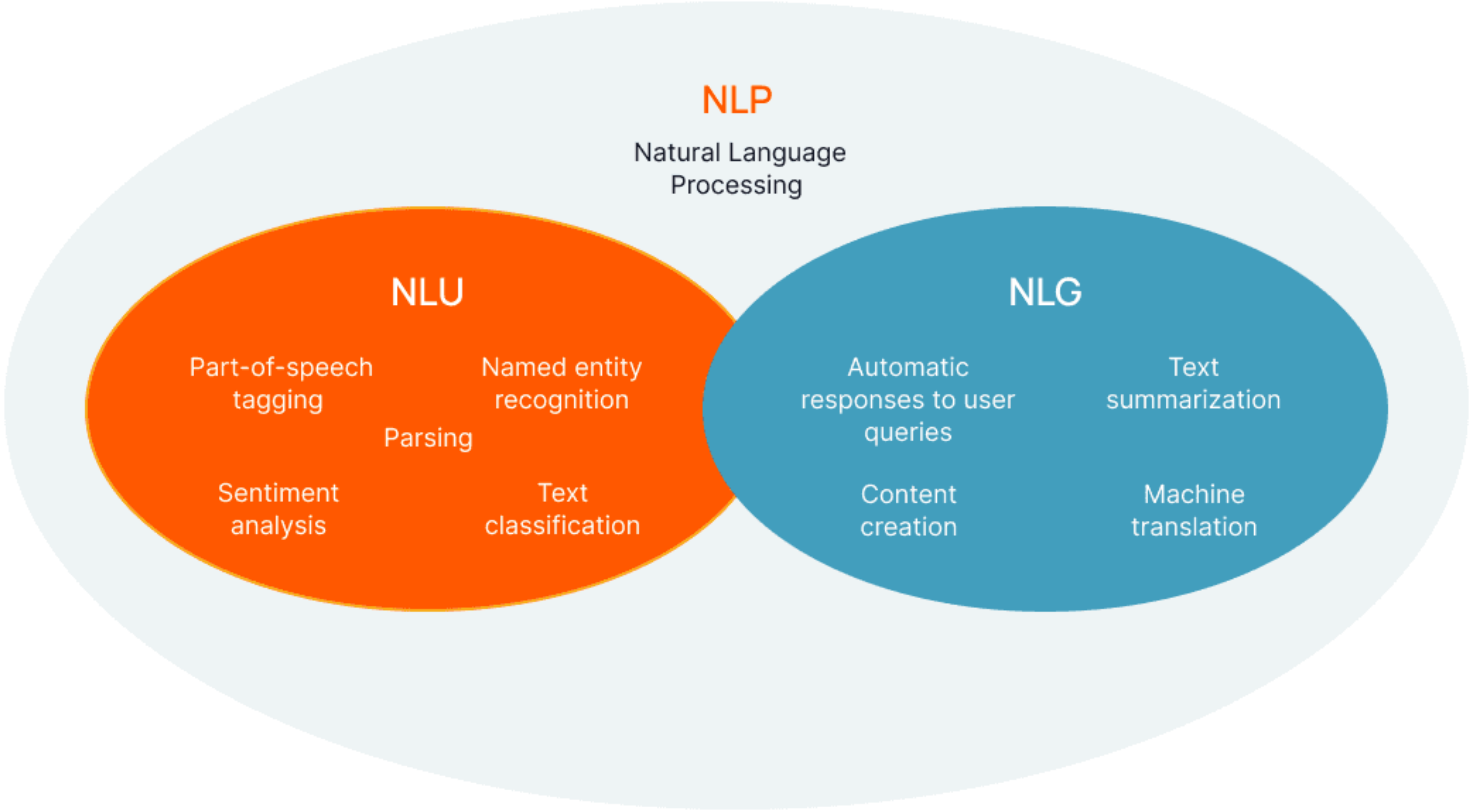
Haw-Shiuan Chang

Deadlines

- **https://people.cs.umass.edu/~hschang/cs685/schedule.html**

- **3/3: Quiz 2 due**
- **3/7**: Project proposals due
  - Please submit only one proposal for each group
  - If the score of the proposal is lower than the final report, we will use the final report score.
    - However, if you don't submit the proposal, we won't give you the feedback and provide you with some LLM credit support.
  - **In your proposal, please estimate the cost of API credit you need and which LLM and service provider you plan to use.**
    - I know it is hard, but please try.
    - We only have $500 for the whole class. Try not to have a project that needs hundreds of dollars (unless you are willing to pay by yourself).
    - You might get more money if your proposal looks better and more feasible
- **3/14**: HW 1 due
- **3/17:** Quiz 3

NLU vs NLG



https://www.kdnuggets.com/natural-language-processing-bridging-human-communication-with-ai

