

# From Text to Thought

## How Modern AI Models Learn to Reason

**Pittawat Taveekitworachai**

*Research Scientist, SCB 10X R&D*

3 February 2025 @ SIT, KMUTT

# About Me – Pittawat Taveekitworachai (Pete)

## Research Scientist @ SCB 10X R&D (Typhoon)

- Typhoon T1 – Open Reasoning Models

**Research Interests:** Reasoning in LMs, Prompt Engineering, Inference-time Scaling, and LM Agents

- Over 25 publications about NLP (h-index: 7, citations: 107)

## Me & KMUTT

- CS@SIT #19, ex-GDSC Lead for KMUTT, 2B-KMUTT #13-14, ComCamp#29, WipCamp #9, JPC XV technical lead



Open source AI technology for  
Thai language, knowledge, and  
culture

# What Is Typhoon?

Typhoon is an advanced research initiative focused on developing open-source large language technologies for the Thai language. We provide models, datasets, tools, and research to advance Thai language AI and multimodal capabilities



Efficient Speed & Cost



Improved Thai Knowledge  
and Instruction-Following  
Performance



Open Source

*Open access to resources fosters collaboration and drives AI innovation*

# PARTNERSHIPS



together.ai



# Agenda (1/2)

## Part I: Introduction

- The Current Landscape of Reasoning Models

## Part II: “Thought”

- What is a “Thought”?
- Human Thinking System
- Chain-of-Thought Prompting
- Beyond CoT Prompting
- Reasoning Model vs \*-of-Thought Prompting
- Can’t We Train CoT Back to LLMs?

# Agenda (2/2)

## Part III: How to Train a Reasoning Model?

- Machine Learning Paradigms
- How to Train a Large Language Model?
- How to Train a Reasoning Model?
- Do We Really Need RL?

## Part IV: Typhoon T1-A Case Study on Developing A Reasoning Model

- Lessons Learned

## Part V: What's Next?

- Future of Reasoning Models

# Introduction

Part I

# How Would You Solve This Problem?

*Think it out loud!*

*If a doctor gives you three pills, telling you to take one every half hour, how many minutes will pass from taking the first pill to the last pill?*

# Okay, Let's See How an LLM Think?

To determine how many minutes will pass from taking the first pill to the last pill, let's break it down step by step:

1. **First Pill:** You take the first pill at time **0 minutes**.
2. **Second Pill:** You take the second pill after **30 minutes** (half an hour). So, the second pill is taken at **30 minutes**.
3. **Third Pill:** You take the third pill after another **30 minutes** (half an hour). So, the third pill is taken at **60 minutes**.
4. **Total Time:** The time from the first pill (0 minutes) to the last pill (60 minutes) is **60 minutes**.

Thus, **60 minutes** will pass from taking the first pill to the last pill.

*Is this thought?*

# How About a Reasoning Model?

Okay, let's see. The doctor gives me three pills and tells me to take one every half hour. I need to figure out how many minutes will pass from taking the first pill to the last one. Hmm, let's break this down.

First, if I have three pills, that means I have to take them at different times, right? The first pill is taken immediately, I guess. Then the second one half an hour later, and the third one another half hour after that. Wait, but how many intervals are there between the pills?

Let me think. If I take the first pill at time zero, the second pill would be at 30 minutes, and the third pill at 60 minutes. So from the first to the third pill, that's 60 minutes. But wait, maybe I need to check that again.

Let me visualize it. Let's say I take pill 1 at 12:00 PM. Then, since I need to take one every half hour, the next one would be at 12:30 PM, which is pill 2. Then the third pill would be at 1:00 PM. So from 12:00 PM to 1:00 PM is one hour, which is 60 minutes. So the total time from first to last is 60 minutes. That seems straightforward.

But sometimes these problems can be tricky. Let me consider another approach. If I have three pills, the number of intervals between taking them is two. Because between pill 1 and pill 2 is one interval, and pill 2 and pill 3 is another. Each interval is 30 minutes. So 2 intervals times 30 minutes each is 60 minutes. That matches the previous answer.

Wait, but I've seen similar problems where people might think it's 90 minutes. Maybe they thought three pills, each 30 minutes apart, so 3 times 30. But that would be 90 minutes. But that's incorrect because the first pill is taken at the start, so the time starts counting from there. So the first interval is between the first and second pill, then the second interval is between the second and third. So two intervals. Hence, 30 times 2 is 60.

Let me test with a smaller number. Suppose I have two pills, take one every half hour. Then, take the first at time zero, the second at 30 minutes. So total time is 30 minutes. Which is one interval. So number of pills minus one gives the number of intervals. Therefore, three pills would have two intervals. So two times 30 is 60.

Yes, that makes sense. So the answer should be 60 minutes.

But wait, just to confirm. Let's take another example. If I had four pills, taking one every half hour. Then the intervals would be three, so 3 times 30 is 90 minutes. So from first to last would be 90 minutes. So the formula is (number of pills - 1) multiplied by interval time.

Therefore, applying that here: 3 pills, minus 1 is 2, times 30 minutes equals 60 minutes. Correct.

So the confusion might come if someone counts the number of pills and multiplies by the interval, but that's not the case here because the first pill starts the timer, and each subsequent pill is after an interval. So the total time is intervals between the pills, not the number of pills.

So yeah, the answer is 60 minutes. So when the doctor gives you three pills and tells you to take one every half hour, the duration from the first to the last pill is 60 minutes.

**Final Answer**

The total time from the first to the last pill is \boxed{60} minutes.

*The doctor gives you three pills and instructs you to take one every half hour. To determine the total time from taking the first pill to the last pill, we need to consider the intervals between each pill.*

1. First pill: Taken at time 0 minutes.
2. Second pill: Taken 30 minutes after the first pill.
3. Third pill: Taken 30 minutes after the second pill.

*The intervals between the pills are calculated as follows:*

- Between the first and second pill: 30 minutes.
- Between the second and third pill: 30 minutes.

*There are 2 intervals between 3 pills. Each interval is 30 minutes, so the total time is:  

$$[2 \times 30 = 60 \text{ minutes}]$$*

*Thus, the total time from taking the first pill to the last pill is \boxed{60} minutes.*

 **Is this thought?**

# A Closer Look At The “Thought” of The Reasoning Model

Okay, let's see. The doctor gives me three pills and tells me to take one every half hour. I need to figure out how many minutes will pass from taking the first pill to the last one. Hmm, let's break this down.

→ **Planning**

First, if I have three pills, that means I have to take them at different times, right? The first pill is taken immediately, I guess. Then the second one half an hour later, and the third one another half hour after that. Wait, but how many intervals are there between the pills?

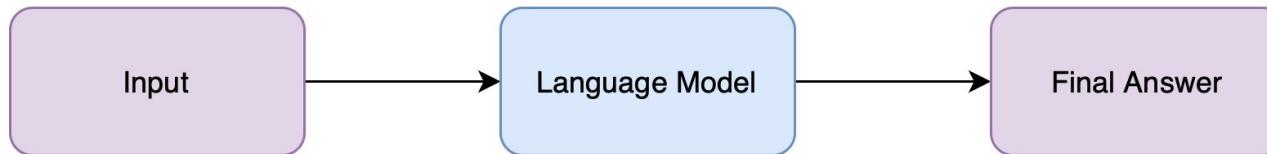
Let me think. If I take the first pill at time zero, the second pill would be at 30 minutes, and the third pill at 60 minutes. So from the first to the third pill, that's 60 minutes. But wait, maybe I need to check that again.

→ **Evaluation**

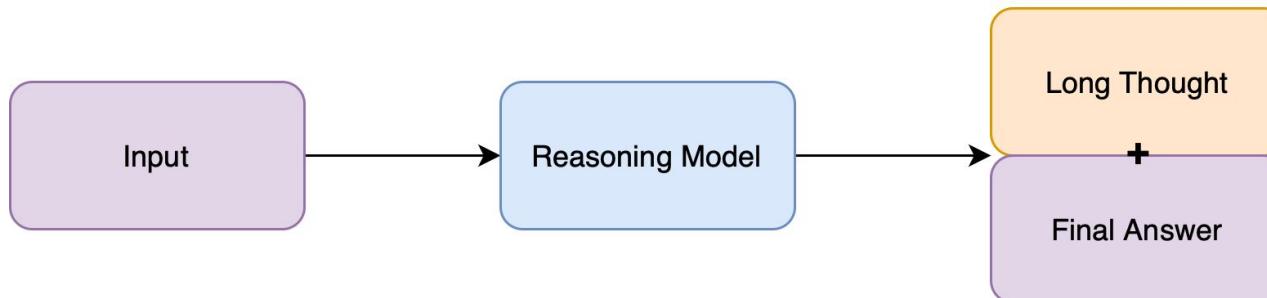
...more...

