APSEC 2023 December 4-7, 2023 | Grand Walkerhill Seoul, Korea
STAR LAB Software Testing and Analysis Research Lab @ UNIST

# ScenaVRo : Ensuring Variety of Scenario Set for Testing Autonomous Driving Systems

Seongdeok Seo

ULSAN NATIONAL INSTITUTE OF SCIENCE AND TECHNOLOGY 2007

## 1. MOTIVATION

### Existing Generation-based Fuzzing in ADS Testing

- Generation-based Fuzzing

- Focusing on Ego-Vehicle's Route Diversity.
- Generating routes of ego-vehicle to test diverse situations.

ScenoRITA [TSE' 2023], DoppelTest [ICSE' 2023] : Random-based fuzzing

- Randomly configuring the ego-vehicle's driving route across the entire map.

ComOpT [ICRA' 2022], ATLAS [ASE' 2021] : Model-based fuzzing

- Automatically generating scenarios by categorizing only lane of junction.

**LIMITATION 1**

Not considering the route characteristics. (only junction lane characteristics)

**LIMITATION 2**

No method to evaluate variety of the driving scenarios from simulation result.

### Existing Mutation-based Fuzzing in ADS Testing

- Mutation-based Fuzzing

- Focusing on NPC Diversity.
- Perturbing parameters of scenario seed to find violated scenarios.

AV-Fuzzer [ISSRE' 2020], EMOOD [ASE' 2021], Drivefuzz [CCS' 2022], ADFuzz [TSE' 2022], PlanFuzz [NDSS' 2022] ...

- To finding safety violated scenario, they apply evolutionary fuzzing.

CRT [ICRA' 2021], BehAVExplor [ISSTA' 2023]

- Consider covering the widest possible range of the ego-vehicle's driving behaviors alongside safety violation verification. (No Change in route of Ego-Vehicle)

MOSAT [FSE' 2022]

- Standardize NPC vehicle behaviors and generate various NPC actions.

**LIMITATION 1**

The ego-vehicle's routes of mutated scenarios are same as seed scenarios'.

**LIMITATION 2**

No method to evaluate variety of driving scenario sets.

## 2. GOAL

Challenge 1 : How to ensure variety of scenario sets?

Challenge 2 : How to evaluate the uniqueness of scenarios?

### Contributions

- Ego-Vehicle's Route Variation (Section III-A)

- Generate various driving routes with classification automatically.

- Variety Evaluation (Section III-C1)

- Abstract results of scenario execution into essential sequences.

- Test Suite Minimization (Section III-C3)

- Minimize corpus by removing similar scenarios.

## 3. APPROACH

### Input Generation : Model-based Fuzzing

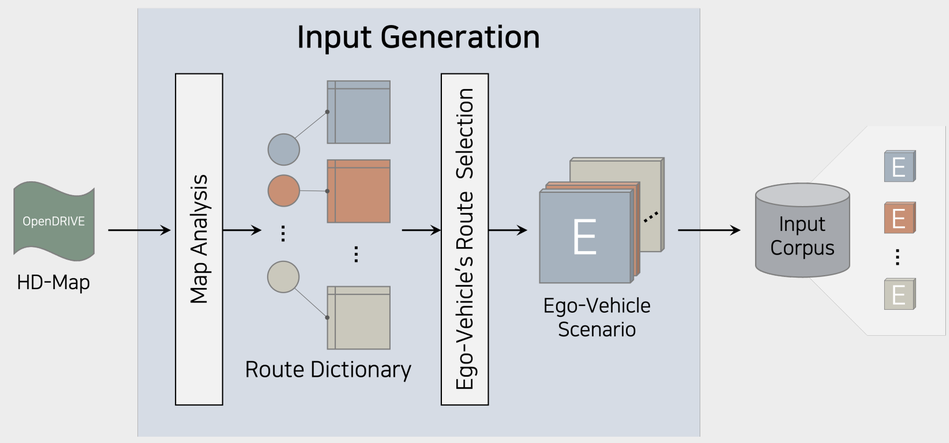



Figure 1. Input generation phase of ScenaVRo

- Map Analysis

- To Construct Road Graph, Parse HD-Map and Analyze it.