Architecture Comparison

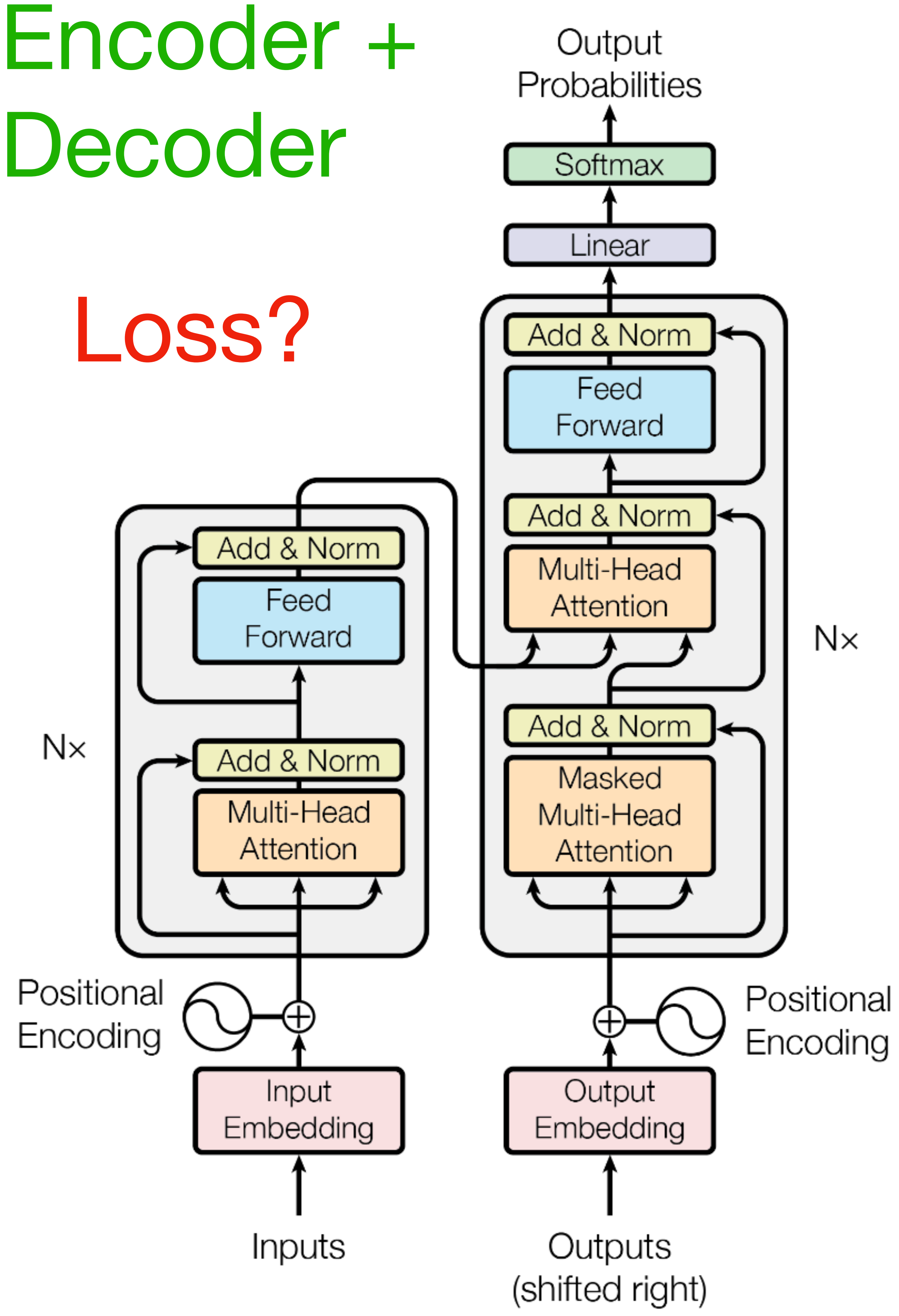
Encoder
Decoder
Encoder + Decoder

XX BERT
GPT
T5

ChatGPT
Claude
Gamini
Llama
Deepseek

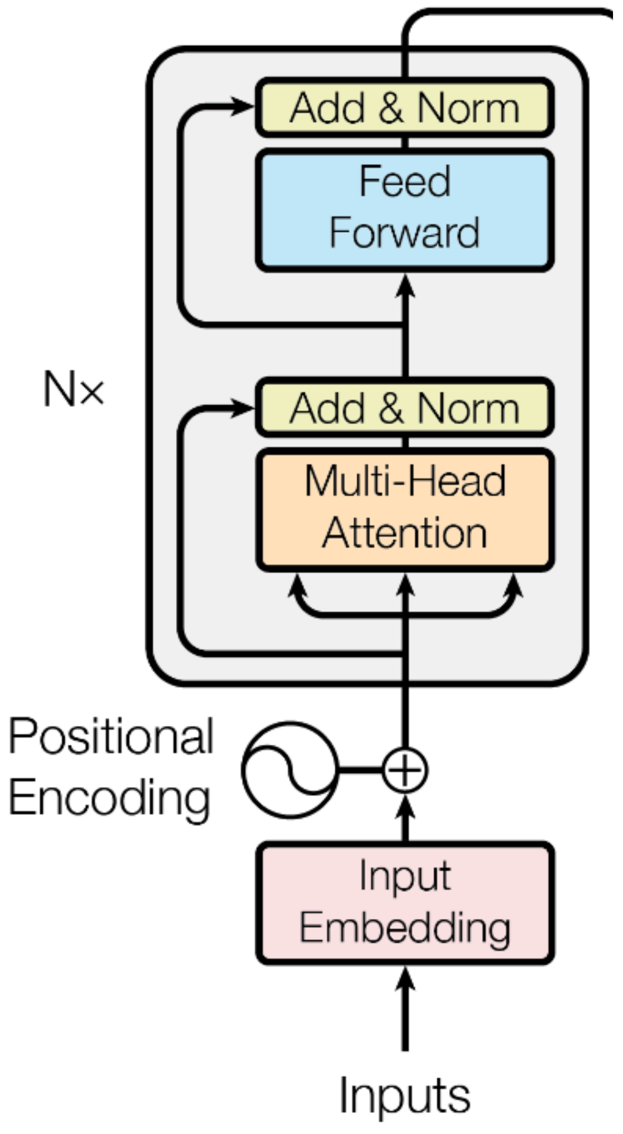
Encoder + Decoder

Loss?



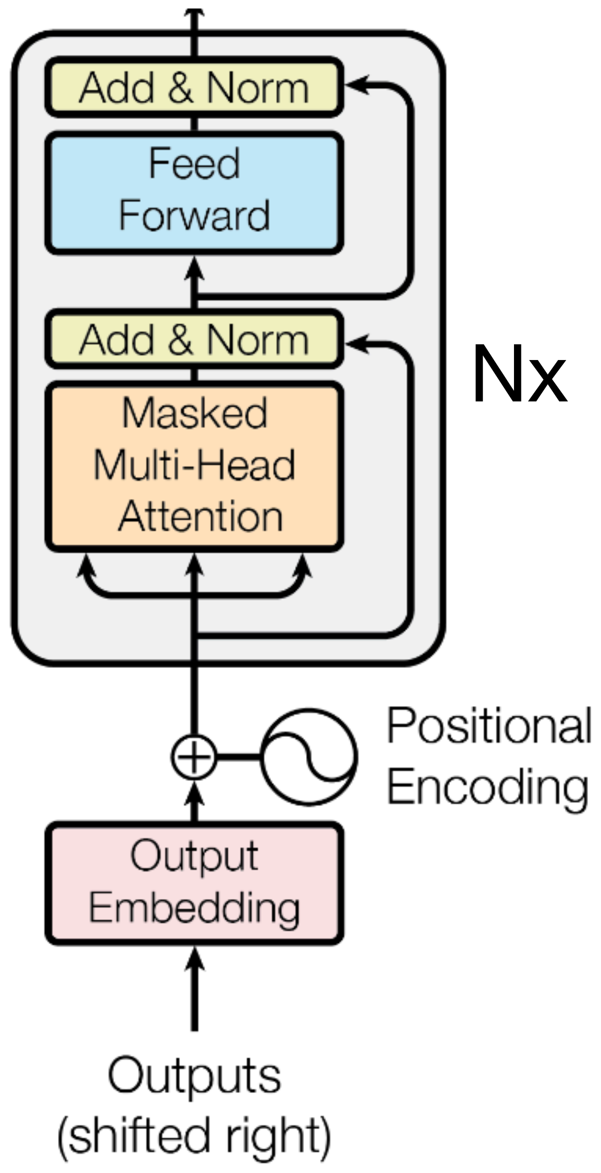
Encoder

Loss?