# Reasoning Models

## Language Model



## Reasoning Model



# Reasoning Models



# “Thought”

Part II

# What Exactly is “Thought”?

**thought** n. /θɔ:t/

1. something you think
2. mind/ideas
3. **process/act of thinking**

# Human Thinking Systems



## System 1 “Fast”

*Assess the situation and deliver updates*

- Unconscious
- Effortless
- Automatic



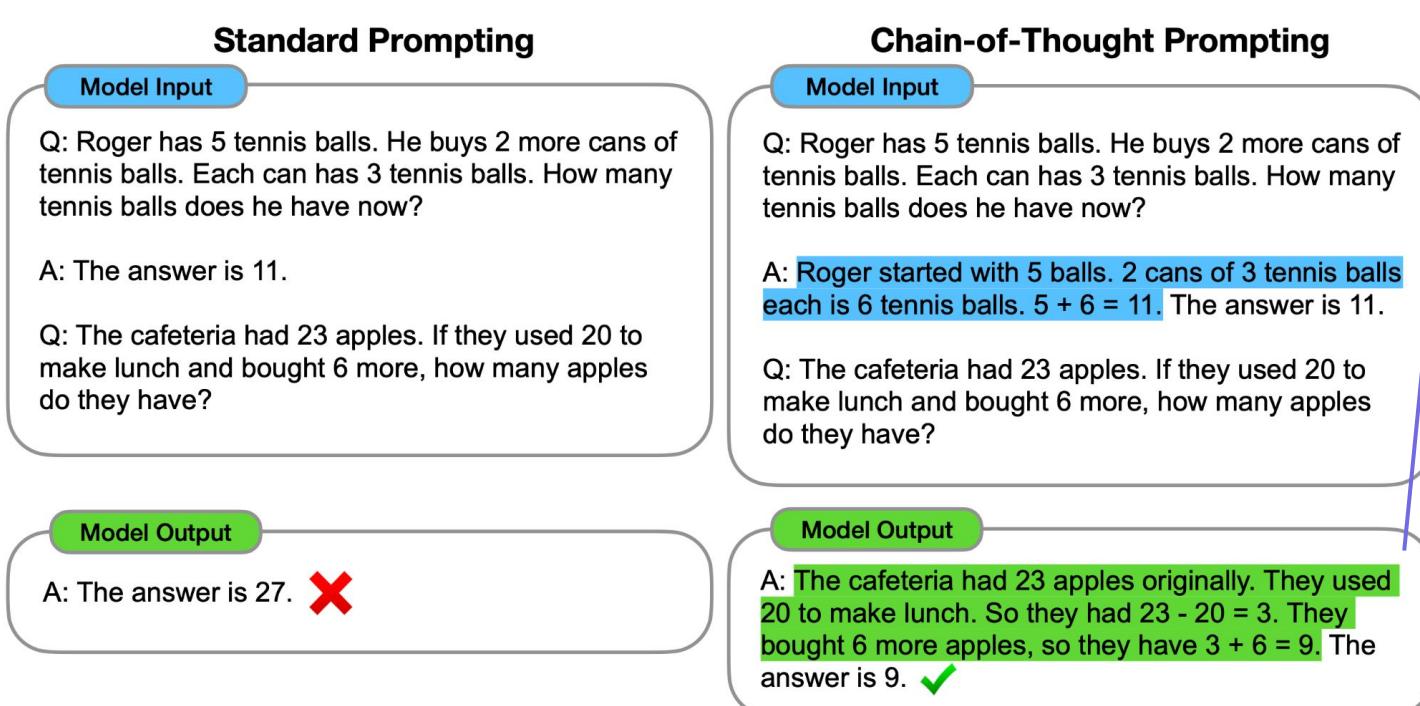
## System 2 “Slow”

*Seek new information and makes decisions*

- Deliberate and conscious
- Effortful
- Controlled mental process

# Chain-of-Thought Prompting

*Here's the "thoughts"!*



# LLMs Know How to Think Without Examples

*Zero-shot Chain-of-Thought Prompting*

*Here's the "thoughts"!*

(a) Few-shot

Q: Roger has 5 tennis balls. He buys 2 more cans of tennis balls. Each can has 3 tennis balls. How many tennis balls does he have now?  
A: The answer is 11.

Q: A juggler can juggle 16 balls. Half of the balls are golf balls, and half of the golf balls are blue. How many blue golf balls are there?

A:

*(Output) The answer is 8. X*

(b) Few-shot-CoT

Q: Roger has 5 tennis balls. He buys 2 more cans of tennis balls. Each can has 3 tennis balls. How many tennis balls does he have now?

A: Roger started with 5 balls. 2 cans of 3 tennis balls each is 6 tennis balls.  $5 + 6 = 11$ . The answer is 11.

Q: A juggler can juggle 16 balls. Half of the balls are golf balls, and half of the golf balls are blue. How many blue golf balls are there?

A:

*(Output) The juggler can juggle 16 balls. Half of the balls are golf balls. So there are  $16 / 2 = 8$  golf balls. Half of the golf balls are blue. So there are  $8 / 2 = 4$  blue golf balls. The answer is 4. ✓*

(c) Zero-shot

Q: A juggler can juggle 16 balls. Half of the balls are golf balls, and half of the golf balls are blue. How many blue golf balls are there?

A: The answer (arabic numerals) is

*(Output) 8 X*

(d) Zero-shot-CoT

Q: A juggler can juggle 16 balls. Half of the balls are golf balls, and half of the golf balls are blue. How many blue golf balls are there?

A: **Let's think step by step.**

*(Output) There are 16 balls in total. Half of the balls are golf balls. That means that there are 8 golf balls. Half of the golf balls are blue. That means that there are 4 blue golf balls. ✓*

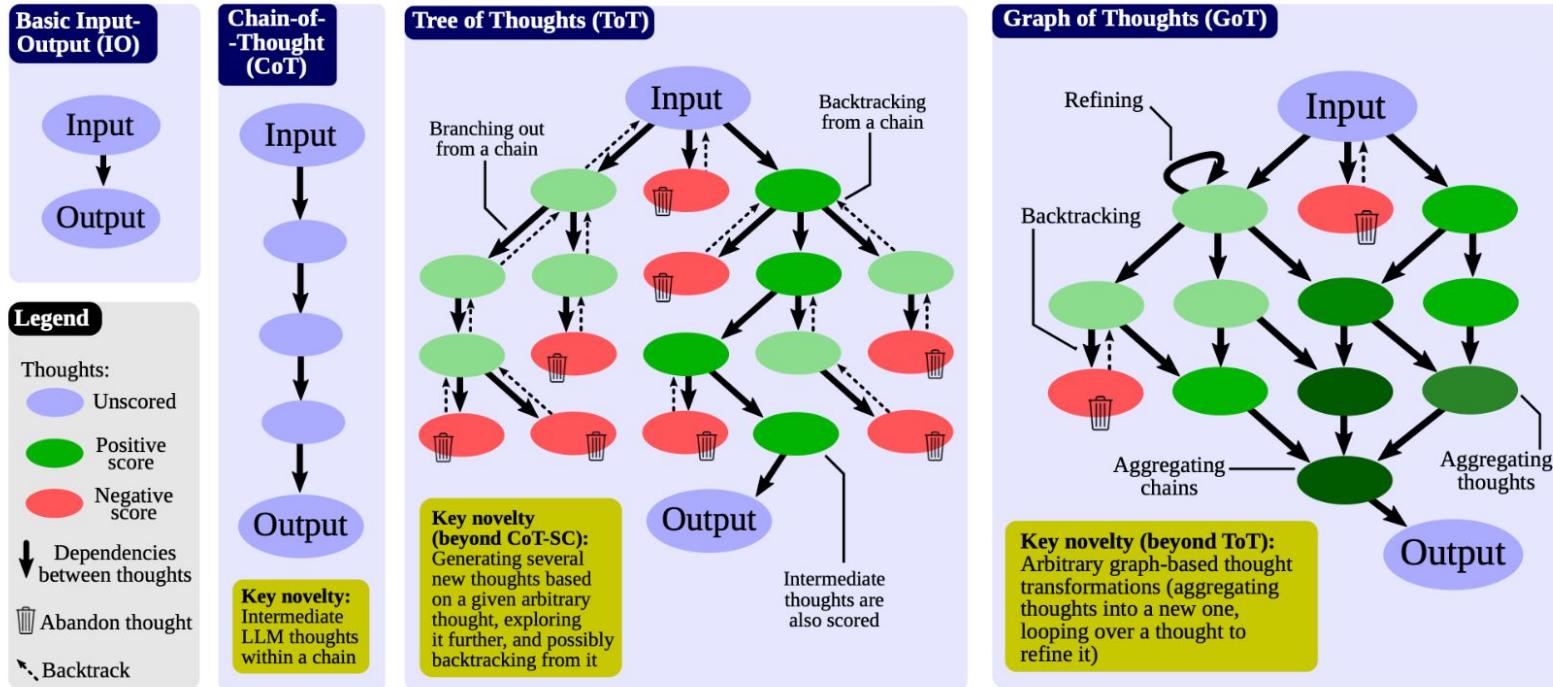
# What Exactly is “Thought” in LLM’s Context?

