

# **Fine-tuning LLM for Solving Competitive Programming Problems**

B10902028 王 勻

B10902060 翁苑矜

B10902081 王政祺



# **Introduction**

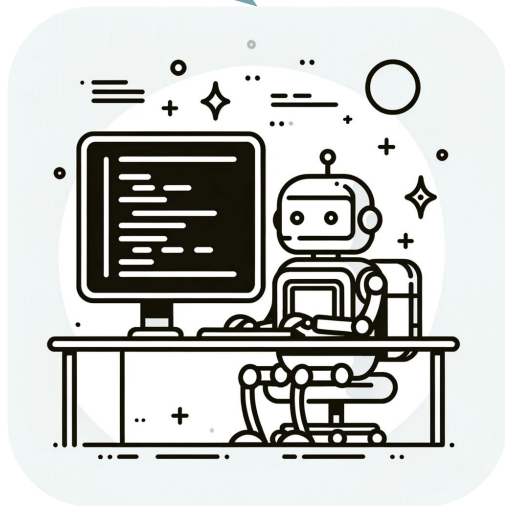
## Problem Description

在最近的字串課程裡，小普學習到了一個新的字串定義，那就是對等。對於兩個長度相同的字串  $SA$  與  $SB$ ，若他們互相對等，則滿足以下兩種條件其中之一：

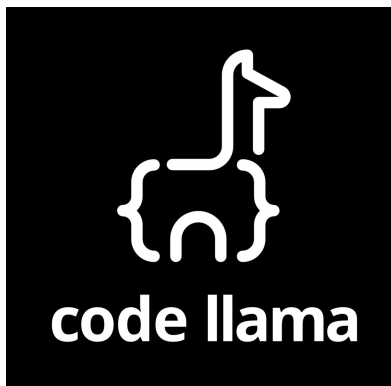
- 1.  $SA$  與  $SB$  完全相同
- 2. 若  $SA$  與  $SB$  的長度都是偶數，我們將字串  $SA$  從中間切開，.....

## Solution Code

```
#include <bits/stdc++.h>
using namespace std;
string f(string s) {
    if(s.size()%2!=0) return s;
    .....
```