Decoder

Loss: Predicting next token



Last Year Notes

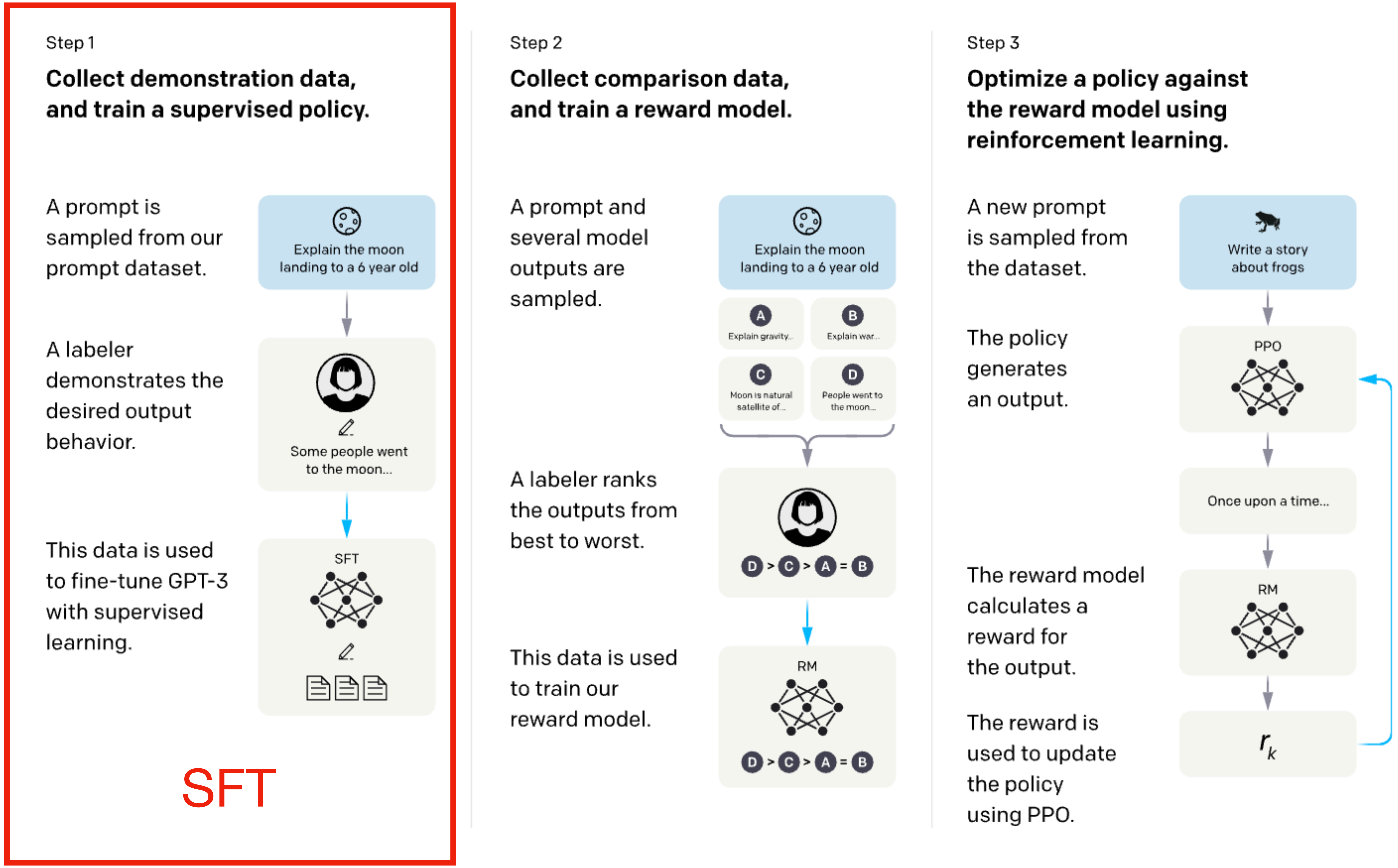
LLM Development

- Architectures
  - MLP
  - RNN
  - Transformer

- Training Stages
  - Pretraining → Usually more expensive
  - **Supervised Fine-tuning (SFT)**
  - Alignment
    - Learning from Human Feedback (LHF)
    - Reasoning

Post-training stage

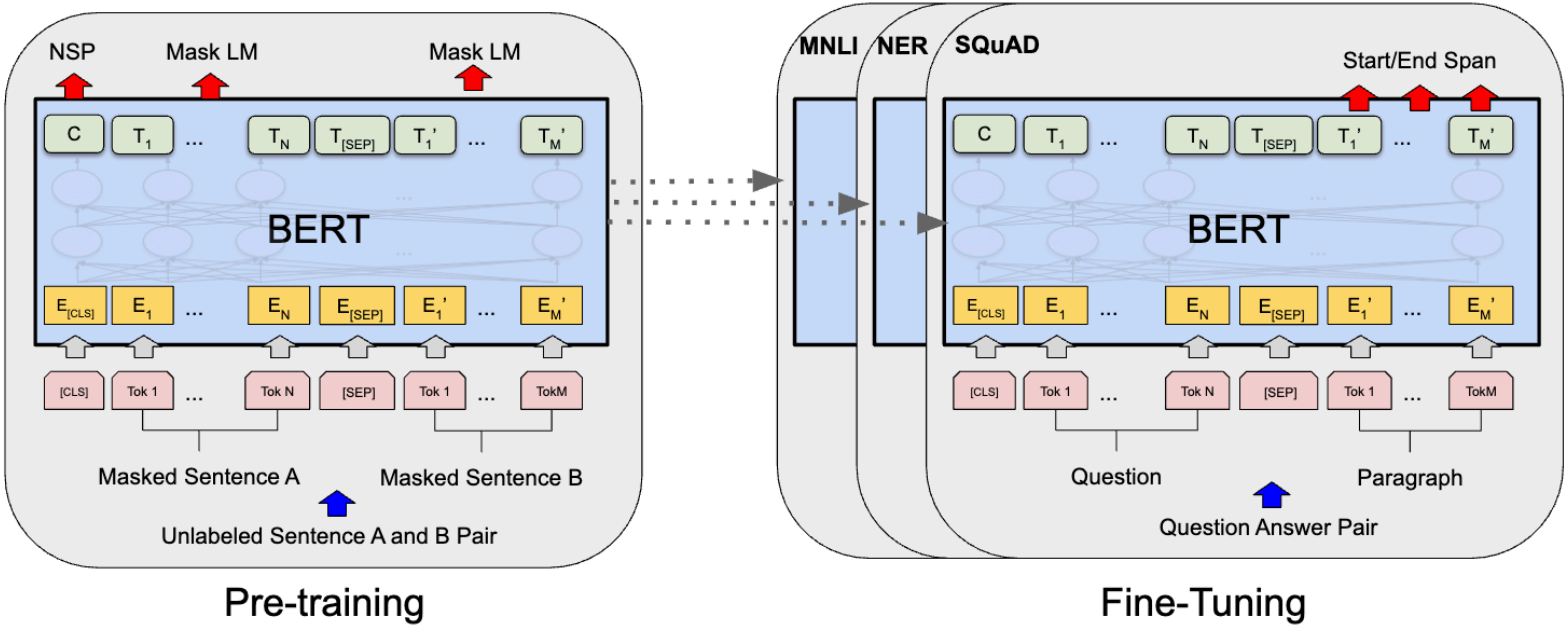
Supervised Fine-Tuning (SFT)