## thought

1. A thinking step

Since there are *many* steps in the response,  
That's why we called it's a **chain-of-thought** 

# Beyond Linear Thoughts



Shunyu Yao, et al. "Tree of Thoughts: Deliberate Problem Solving with Large Language Models." Thirty-seventh Conference on Neural Information Processing Systems. 2023.

Maciej Besta, et al. "Graph of Thoughts: Solving Elaborate Problems with Large Language Models". AAAI Conference on Artificial Intelligence 2024 21

# Can't We Just Fine-Tune CoT Responses Back to LLMs?

Yes, we can [1] 

However, this is what we got:

(c) Zero-shot

Q: A juggler can juggle 16 balls. Half of the balls are golf balls, and half of the golf balls are blue. How many blue golf balls are there?

A: The answer (arabic numerals) is

(d) Zero-shot-CoT

(Output) There are 16 balls in total. Half of the balls are golf balls. That means that there are 8 golf balls. Half of the golf balls are blue. That means that there are 4 blue golf balls. ✓



This is not exactly the reasoning model we want 

# However, Reasoning Models Are Different

- It's just a ***single*** LLM generates only ***once*** 
  - Unlike ToT or GoT promptings, which involve additional components and require multiple generation
- The reasoning model generates a ***very long chain-of-thought*** 
- The characteristics of the reasoning model's thoughts are ***different***, as they must comprise the following attributes:
  - **Evaluation:** self-verification
  - **Reflection:** self-correction
  - **Exploration:** alternative proposal
  - **Planning:** breakdown a problem into multiple subproblems

# How to Train a Reasoning Model?

Part III

# Machine Learning Paradigms



## Supervised Learning

Given **(input, label)**, the model learns to predict a **(label)** given the **(input)**



## Self-supervised Learning

Given **(input)**, the model learns to predict **(pseudo-label)**, which is generated from **(input)**



## Unsupervised Learning

Given **(input)**, the model learns to discover **(patterns)** or **(structures)**



## Reinforcement Learning

Given **(state, action, reward)**, the model learns to predict **(action)** that will maximize **(reward)**

# How Do We Typically Train An LLM?

## Post-training



### Pre-training

- Teach the model about world knowledge
- Approach: Language modeling: predicting a next token → self-supervised learning



### Mid-training

- a.k.a. Continuous pre-training
  - Correct/add knowledge
  - Improve multilingual abilities
  - Highly similar to pre-training



### Supervised Fine-tuning

- a.k.a. Instruction-tuning
  - Change output format of the model from next token prediction to answer a query
  - Approach: supervised learning



### Preference tuning

- a.k.a Alignment
  - Encourage positive behaviors and discourage negative behaviors
  - Approach: Reinforcement Learning with Human Feedback → RL

# Pre-training

**Why?**

Teaching ***foundational*** language skills to the model, such as ***syntax*** and ***semantic***

**Training Objective**

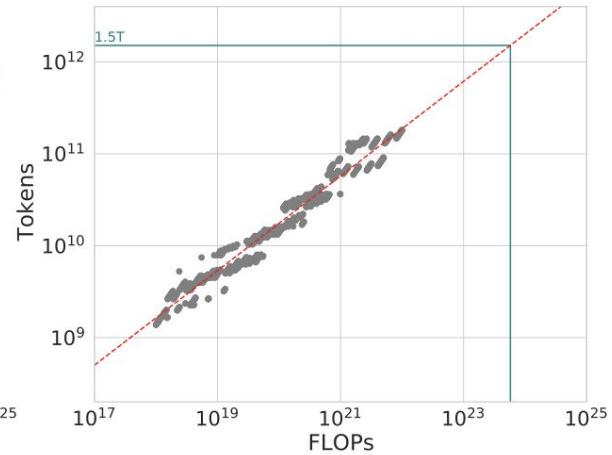
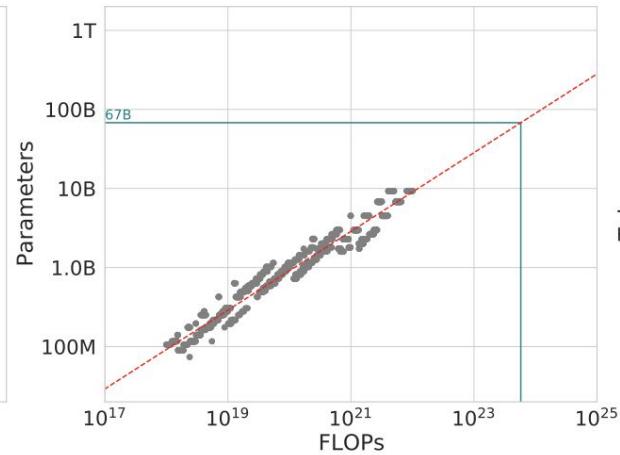
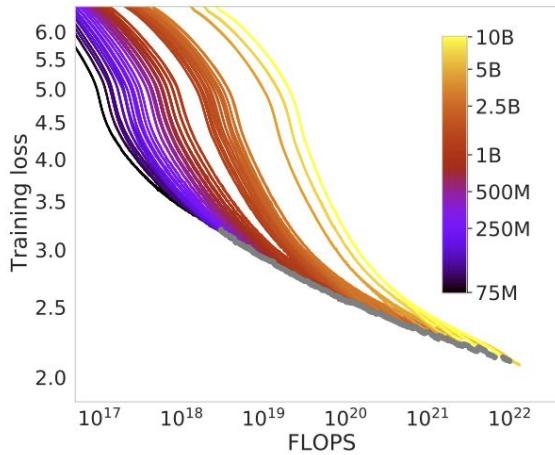
***Language modeling***

Train the model to predict the ***next token*** given the input sequence

We dump the “**filtered**” Internet, humanity’s knowledge, onto the model.

We usually call the result of this process: **Base model**

# Why We Want A “Large” Language Model?: Scaling Law



# Mid-training

Why?

Adding/patching **new** knowledge/skills

Training Objective

*Language modeling*

Train the model to predict the **next token** given the input sequence

Same as *pre-training*

However, splitting pre-training into **two stages** show effectiveness for different purposes

# Post-training

**Why?**

Making the model useful; instruction following, aligning with human preferences

**Training Objective**

*Depends on the stage*

Two primary stages: (1) supervised fine-tuning, (2) preference tuning

Much cheaper than pre-training

We usually call the result of these processes: **Instruct model**

# Supervised Fine-tuning

**Why?**

Teaching a model to follow an instruction, instead of predicting a next token

**Training Objective**

Labelled language modeling

a.k.a. Instruction-tuning

We usually call the result of these processes: **Instruct model**

However, the model *may not be aligned with human preferences yet*

# Preference Tuning

Why?

**Aligning models with human preferences**

Encouraging *preferred* responses and discouraging *undesired* ones

Training Objective

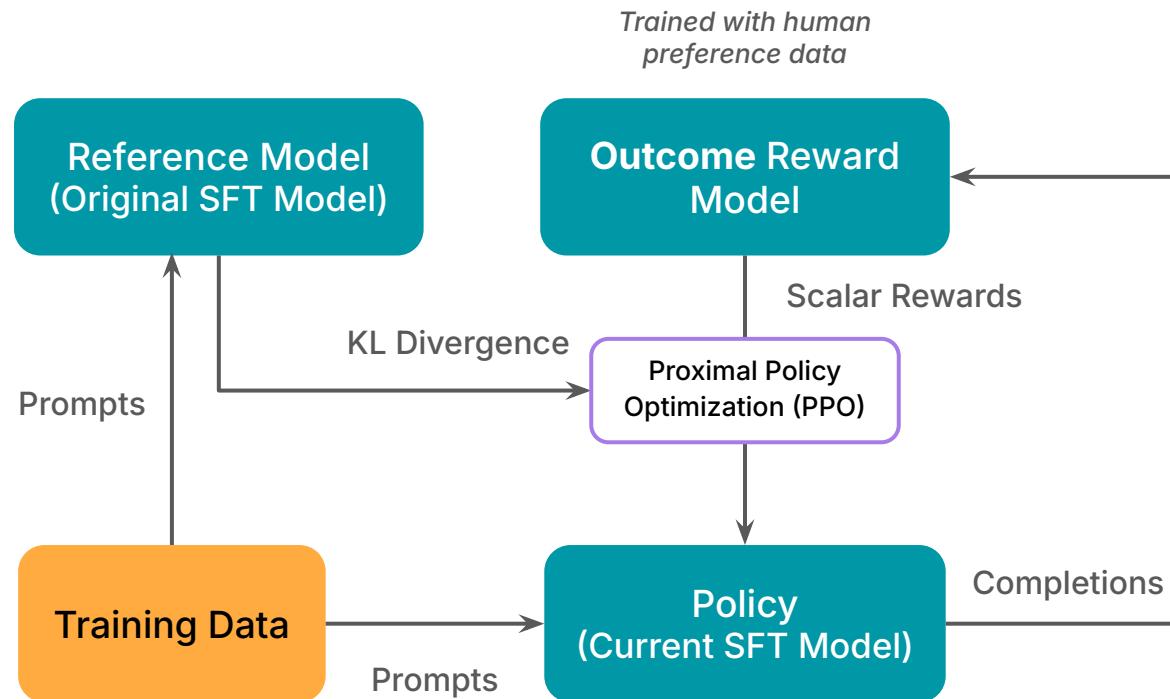
Reinforcement Learning with Human Feedback