- Route Dictionary

- **Minimum Unit Scenario Generation**

  Generate Driving Route which cross one junction (General seed form in prior work)

- **Similar Route Clustering**

  Classify not only lane of junction but also previous lane and next lane.

### Fuzzing Cycle of ScenaVRo : Mutation-based Fuzzing

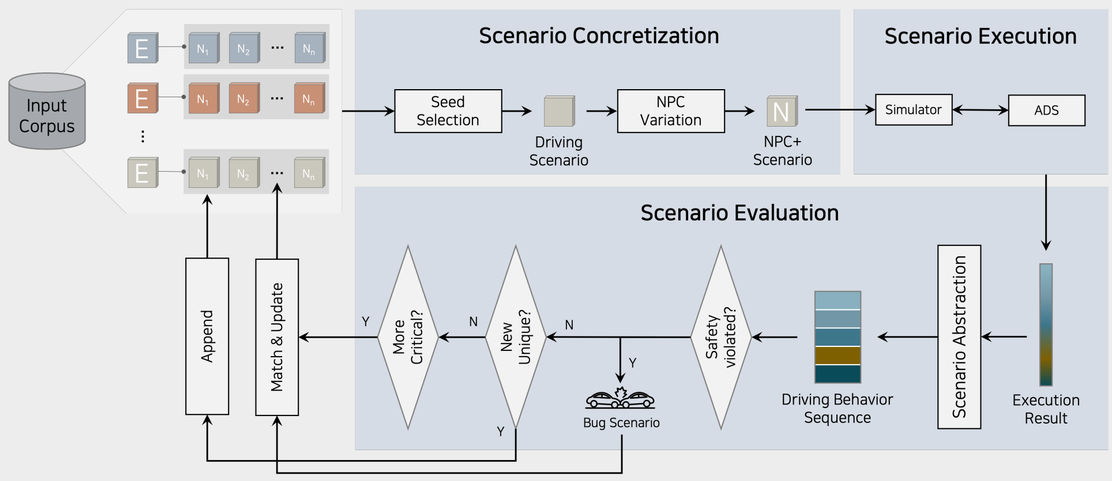
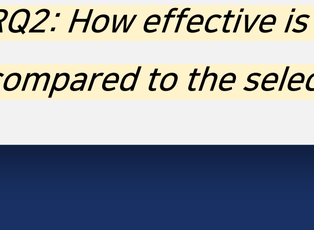



Figure 2. Fuzzing cycle overview of ScenaVRo

- Scenario Concretization

- Appending NPC : Add dynamic NPC using intersection point with ego-vehicle
- Perturbing NPC : Mutable parameter = model, coordinate, speed, trigger ···

- Scenario Evaluation




▼ Figure 4. BNF notation of abstract driving scenario

$$
\begin{aligned}
\langle Scenario \rangle &::= \text{start } \langle \text{LF} \rangle \ \langle tmp \rangle \\
\langle tmp \rangle &::= \langle \text{LF} \rangle \ \langle tmp \rangle \mid \langle \text{LC} \rangle \ \langle tmp \rangle \mid \langle \text{JC} \rangle \ \langle tmp \rangle \\
&\quad \mid \langle \text{SD} \rangle \ \langle tmp \rangle \mid \langle \text{CO} \rangle \ \langle tmp \rangle \mid /\text{exit}
\end{aligned}
$$

▲ Figure 5. ASM for representing outcome of driving scenario execution

- **Scenario Abstraction**

  Translate simulation results into driving behavior sequence (BS).

- **Safety Evaluation**

  Classify highly dangerous driving actions (e.g. collision, lane invasion)

- **Feedback in Fuzzing Cycle**

  Minimize input corpus by eliminating BS-duplicated scenarios.

*RQ1: How effective is ScenaVRo in finding safety violations of ADS compared to the selected baselines?*

*RQ2: How effective is ScenaVRo in generating various driving scenarios compared to the selected baselines?*