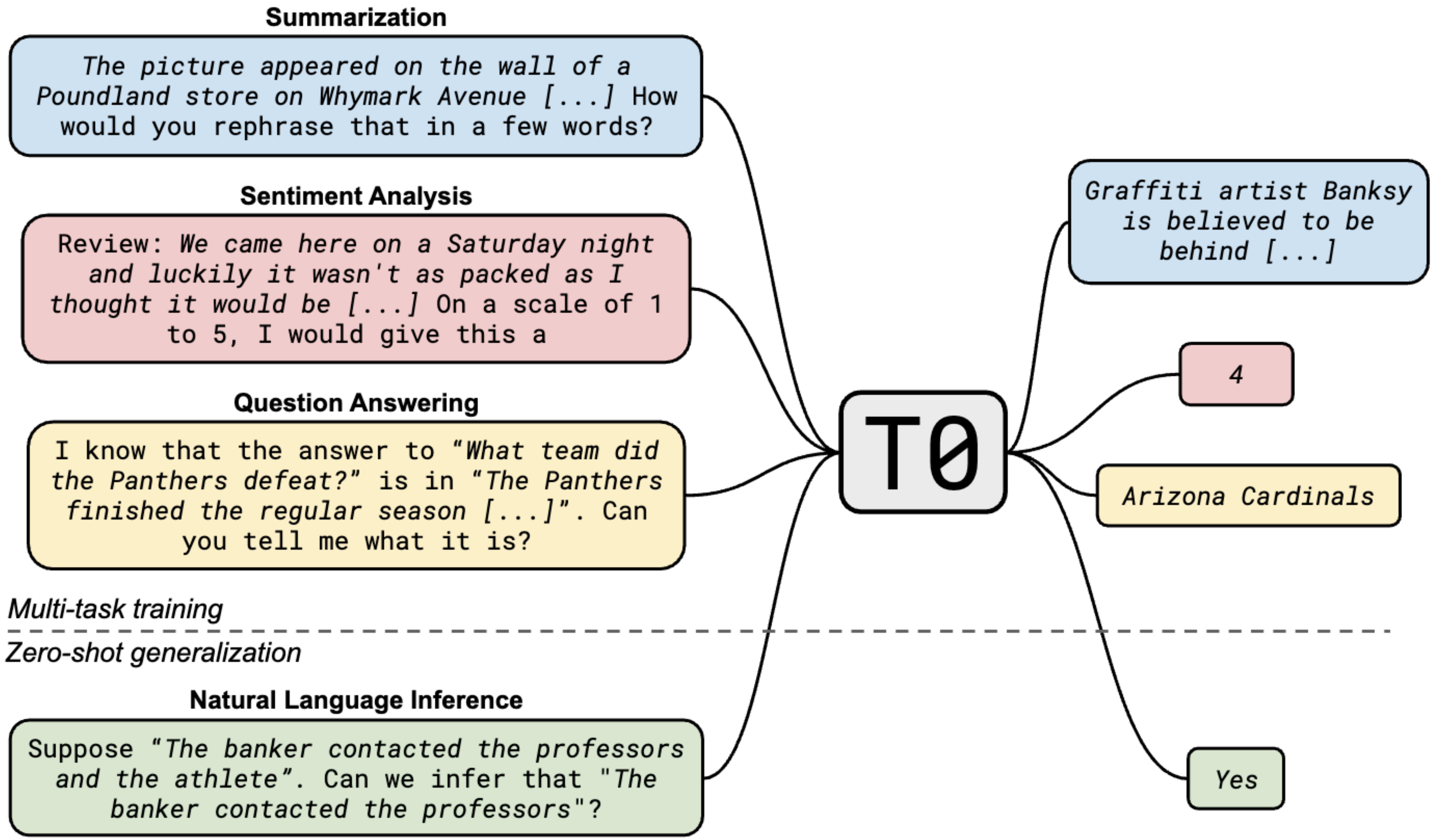
What Fine-Tuning Data Should we Collect/Use?

The answer is complicated

Old Fine-Tuning

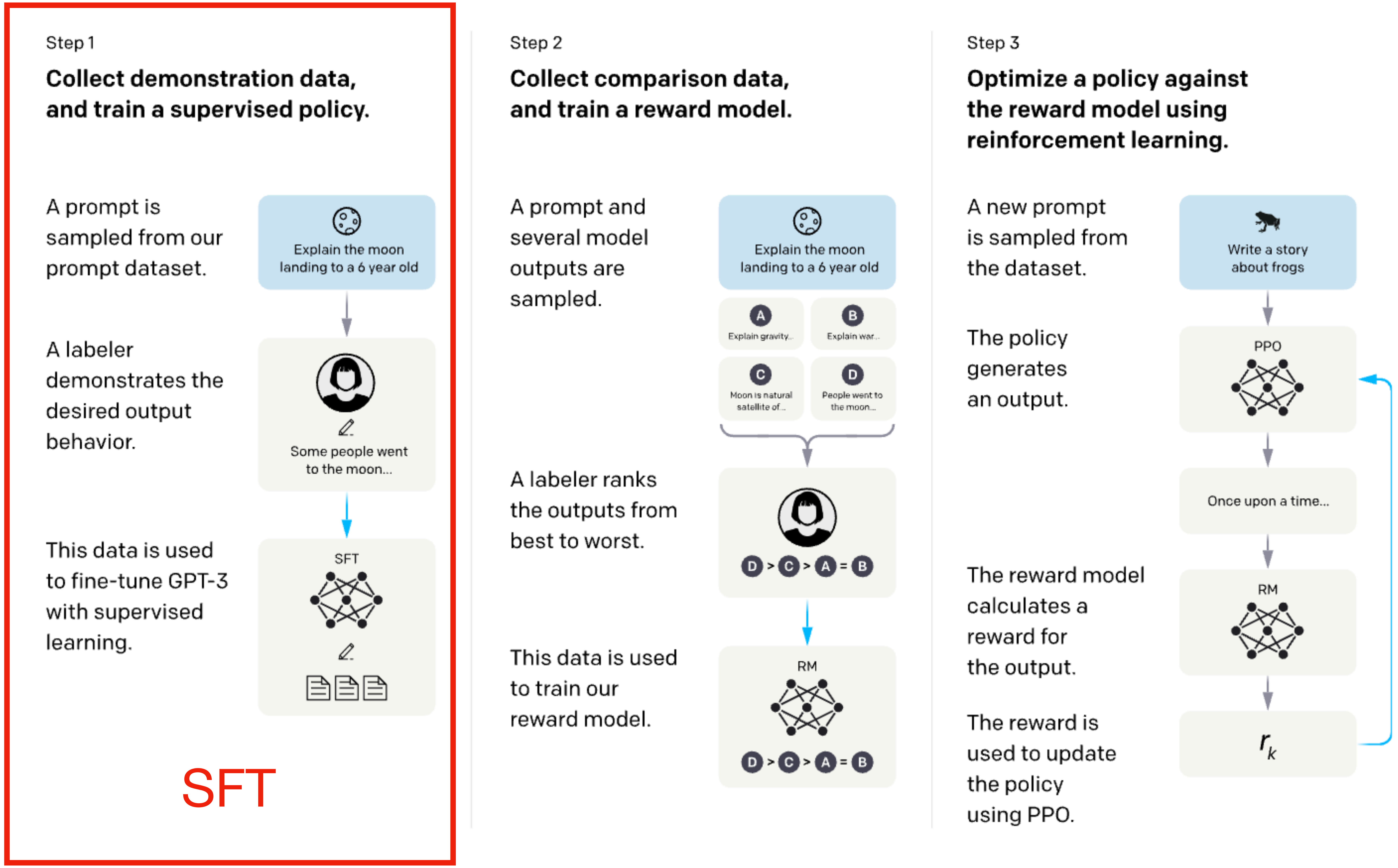


Instruction Tuning



MULTITASK PROMPTED TRAINING ENABLES ZERO-SHOT TASK GENERALIZATION (https://arxiv.org/pdf/2110.08207)