a.k.a. Alignment

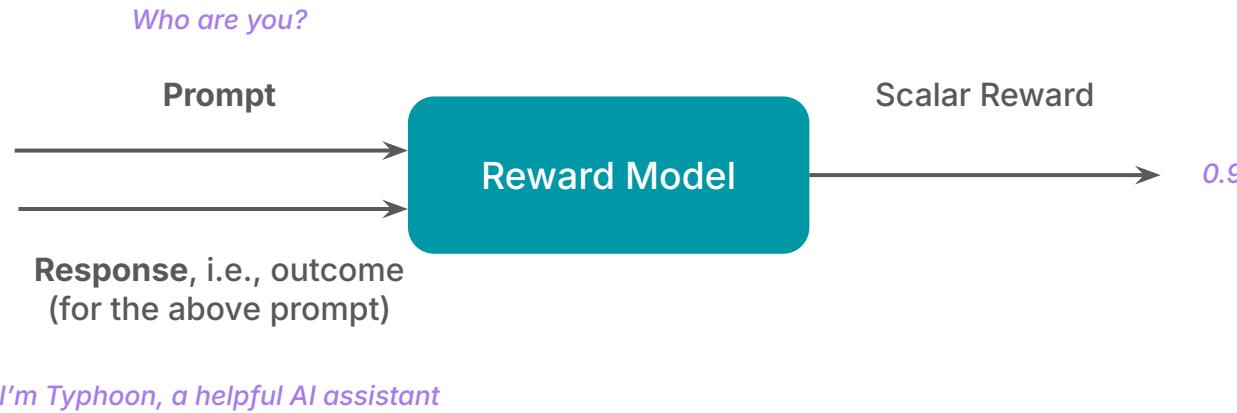
We usually call the result of these processes: **Instruct model**

# Preference Tuning

*Simplified*

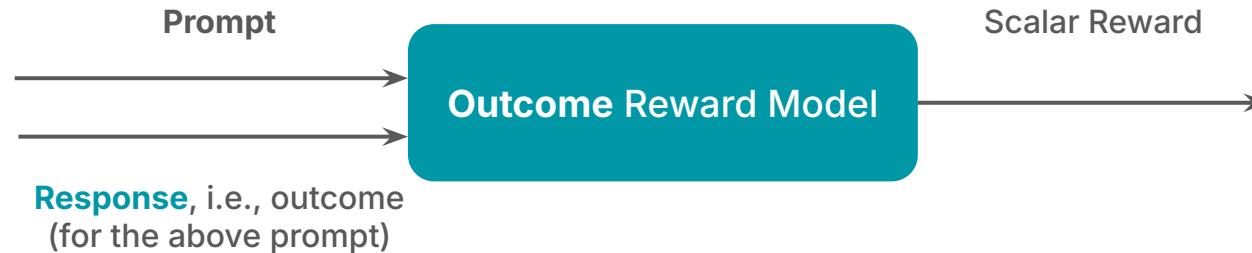


# Reward Models

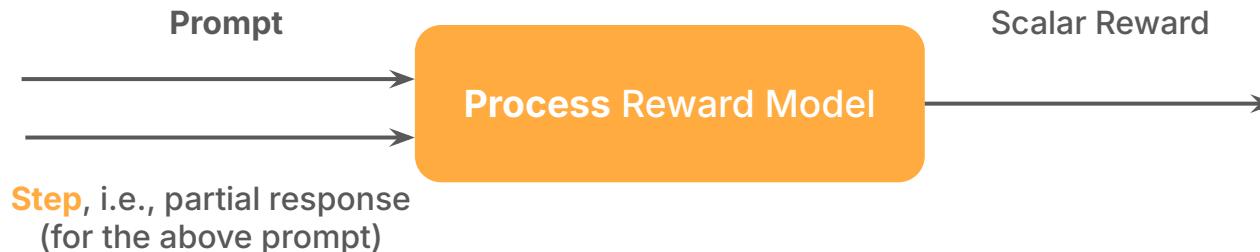


# Different Types of Reward Models

## 1. Outcome Reward Model (ORM)

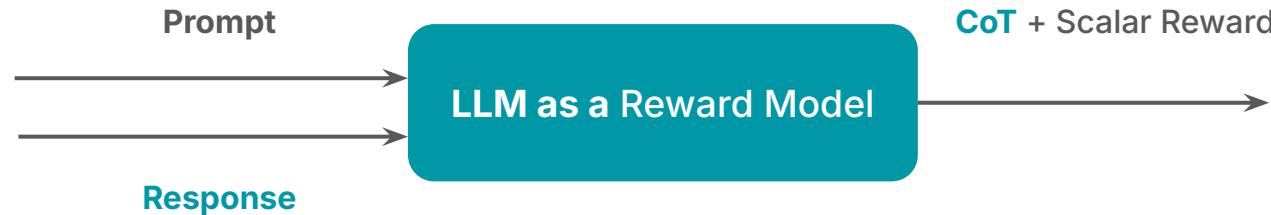


## 2. Process Reward Model (PRM)

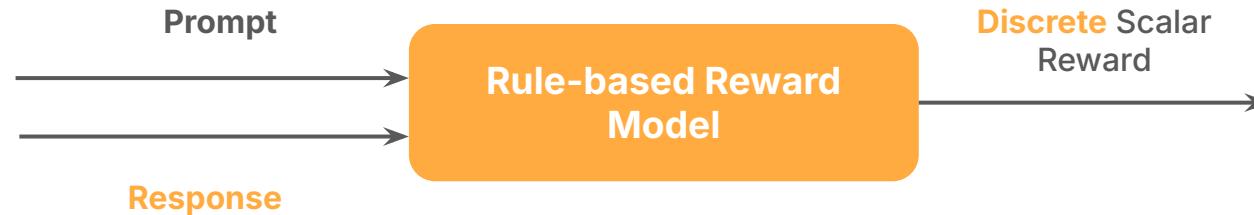


# Different Types of Reward Models

## 3. Generative Reward Model



## 4. Rule-based Reward Function



# So far, so good?



## Pre-training

- Teach the model about world knowledge
- Approach: Language modeling: predicting a next token → self-supervised learning



## Mid-training

- a.k.a. Continuous pre-training
  - Correct/add knowledge
  - Improve multilingual abilities
  - Highly similar to pre-training



## Supervised Fine-tuning

- a.k.a. Instruction-tuning
  - Change output format of the model from next token prediction to answer a query
  - Approach: supervised learning



## Preference tuning

- a.k.a Alignment
  - Encourage positive behaviors and discourage negative behaviors
  - Approach: Reinforcement Learning with Human Feedback → RL

# So, Where Is The Reasoning?

The entire process got us this:

To determine how many minutes will pass from taking the first pill to the last pill, let's break it down step by step:

1. **First Pill:** You take the first pill at time **0 minutes**.
2. **Second Pill:** You take the second pill after **30 minutes** (half an hour). So, the second pill is taken at **30 minutes**.
3. **Third Pill:** You take the third pill after another **30 minutes** (half an hour). So, the third pill is taken at **60 minutes**.
4. **Total Time:** The time from the first pill (0 minutes) to the last pill (60 minutes) is **60 minutes**.

Thus, **60 minutes** will pass from taking the first pill to the last pill.

# So, Where Is The Reasoning?

## And not this:

Okay, let's see. The doctor gives me three pills and tells me to take one every half hour. I need to figure out how many minutes will pass from taking the first pill to the last one. Hmm, let's break this down.

First, if I have three pills, that means I have to take them at different times, right? The first pill is taken immediately, I guess. Then the second one half an hour later, and the third one another half hour after that. Wait, but how many intervals are there between the pills?

Let me think. If I take the first pill at time zero, the second pill would be at 30 minutes, and the third pill at 60 minutes. So from the first to the third pill, that's 60 minutes. But wait, maybe I need to check that again.

Let me visualize it. Let's say I take pill 1 at 12:00 PM. Then, since I need to take one every half hour, the next one would be at 12:30 PM, which is pill 2. Then the third pill would be at 1:00 PM. So from 12:00 PM to 1:00 PM is one hour, which is 60 minutes. So the total time from first to last is 60 minutes. That seems straightforward.

But sometimes these problems can be tricky. Let me consider another approach. If I have three pills, the number of intervals between taking them is two. Because between pill 1 and pill 2 is one interval, and pill 2 and pill 3 is another. Each interval is 30 minutes. So 2 intervals times 30 minutes each is 60 minutes. That matches the previous answer.