# Pretrained Models



# Datasets

**4000+**  
**AC Submission**

Problem  
Description

Solution  
Code

Problem Tag

**Problem Tag  
Examples**

loop

array

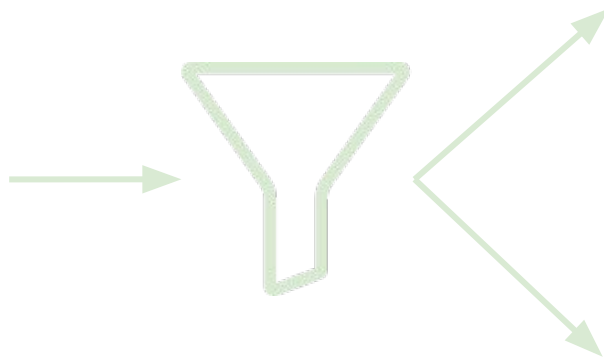
graph

**Train**

Exclude  
test data

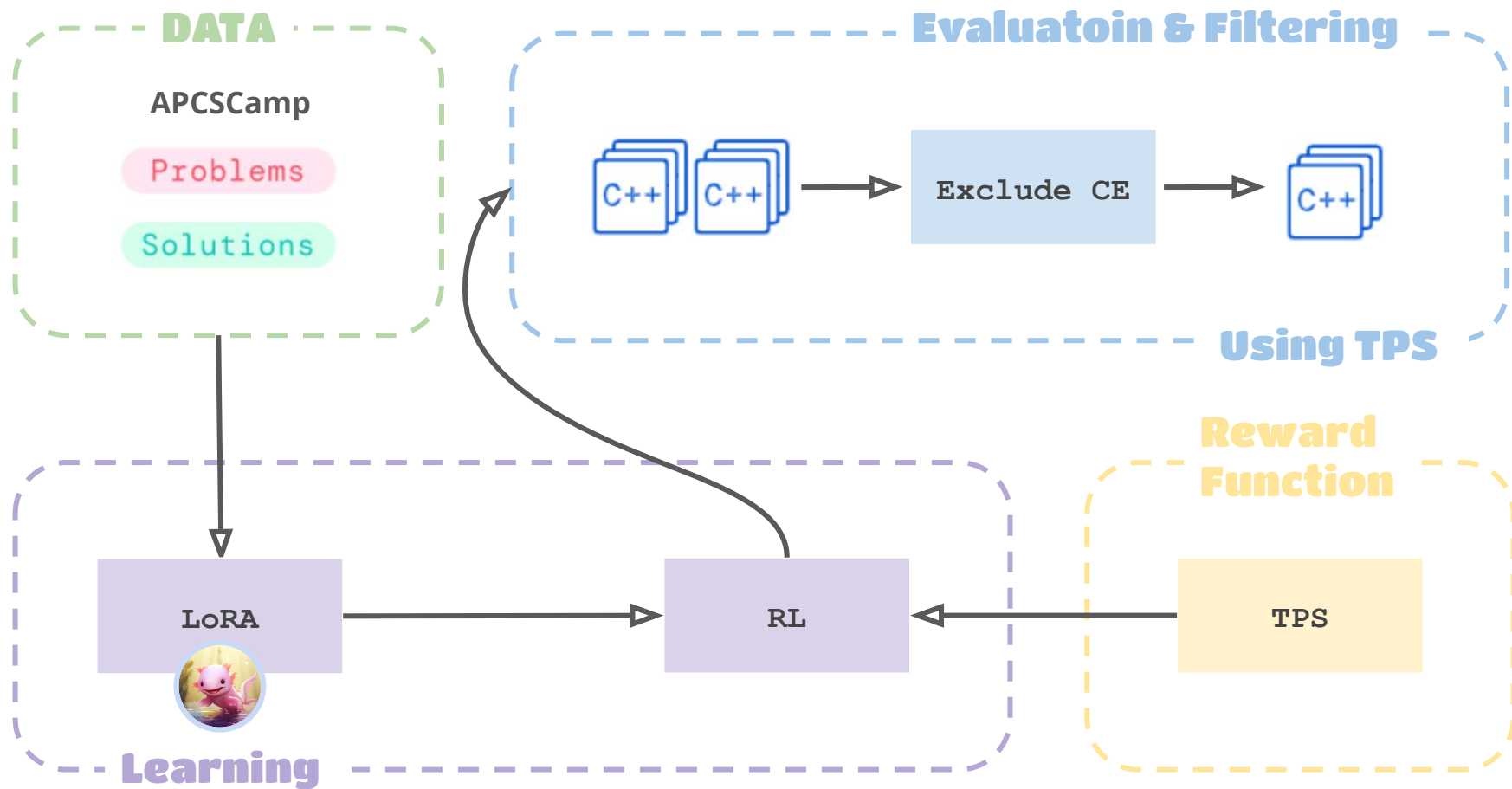
**Test**

5 problems  
for each tag





# **Training Process**



# Reward function

## Compile Error

`-number of errors`

+

`first error position / code length`

## Compile Successfully

`2 + 5 * passed task cases ratio`

+