Supervised Fine-Tuning (SFT)



Old Fine-Tuning

- Before ChatGPT/LLM
  - Models like BERT
- Tasks
  - Usually single Natural Language Understanding (NLU) task
- Input
  - No instruction
- Output
  - Often human labeled

Instruction Tuning

- Before ChatGPT/LLM
  - Models like T0
- Tasks
  - Many tasks, NLU + NLG
- Input
  - Instructions from hundreds of tasks
- Output
  - Often human labeled

SFT

- After ChatGPT/LLM
  - Models like GPT 3.5
- Tasks
  - Many tasks, usually Natural Language Generation (NLG) tasks
- Input
  - Mostly free-form Instructions
- Output
  - Sometimes extracted

Old Fine-Tuning

- Specialization
  - One FT Model for one task
- Cost
  - Cheap
- Boundary to Pretraining
  - Clear

Instruction Tuning

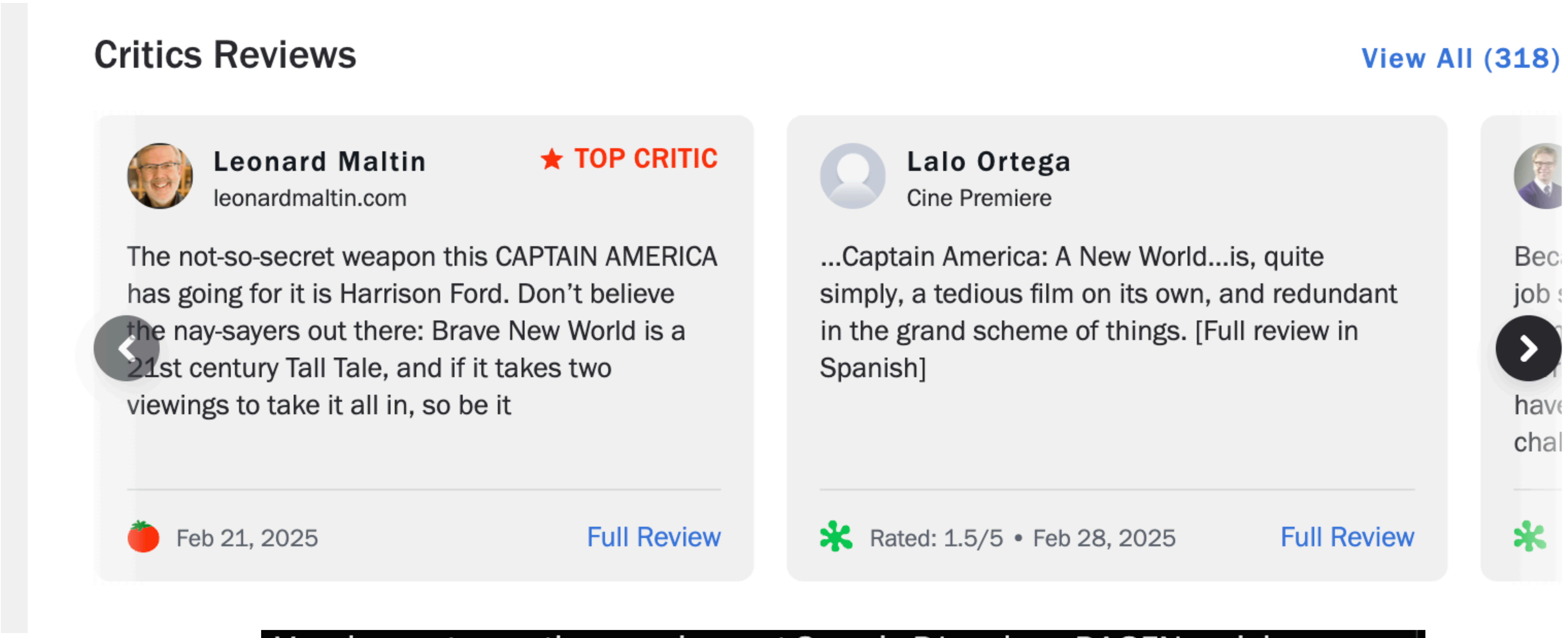
- Specialization
  - One FT Model for all tasks
- Cost
  - Expensive (various tasks)
- Boundary to Pretraining
  - Clear