Wait, but I've seen similar problems where people might think it's 90 minutes. Maybe they thought three pills, each 30 minutes apart, so 3 times 30. But that would be 90 minutes. But that's incorrect because the first pill is taken at the start, so the time starts counting from there. So the first interval is between the first and second pill, then the second interval is between the second and third. So two intervals. Hence, 30 times 2 is 60.

Let me test with a smaller number. Suppose I have two pills, take one every half hour. Then, take the first at time zero, the second at 30 minutes. So total time is 30 minutes. Which is one interval. So number of pills minus one gives the number of intervals. Therefore, three pills would have two intervals. So two times 30 is 60.

Yes, that makes sense. So the answer should be 60 minutes.

But wait, just to confirm. Let's take another example. If I had four pills, taking one every half hour. Then the intervals would be three, so 3 times 30 is 90 minutes. So from first to last would be 90 minutes. So the formula is (number of pills - 1) multiplied by interval time.

Therefore, applying that here: 3 pills, minus 1 is 2, times 30 minutes equals 60 minutes. Correct.

So the confusion might come if someone counts the number of pills and multiplies by the interval, but that's not the case here because the first pill starts the timer, and each subsequent pill is after an interval. So the total time is intervals between the pills, not the number of pills.

So yeah, the answer is 60 minutes. So when the doctor gives you three pills and tells you to take one every half hour, the duration from the first to the last pill is 60 minutes.

Final Answer

The total time from the first to the last pill is  $\boxed{60}$  minutes.

The doctor gives you three pills and instructs you to take one every half hour. To determine the total time from taking the first pill to the last pill, we need to consider the intervals between each pill.

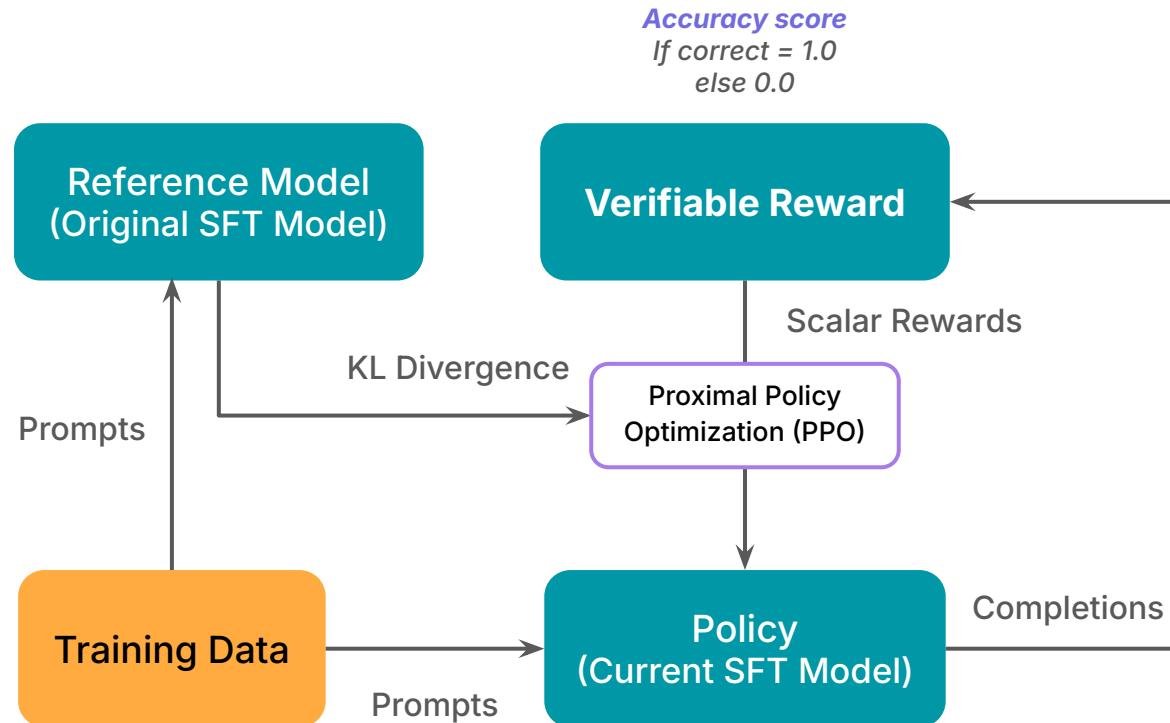
1. First pill: Taken at time 0 minutes.
  2. Second pill: Taken 30 minutes after the first pill.
  3. Third pill: Taken 30 minutes after the second pill.
- The intervals between the pills are calculated as follows:
- Between the first and second pill: 30 minutes.
  - Between the second and third pill: 30 minutes.

There are 2 intervals between 3 pills. Each interval is 30 minutes, so the total time is:  
 $(2 \times 30) = 60$  minutes

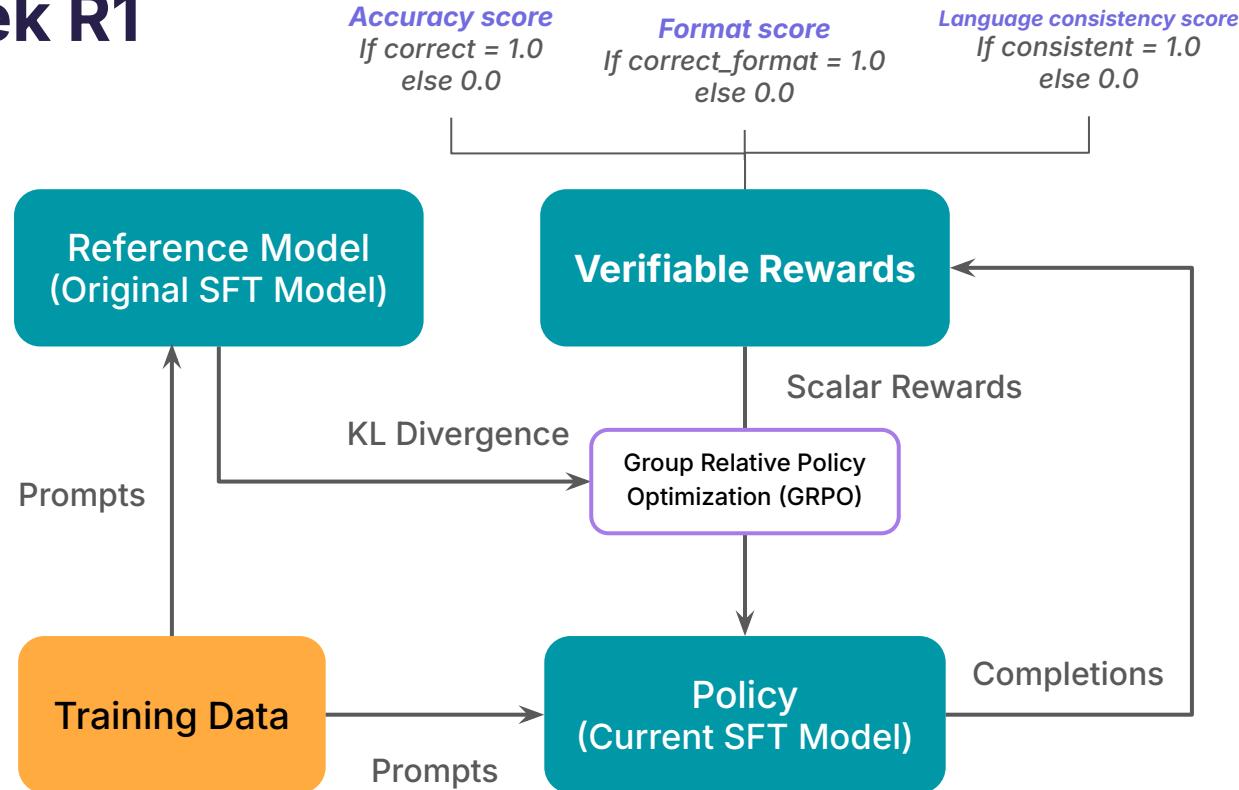
Thus, the total time from taking the first pill to the last pill is  $\boxed{60}$  minutes.

