Queen Mary
University of London

QMUL-BUPT Joint Programme
JP Student Innovation Centre (JPIC)
**Annual Showcase 2021/22**

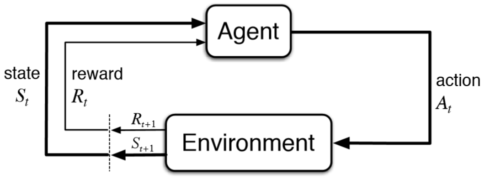
# “Can AI Choose the Right Data Types for us?” Reinforcement Learning for Automatic Software Tuning

Xiao Zheng, Yuanxin Xu, Yujia Guo, Zihao Li
Project Leader: Matthew Tang

## Abstract

It is always a big trouble when we try to analyse codes and programs and try to figure out which is the best way for optimisation. So, what if there is a machine that can automatically help us optimizing our codes?
This project sought for solutions by applying reinforcement learning, where the agent would test different predefined optimisation operations and check their outcomes. If the result is satisfied, the agent would output a Q-table that contains the optimisation suggestions.

## Design



RL State Transition Diagram

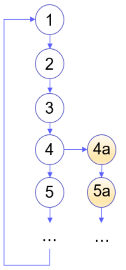
State:
1. Code modification records
2. Current line number being observed

Action: Skip or make changes

Reward: $\frac{(new\ runtime - old\ runtime)}{old\ runtime}$

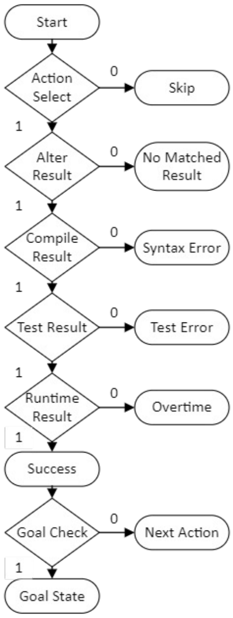
```
1   #include <stdio.h>
2
3   int main() {
4       int n = 10000, sum = 0;
5       for (int i = 0; i < n; ++i) {
6           for (int j = 0; j <= i; ++j) {
7               sum += i + j;
8           }
9       }
10      printf("%d", sum);
11      return 0;
12  }
```

Change: using float instead of int

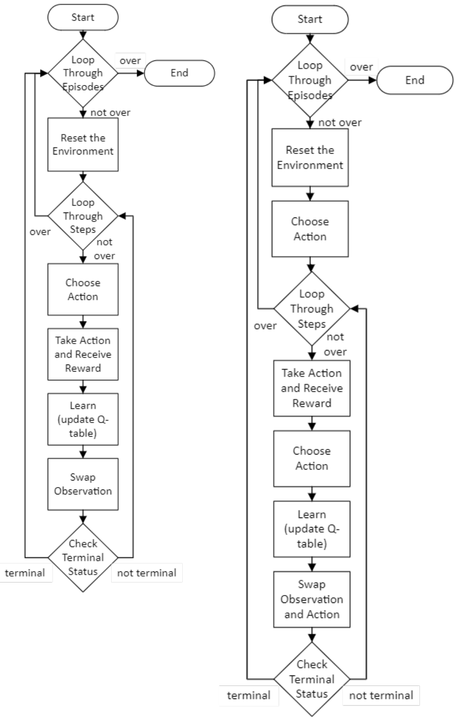


Mapping code into RL Environment

## Implementation



Core Judgement Sequence in Step() (action function)

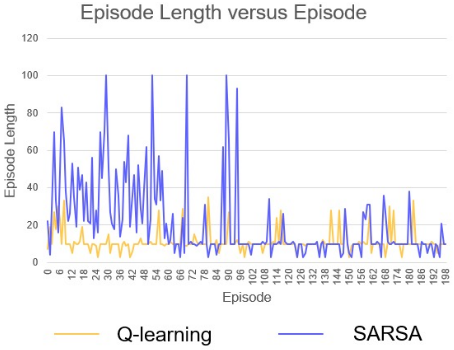


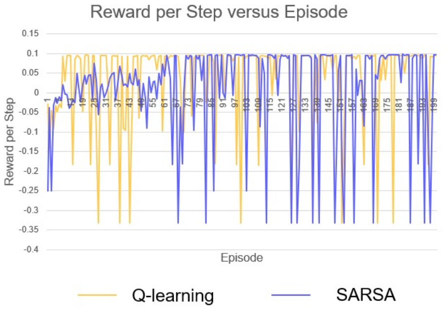
Workflow of Q-learning and SARSA

## Result

| Action / State Reward | **0** | **1** |
|---|---|---|
| **0** | 0.007804 | 0.000108 |
| **1** | 0.019979 | 0.000387 |
| **2** | 0.041704 | -0.08648 |
| **3** | 0 | 0.071095 |
| **3(3)** | 0.006085 | -0.10466 |
| **4(3)** | 0.021464 | 0.000156 |
| **5(3)** | 0.065752 | 0.001038 |
| **6(3)** | 0.001803 | 0.173146 |
| **6(3,6)** | 0.337665 | -0.07726 |
| **7(3,6)** | 0 | 0.636504 |
| **7(6)** | 0 | 0.001732 |
| **2(3,7)** | 0 | -0.0199 |
| **5(3,7)** | 0.000217 | 0 |
| **6** | 0 | -0.00024 |
| **7** | 0 | 0.002987 |

Q-table Result (part)





## Conclusion

- Reinforcement Learning can be used for optimizing codes and programs, if condition permits.
- Q-learning is better than SARSA in the experimental environment.