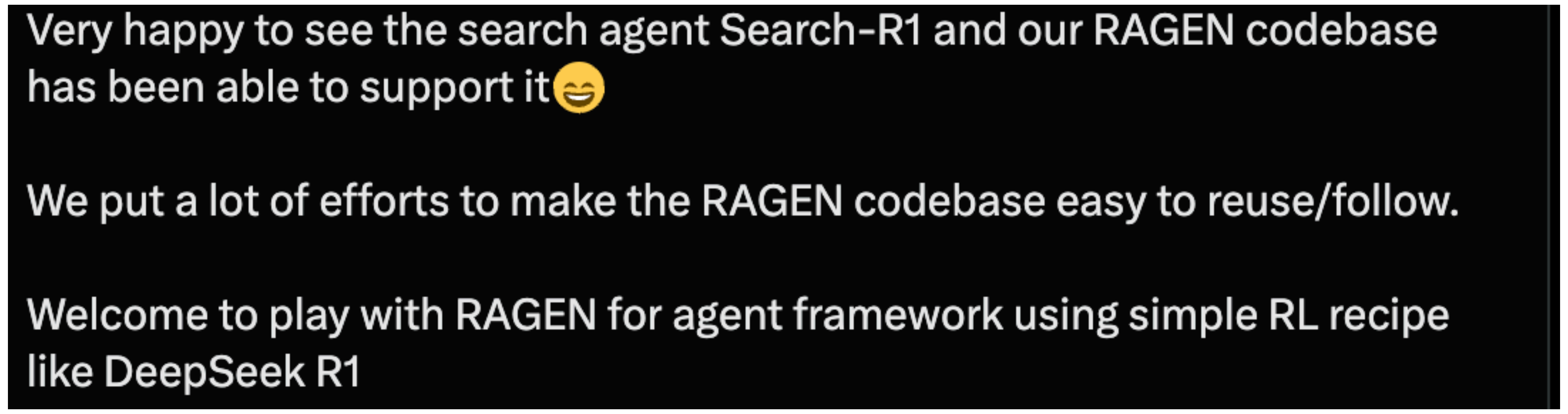
SFT

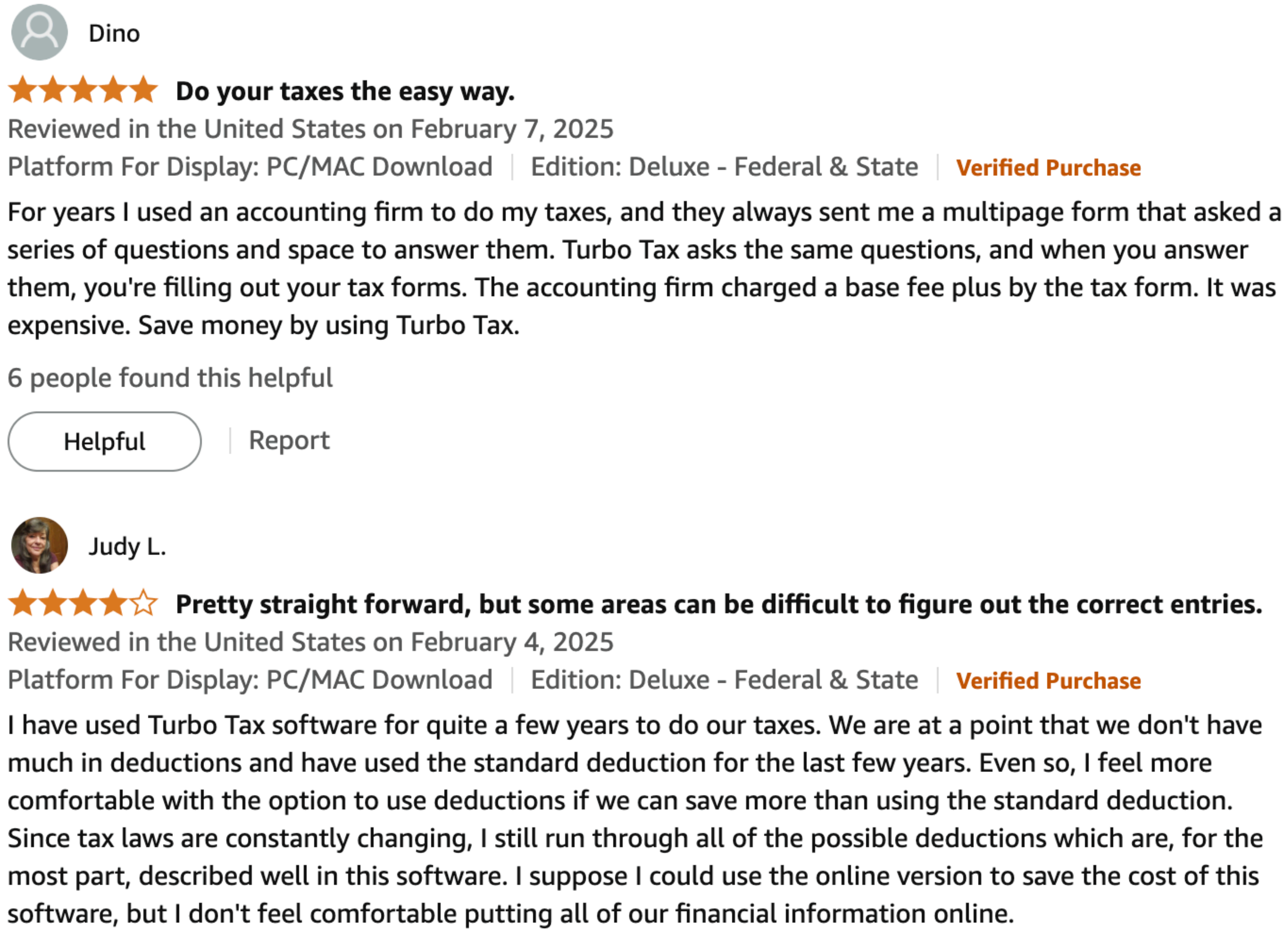
- Specialization
  - One FT Model for all tasks
- Cost
  - Expensive (LLM)
- Boundary to Pretraining
  - Blurry

Difference between pretraining and fine-tuning is not loss different. It is data difference and hyperparameter difference.

The Blurry Boundary between Pretraining and SFT





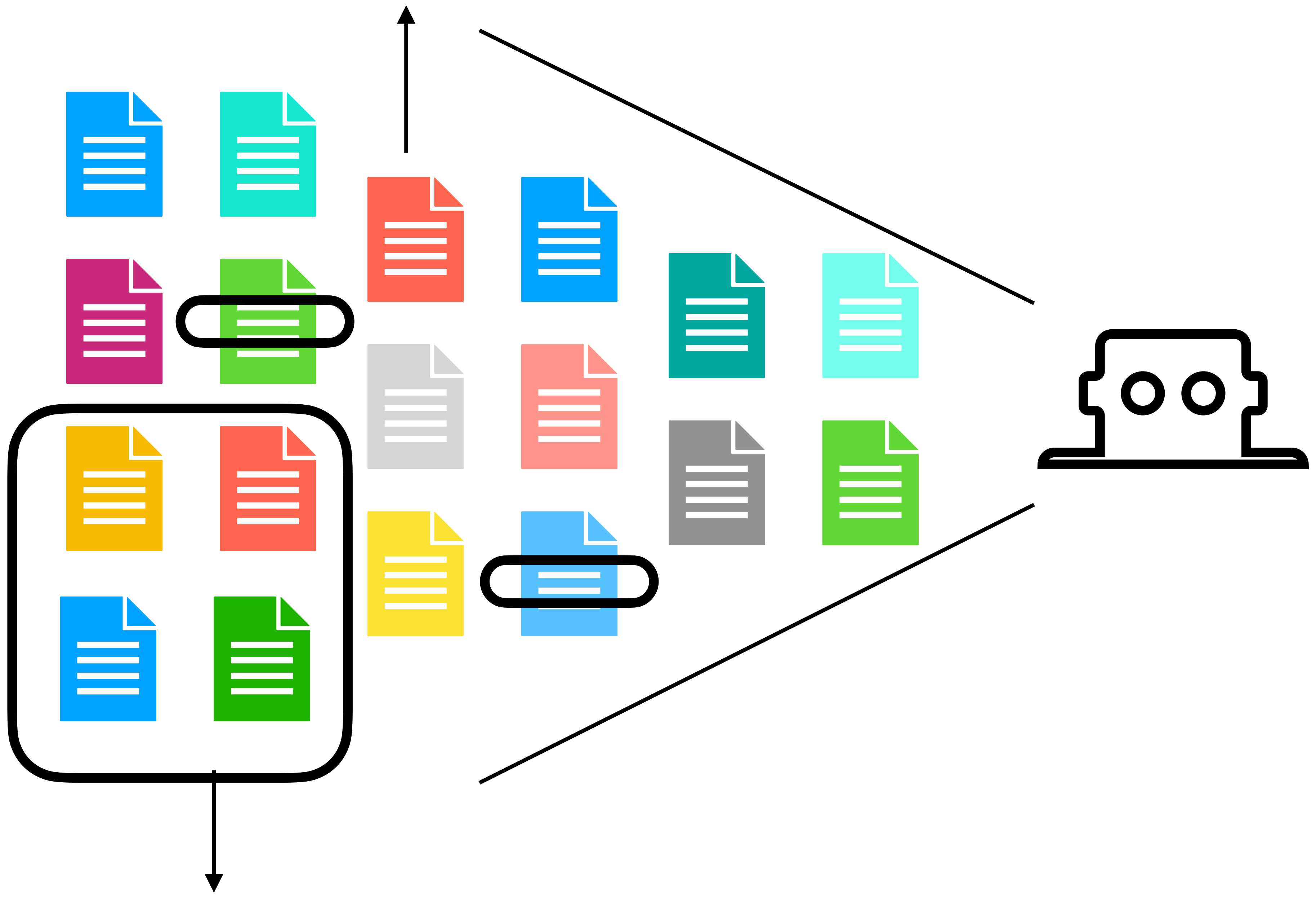


- "Transfer" to Sentiment Analysis
- Controlled Review Generation

Perspectives != Facts

I will present my conclusions derived from existing papers and my experience.
Some of them have not been universally accepted by the NLP community yet