*Where is the missing piece?*

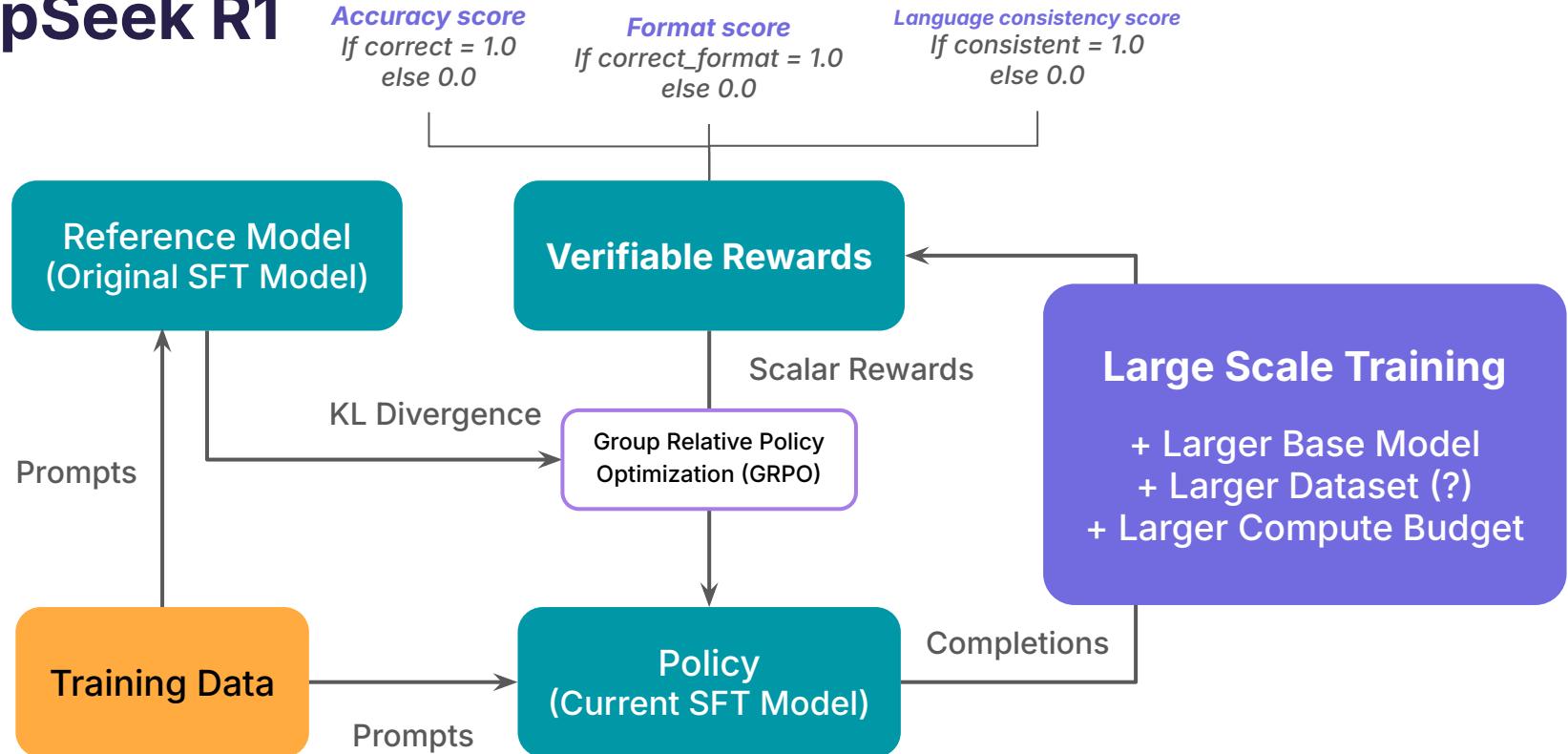
# Reinforcement Learning with Verifiable Reward (RLVR)



# DeepSeek R1

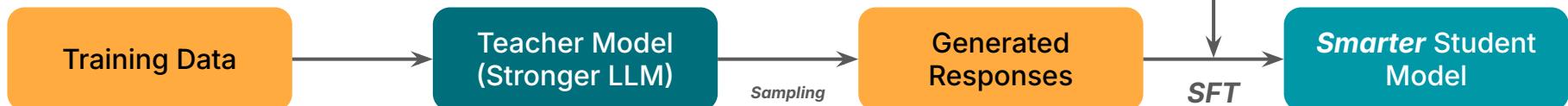


# DeepSeek R1



# Any Alternative Approaches?

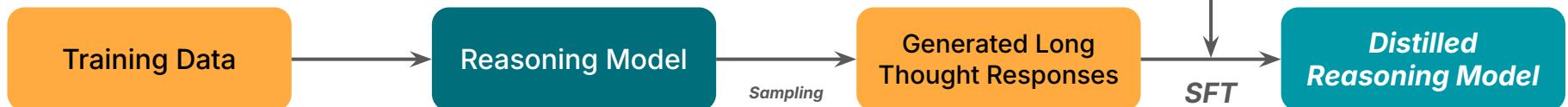
## Knowledge Distillation



Xiaohan Xu, et al. "A Survey on Knowledge Distillation of Large Language Models." (2024).  
DeepSeek-AI, et al. "DeepSeek-R1: Incentivizing Reasoning Capability in LLMs via Reinforcement Learning." (2025).  
SCB 10X R&D. "Typhoon T1 — Our Experiment Towards State-of-the-Art Reasoning Model." (2025).

# Any Alternative Approaches?

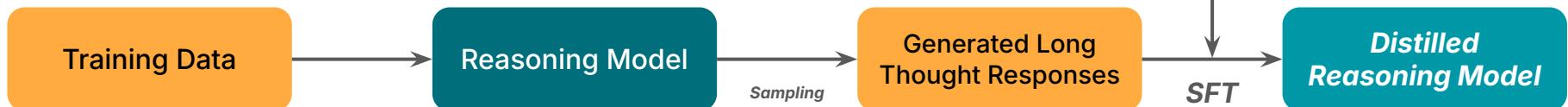
## Knowledge Distillation



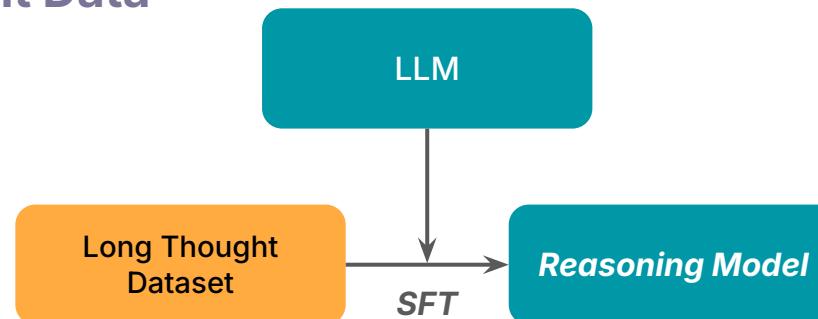
Xiaohan Xu, et al. "A Survey on Knowledge Distillation of Large Language Models." (2024).  
DeepSeek-AI, et al. "DeepSeek-R1: Incentivizing Reasoning Capability in LLMs via Reinforcement Learning." (2025).  
SCB 10X R&D. "Typhoon T1 — Our Experiment Towards State-of-the-Art Reasoning Model." (2025).

# Any Alternative Approaches?

## Knowledge Distillation



## SFT On Long Thought Data



*How to Construct A  
Long Thought Dataset?*

Xiaohan Xu, et al. "A Survey on Knowledge Distillation of Large Language Models." (2024).  
 DeepSeek-AI, et al. "DeepSeek-R1: Incentivizing Reasoning Capability in LLMs via Reinforcement Learning." (2025).  
 SCB 10X R&D. "Typhoon T1 — Our Experiment Towards State-of-the-Art Reasoning Model." (2025).

# Typhoon T1

## A Case Study on Developing A Reasoning Model

### Part IV

# Typhoon T1: Overview



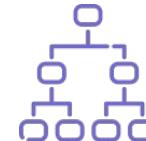
## First “Open” Reasoning Model in SEA

We open our recipe on data and training along with our model weights of Typhoon T1



## No Distillation From Other Reasoning Models

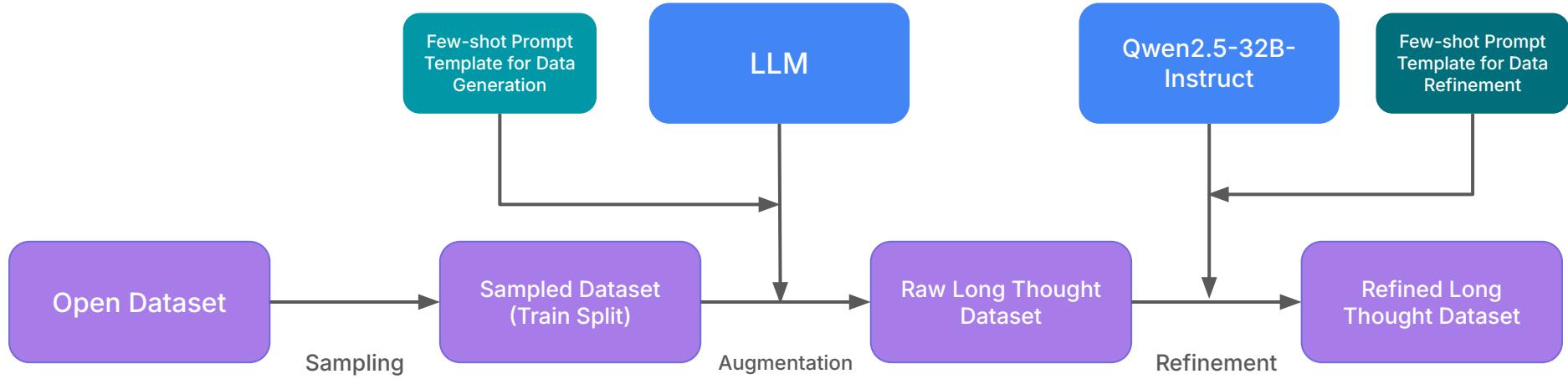
Unlike the other T1, we **did not distill** from other reasoning models to train our Typhoon T1. Typhoon T1 works *across domains*, not limited to math and code