`5 * (judge scores / 100)`





# **Hyperparameters**

## LoRA

`max_sequence_len = 2048`

`lora_r = 32`

`lora_alpha = 16`

`learning_rate = 2e-4`

`batch_size = 4`

`gradient_accumulation_steps=2`

## RL

`max_sequence_len = 2000`

`learning_rate = 6e-6`

`total_steps = 30000`

`update_interval = 200steps`

`temperature = 1`

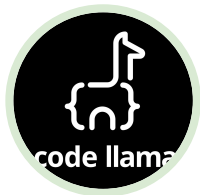


**Result**

# Compile Success rate



**81.3%**



**pretrained**

**1.3 %**

**LoRA**

**92%**

**LoRA + RL**

**91%**



**pretrained**

**0 %**

**LoRA**

**84%**

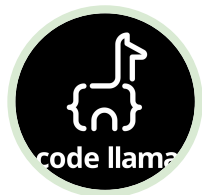
**LoRA + RL**

**67%**

# Accepted rate



**36%**



**pretrained**

**0 %**

**LoRA**

**20%**

**LoRA + RL**

**21%**



**pretrained**

**0 %**

**LoRA**

**33%**

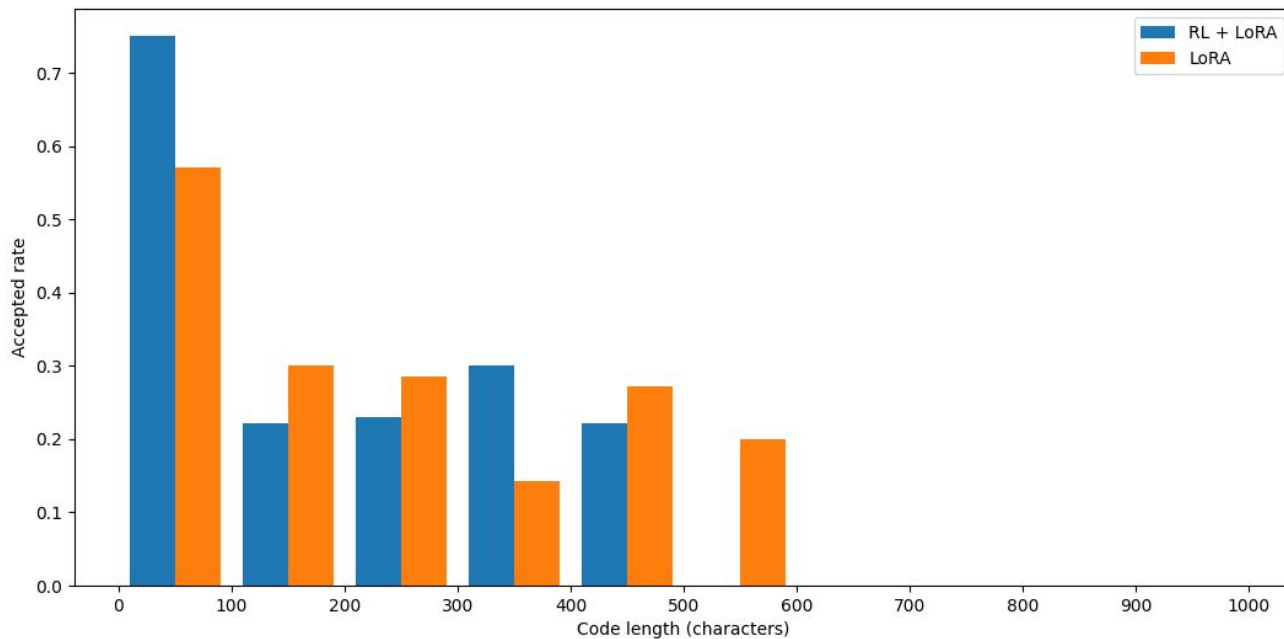
**LoRA + RL**

**28%**

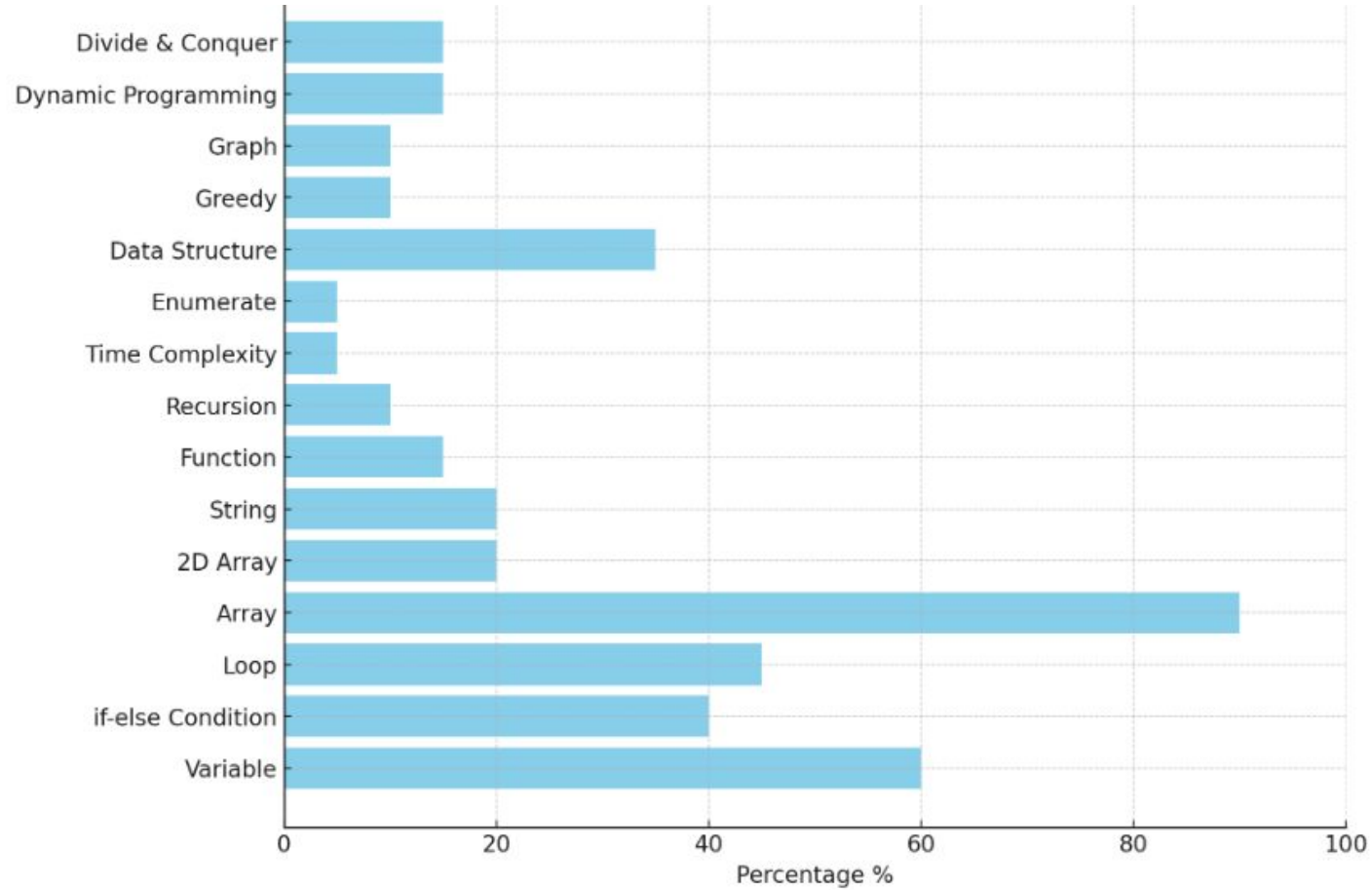


# **Discussion**

# Accepted rate vs Code length



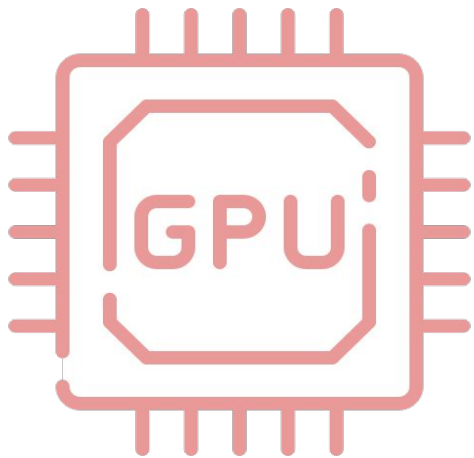
# Accepted rate vs Problem tag







**Limitation**



**GPU memory**

**Dataset size**





**End !**