LLM Development



- Architectures
  - MLP
  - RNN
  - Transformer
- Training Stages
  - Pretraining
  - **Supervised Fine-tuning (SFT)**
  - Alignment
    - Learning from Human Feedback (LHF)
    - Reasoning

Post-training stage
(Filtering process)

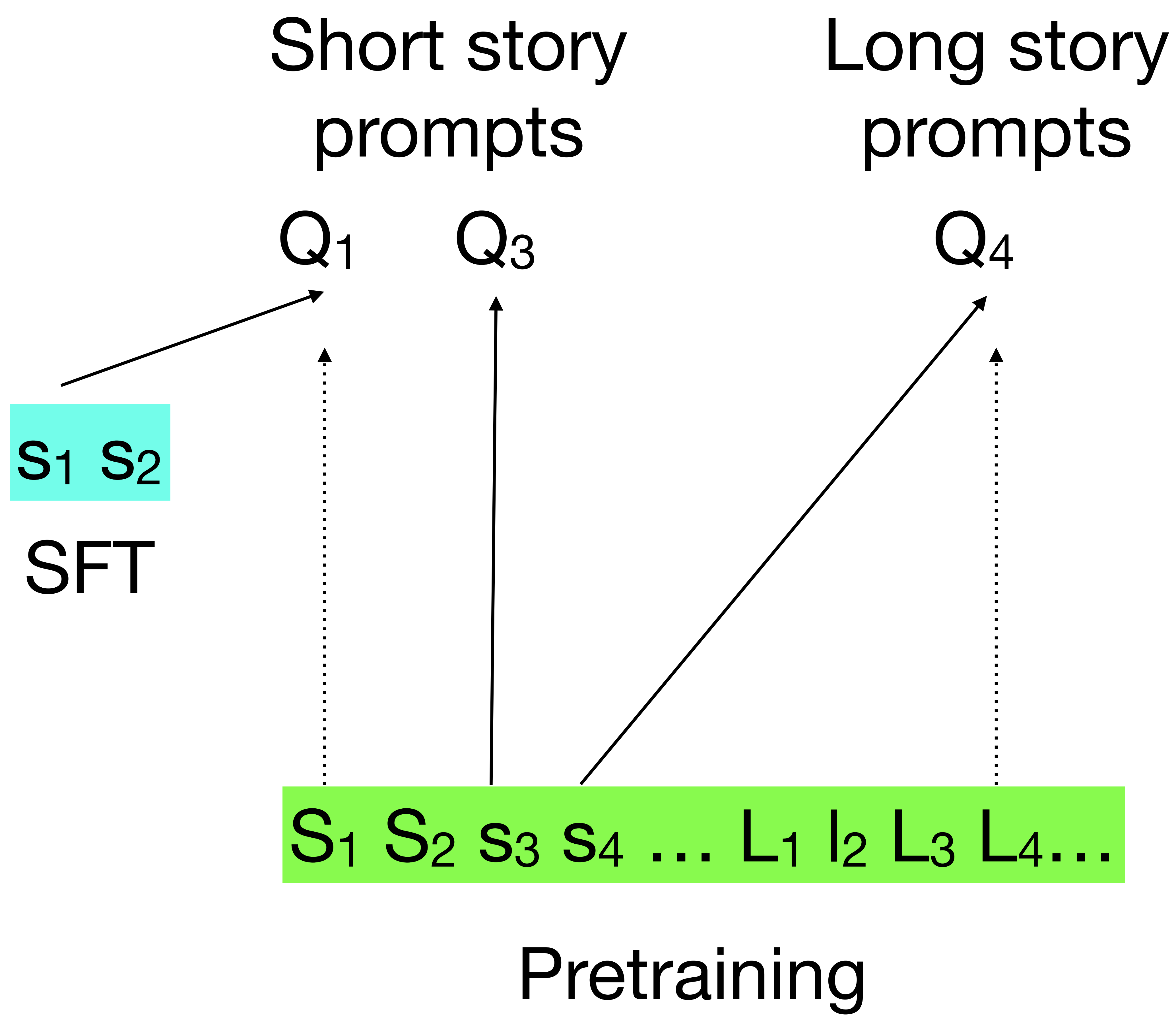
Question

- Assuming you are in an SFT team at a large company. You recently collected 1k high-quality (constraints, short stories) to improve an LLM.
- Given the context is “Please output a short story with the following constraints: {constraints}.”, you fine-tune the LLM to output the collected short story.
- During the testing time, you compare the LLM’s response before and after your fine-tuning given the prompt “Please output a long story.”
- After fine-tuning, will the story length be reduced a lot?

Fine-tuning LLM for a Target Task

- Target task itself
  - Response from the fine-tuning data
    - Could also overwrite the good responses in the pretraining data
  - Response learnt from pretraining
    - Inducing output that is similar to the output of the fine-tuning data

- Other tasks
  - Could make the output closer to the output of the target task
    - e.g., after training on a short story, the story length would decrease when you want LLM to give you a long story



LLM Paradigm Shift

- Before LLM, instruction tuning should also use as many tasks/data in the fine-tuning stage as possible
  - Example: Flan-T5:
    - Smaller encoder-decoder model, less pretraining -> many fine-tuning tasks
- After LLM, SFT should only use a few high-quality fine-tuning data
  - Example ChatGPT:
    - Larger decoder-only model, more pretraining -> fewer fine-tuning tasks