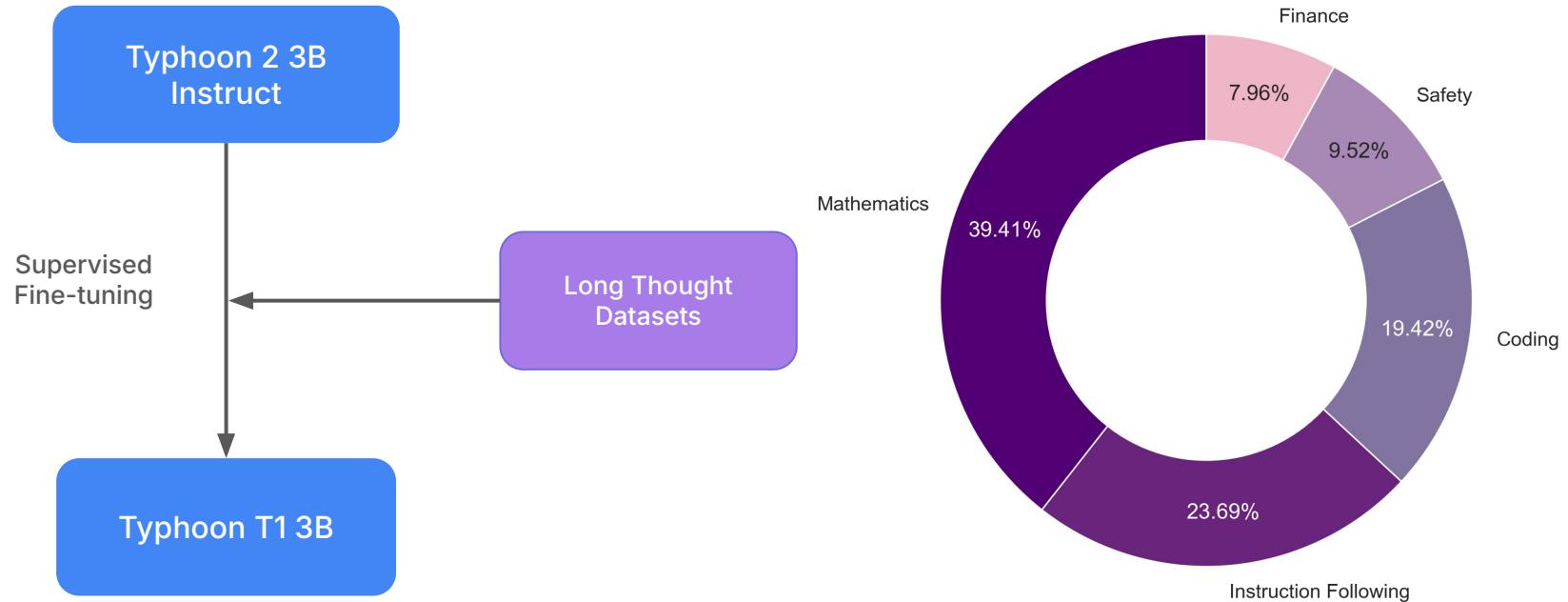
## Structured Thinking for Improved Performance

Structured thinking offers systematic approach with additional auxiliary structural tags for the model to follow outperform its variant without auxiliary tags

# Typhoon T1: Data Pipeline



# Typhoon T1: Supervised Fine-Tuning



# Typhoon T1: Performance

Model name	GSM8K (↑) 8-shot	HumanEval+ (↑) Pass@10	GPQA (↑) 0CoT	MMLU Pro (↑), 5-shot	AIME (↑)
Typhoon 2 3B Instruct	56.63	66.00	27.01	26.70	0.00
Typhoon T1 3B (semi)	59.59	68.99	25.89	-	0.00
<b>Typhoon T1 3B (Research Preview)</b>	<b>62.40</b>	<b>69.87</b>	<b>31.70</b>	<b>30.65</b>	<b>2.22</b>

# Coding Ability

## Prompt

Create a footer with three columns of equal width:

- The first column should contain contact information
- The second column should contain a list of links.
- The third column should contain social media icons (use placeholder icons).

## Typhoon T1

Contact Information	Links	Social Media
Phone: 555-1234 Email: example@example.com	<a href="#">Home</a> <a href="#">About</a> <a href="#">Services</a> <a href="#">Contact</a>	  

*Typhoon T1 is able to generate the correct footer*

## Typhoon 3B Instruct

### Contact Information

123 Main St, Anytown, USA

Phone: (555) 123-4567

Email: info@example.com

### Links

- [Home](#)
- [About Us](#)
- [Services](#)
- [Blog](#)
- [Contact](#)

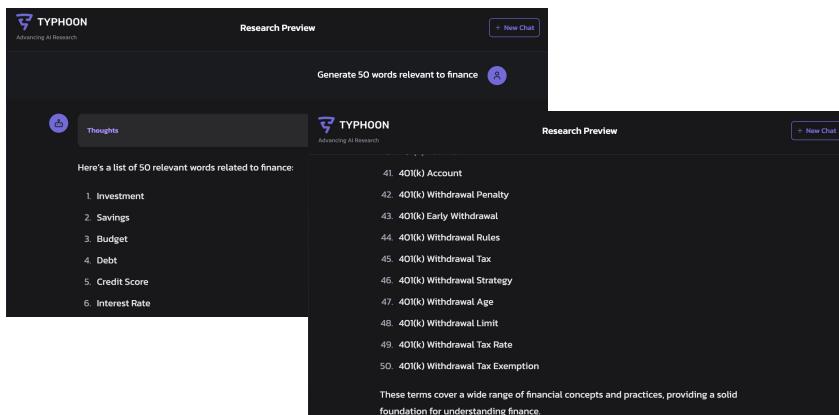
### Social Media

# Domain-specific Jargon

## Prompt

Generate 50 words relevant to finance.

## Typhoon T1



The screenshot shows two side-by-side interfaces of the Typhoon AI platform. Both have a dark header with the 'TYPHOON Advancing AI Research' logo and a 'Research Preview' tab. Below the tabs is a search bar with the placeholder 'Generate 50 words relevant to finance'. The left interface has a 'Thoughts' tab selected, showing a list of 50 terms. The right interface also shows the same list of 50 terms. At the bottom of both interfaces is a note: 'These terms cover a wide range of financial concepts and practices, providing a solid foundation for understanding finance.'

Term	Term
1. Investment	41. 401(k) Account
2. Savings	42. 401(k) Withdrawal Penalty
3. Budget	43. 401(k) Early Withdrawal
4. Debt	44. 401(k) Withdrawal Rules
5. Credit Score	45. 401(k) Withdrawal Tax
6. Interest Rate	46. 401(k) Withdrawal Strategy
	47. 401(k) Withdrawal Age
	48. 401(k) Withdrawal Limit
	49. 401(k) Withdrawal Tax Rate
	50. 401(k) Withdrawal Tax Exemption

## Typhoon 3B Instruct



The screenshot shows the Typhoon 3B Instruct interface. It features a dark header with the 'scb10x/llama3.2-typhoon2-3b-instruct' logo and a 'Research Preview' tab. Below the tabs is a search bar with the placeholder 'Generate 50 words relevant to finance'. The main content area displays a list of 50 terms, which are mostly identical to the ones generated by Typhoon T1, with minor variations in some terms like 'Valuation' appearing multiple times. The list is presented in two columns.

