# Fuzzing and Static Analysis: A Reinforcement Learning Perspective

Dongjae Lee

## Abstract

This essay explains the relationship between fuzzing and static analysis from the reinforcement learning (RL) perspective. Recent fuzzing research has evolved to conduct efficient fuzzing campaigns based on static analysis information. However, fuzzing results are not being utilized in static analysis. In this essay, I propose utilizing fuzzing results in static analysis and explain its significance from an RL perspective.

Recent fuzzing performs efficient fuzzing campaigns based on static analysis information. However, when static analysis goes wrong, the fuzzing campaign can be severely compromised. This is because when using unsound options in static analysis, it heavily relies on human heuristics and intuition.

What if we could refine static analysis results based on actual execution information? We could recover information missed due to unsound analysis or remove unnecessary information included due to overly estimated analysis.

From an RL perspective, static analysis results serve as a model of the actual environment, which is the source code. The fuzzer executes fuzzing by selecting appropriate policies (e.g., selective coverage, energy distribution) based on this model. When sufficient fuzzing results are collected, we can improve the model (static analysis results) based on that information. Fuzzing using the improved model would be more efficient than before.

This problem can be formulated as a Model-based Off-policy RL problem. This problem definition is completelydifferent from the existing RL-based fuzzing researches, which only consider the fuzzing process without static analysis. By combining knowledge from the existing RL field with knowledge from fuzzing and static analysis perspectives, I believe we can solve this problem.