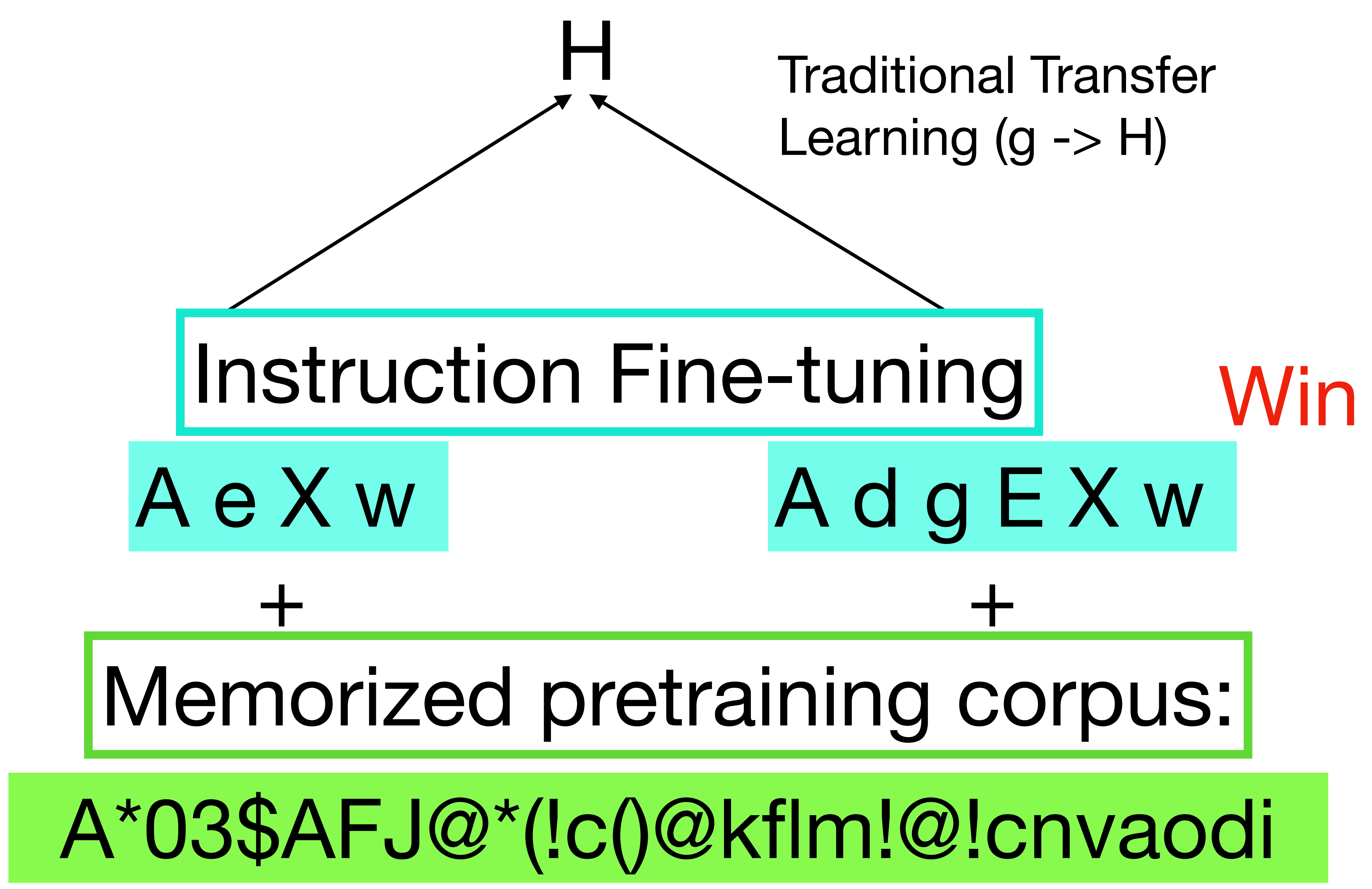
LIMA: Less Is More for Alignment (https://proceedings.neurips.cc/paper_files/paper/2023/file/ac662d74829e4407ce1d126477f4a03a-Paper-Conference.pdf)

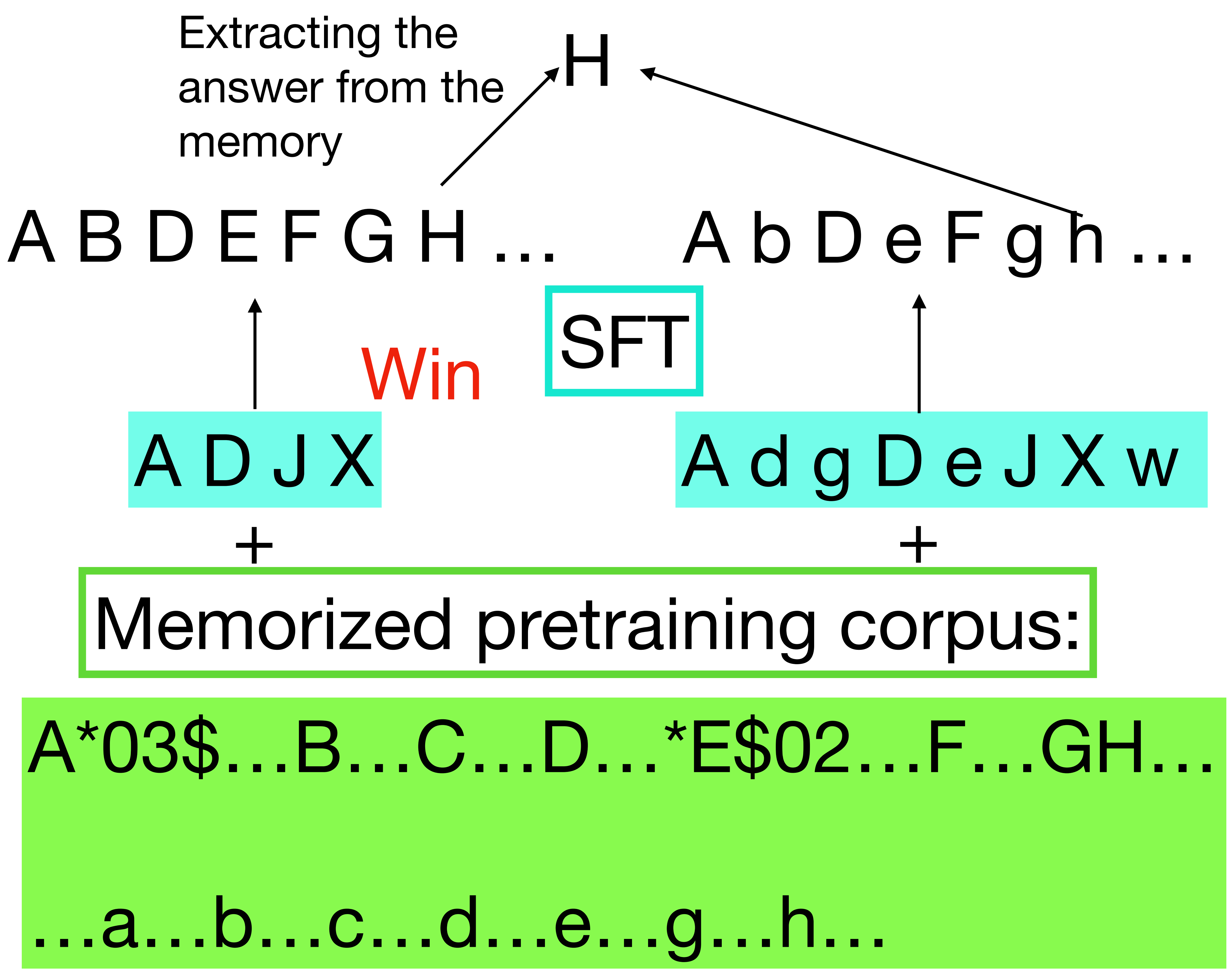
Question

- Remember that the loss function for SFT/pretraining/instruction-tuning is the same.
  - A larger pretraining dataset is better.
  - A larger instruction-tuning dataset is better.
  - Then, why could a larger SFT dataset degrade the performance?

Why Could Fewer Data be Better?

- First task -> A: high-quality data, a: low-quality data





Recent studies show that such transfer learning does not actually work generally. See this paper:

Do Models Really Learn to Follow Instructions? An Empirical Study of Instruction Tuning (https://arxiv.org/pdf/2305.11383)