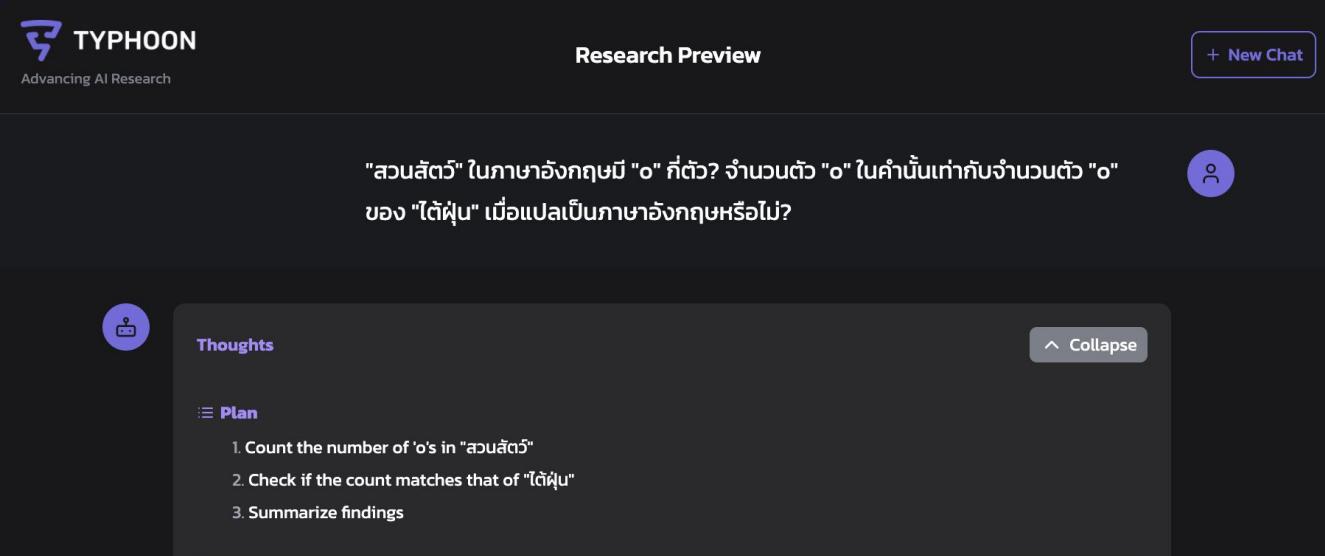
Term	Term
1. Investment	36. Valuation
2. Portfolio	37. Valuation
3. Risk	38. Valuation
4. Return	39. Valuation
5. Market	40. Valuation
6. Stock	41. Valuation
7. Bond	42. Valuation
8. Fund	43. Valuation
9. Interest	44. Valuation
10. Dividend	45. Valuation
11. Inflation	46. Valuation
12. Currency	47. Valuation
13. Exchange	48. Valuation
	49. Valuation
	50. Valuation

*Typhoon T1 provides a diverse and comprehensive list of finance terms, while Typhoon 3B instruct struggles with repetition and lacks variety.*

# Multi-step Reasoning

## Prompt

"สวนสัตว์" ในภาษาอังกฤษมี "o" กี่ตัว? จำนวนตัว "o" ในคำนั้นเท่ากับจำนวนตัว "o" ของ "ໄຕັ້ງບຸນ" เมื่อแปลเป็นภาษาอังกฤษหรือไม่?



The screenshot shows the Typhoon AI Research platform. At the top, there's a header with the Typhoon logo, the text "Research Preview", and a "+ New Chat" button. Below the header, a user message is displayed in a dark-themed chat window:

"สวนสัตว์" ในภาษาอังกฤษมี "o" กี่ตัว? จำนวนตัว "o" ในคำนั้นเท่ากับจำนวนตัว "o" ของ "ໄຕັ້ງບຸນ" เมื่อแปลเป็นภาษาอังกฤษหรือไม่?

Below the message, a "Thoughts" section is shown, containing a "Plan" section with the following steps:

- 1. Count the number of 'o's in "สวนสัตว์"
- 2. Check if the count matches that of "ໄຕັ້ງບຸນ"
- 3. Summarize findings

A "Collapse" button is located at the bottom right of the "Thoughts" section.

*Planning*

# Multi-step Reasoning

## Prompt

"สวนสัตว์" ในภาษาอังกฤษมี "o" กี่ตัว? จำนวนตัว "o" ในคำนั้นเท่ากับจำนวนตัว "o" ของ "ไต้ฝุ่น" เมื่อแปลเป็นภาษาอังกฤษหรือไม่?

### Step 1 - Count the number of 'o's in "สวนสัตว์"

The word "สวนสัตว์" translates to "zoo" in English.

- When I count the 'o's, I find:
- สวนสัตว์ has 2 'o's.

I've counted the 'o's in "สวนสัตว์" and found there are 2.

### Step 2 - Check if the count matches that of "ไต้ฝุ่น"

Now, I need to check "ไต้ฝุ่น":

- ไต้ฝุ่น translates to "typhoon" in English.
- Counting the 'o's in "typhoon":
- ไต้ฝุ่น has 2 'o's.

I've counted the 'o's in "ไต้ฝุ่น" and found there are also 2.

### Step 3 - Summarize findings

Final conclusion:

- "สวนสัตว์" has 2 'o's.
- "ไต้ฝุ่น" also has 2 'o's.
- Therefore, the answer is yes, the number of 'o's in both words is equal.

The findings confirm that both words have the same number of 'o's, which is 2.



Research Preview

Advancing AI Research

+ New Chat

"สวนสัตว์" ในภาษาอังกฤษมี "o" กี่ตัว? จำนวนตัว "o" ในคำนั้นเท่ากับจำนวนตัว "o" ของ "ไต้ฝุ่น" เมื่อแปลเป็นภาษาอังกฤษหรือไม่?



Thoughts

▼ Expand

ในคำว่า "สวนสัตว์" มี 'o' ก็จะหมด 2 ตัว และเมื่อแปลเป็นภาษาอังกฤษว่า "zoo" ก็ยังมี 'o' อีก 2 ตัวเช่นกัน ดังนั้น คำตอบคือ **ใช่** จำนวนตัว 'o' ใน "สวนสัตว์" เท่ากับจำนวนตัว 'o' ใน "ไต้ฝุ่น" (typhoon) ซึ่งก็คือ 2 ตัว.

# Multi-step Reasoning

## Prompt

0.11 กับ 0.9 อันไหนมากกว่า?

### Thoughts

 Collapse

### Plan

1. Understanding the comparison
2. Analyzing the fractions
3. Converting to decimals
4. Comparing the decimal values
5. Summarizing the findings

### Step 1 - Understanding the comparison

The user is asking which number is greater between 0.11 and 0.9. I need to analyze both numbers to determine their relative sizes.

I'm clarifying the task at hand: comparing two decimal numbers.

### Step 2 - Analyzing the fractions

Both numbers are decimals:

- 0.11 has one digit after the decimal point (1).
- 0.9 has one digit after the decimal point (9).

Next, I will convert these decimals into fractions for easier comparison.

I'm preparing to convert the decimals into fractions to facilitate the comparison.

### Step 3 - Converting to decimals

1. Converting 0.11 to a fraction:

• 0.11 can be expressed as  $\frac{11}{100}$ .

2. Converting 0.9 to a fraction:

• 0.9 can be expressed as  $\frac{9}{10}$  or  $\frac{90}{100}$ .

I've successfully converted both decimals into fractions:  $\frac{11}{100}$  for 0.11 and  $\frac{90}{100}$  for 0.9.

### Step 4 - Comparing the decimal values

Now, I will compare the two fractions:

- $\frac{11}{100}$  is less than  $\frac{90}{100}$ .

Since  $\frac{90}{100}$  is greater than  $\frac{11}{100}$ , it follows that 0.9 is greater than 0.11.

I've compared the fractions and determined that  $\frac{90}{100}$  is greater than  $\frac{11}{100}$ .

### Step 5 - Summarizing the findings

Final conclusion:

- 0.11 is less than 0.9.
- Therefore, 0.9 is greater than 0.11.

I'm summarizing my findings to provide a clear answer to the user's question.

# Multi-step Reasoning

## Prompt

0.11 กับ 0.9 อันไหนมากกว่า?

**TYPHOON**  
Advancing AI Research

**Research Preview**

[+ New Chat](#)

**Model**

typhoon-t1-3b-research-preview

**System Prompt**

You are an AI assistant named Typhoon created by SCB 10X to be helpful, harmless, and honest.  
Typhoon is happy to help with

**Parameters**

Temperature	<input type="range" value="0"/>
Top-P	<input type="range" value="0.9"/>
Top-K	<input type="range" value="50"/>
Repetition Penalty	<input type="range" value="1.05"/>

Now, I will compare the two fractions:

- \*  $\frac{11}{100}$  is less than  $\frac{90}{100}$ .

Since  $\frac{90}{100}$  is greater than  $\frac{11}{100}$ , it follows that 0.9 is greater than 0.11.

I've compared the fractions and determined that  $\frac{90}{100}$  is greater than  $\frac{11}{100}$ .

---

Step 5 - Summarizing the findings

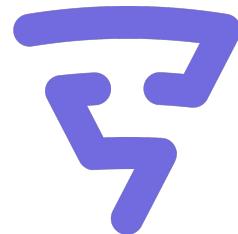
Final conclusion:

- 0.11 is less than 0.9.
- Therefore, 0.9 is greater than 0.11.

I'm summarizing my findings to provide a clear answer to the user's question.

0.11 น้อยกว่า 0.9 เพราะว่า 0.9 มีค่ามากกว่า 0.11 อย่างชัดเจน ดังนั้นค่าตอบคือ 0.9 เป็นจำนวนมากกว่า 0.11

ครับ



# TYPHOON T1

*Research Preview*

Technical Blog



Demo



Model Weights



# What's Next?

Part V

# Open Questions

- Can reasoning models *generalize* across languages?
- *Small* reasoning models
- *Non-distillation approaches* to training reasoning models
- *Better understanding* of reasoning behaviors
- *Generalized reward functions* beyond verifiable tasks
- Teaching models *when not to reason*
- *Efficient* thinking (e.g., critical tokens, O(1)-pruner)
- How can we ensure the *correctness* of reasoning traces?
- ...

“The best way to predict the future is to ***invent*** it.”

— Alan Kay