Generating Intended Scenarios for Testing Autonomous Driving Systems

Search-Based Testing for Autonomous Driving Systems

Paolo Arcaini, Ezequiel Castellano, Ahmet Cetinkaya, Cédric Ho Thanh, Fuyuki Ishikawa, Stefan Klikovits, Xiao-Yi Zhang

Research content

What do we solve ?

Scenario generation [ICECCS'20, ICST'21, SBST'21]

- Which driving scenarios should we generate? Using which *coverage criteria*? Which scenario will most likely expose faults?
- Which objective functions should guide the search generation?

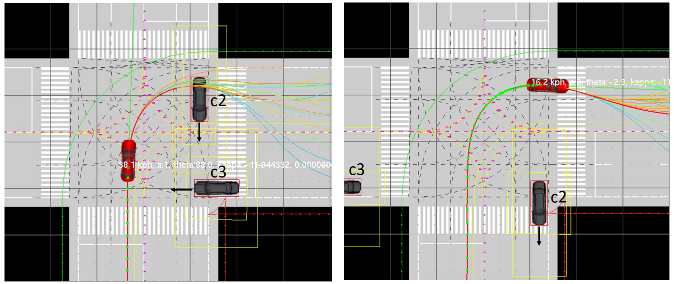
Handling non-determinism of ADSs [QUATIC'21]

- How to handle the noise that is inherent of autonomous driving systems during test generation?

Coverage criteria for scenario generation

We **generate scenarios** to test an industrial path planner. Our research aims to **maximize** the **coverage** of **different decision types** (e.g. avoiding too much curvature) [ICECCS'20] and **different maneuver types** (e.g., turning and accelerating simultaneously) [ICST'21] of the path planner.

Our research introduces objective functions specifying the desired behaviour to cover.

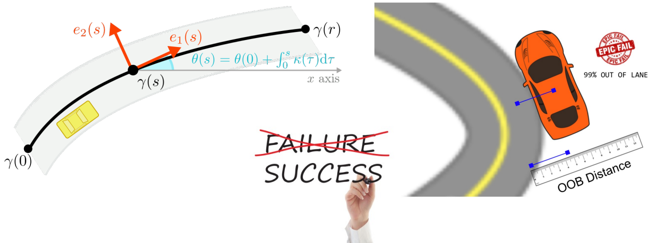


Targeting failures with scenario generation

Frenetic [SBST'21] generates virtual roads to **test the lane-keeping assistant** system of a car in the BeamNG simulator.

The search algorithm tries to **drive the car out of the lane** by maximizing the distance between the car and the center of the lane.

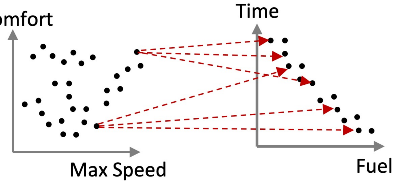
In the ADS testing competition of SBST, Frenetic was among the best ones in **failure detection rate** and produced the most **diverse** test suite.



Handling non-determinism of ADSs

Complex cyber-physical systems (e.g. ADSs) and their simulators are often **non-deterministic**, producing different outputs for the same input, so that observations are **unfaithful**. At the same time repeated simulation of noisy systems is **costly**.

(Non-deterministic) Noisy Objectives



In our work [QUATIC'21], we suggest using observations of nearby solutions to reduce the noise and increase trust in the results without rerunning.

Our research shows that using the **weighted mean** of **k-nearest neighbors** calculates **a more representative value at the same cost.**