Mining Software Repositories 2023 (MSR 2023)
Data and Tool Showcase Track

simula

DeepScenario: An Open Driving Scenario Dataset for Autonomous Driving System Testing

Chengjie Lu[1,2], Tao Yue[1] and Shaukat Ali[1]

[1]Simula Research Laboratory,
[2]University of Oslo

DeepScenario Dataset (Github repository): *https://github.com/Simula-COMPLEX/DeepScenario*

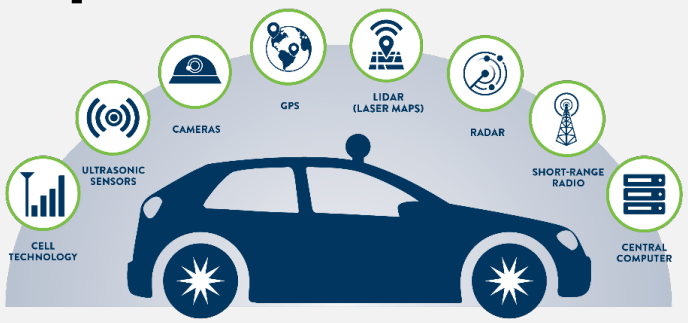
DeepScenario Toolset: https://github.com/Simula-COMPLEX/DeepScenario/tree/main/deepscenario-toolset

CONTENTS

Motivation

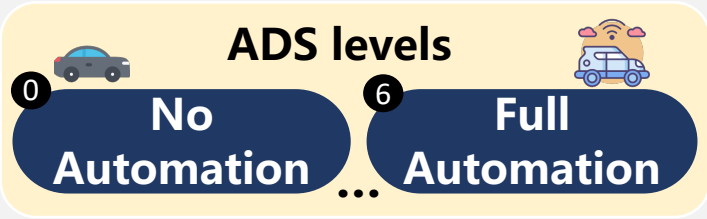
Testing autonomous driving system (ADS) is important and challenging

Importance



Safety-critical system

Continuous evolution



Insufficient test data

No sufficient test data or test scenarios for testing ADS

Limited budget

Exhaustive ADS testing is impossible due to practical constraints

Not adapted to evolution

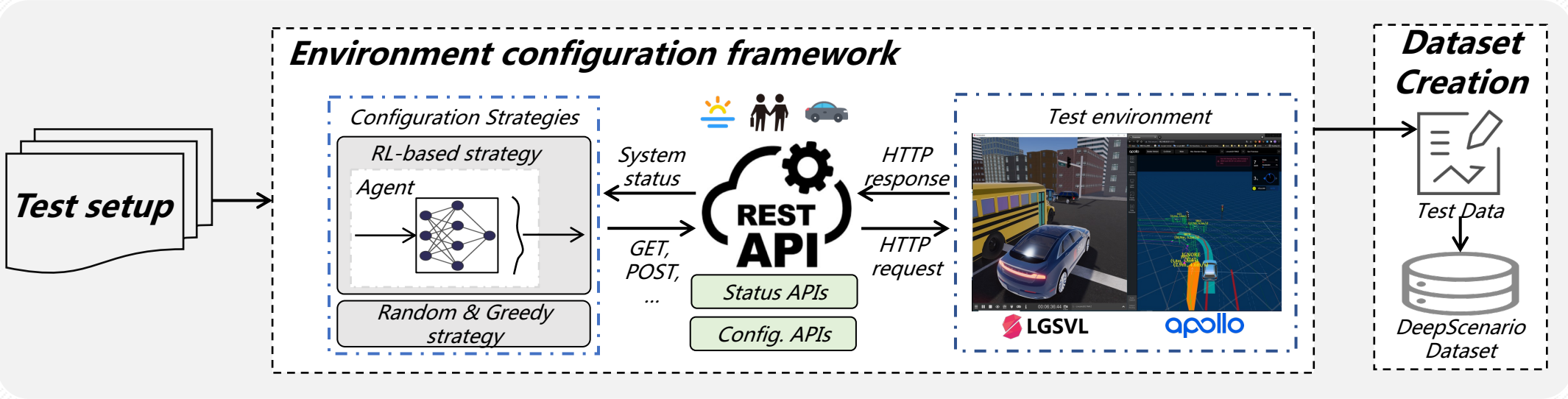
Current approaches are not adapted to system evolution

Challenges

DeepScenario: An Open Driving Scenario Dataset for ADS Testing

Methodology

DeepScenario dataset was created by executing an environment configuration framework

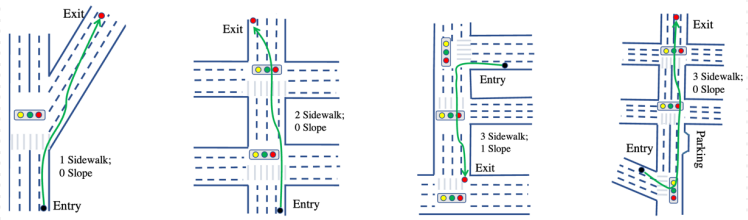


DeepCollision	Baselines
RL-based environment configuration learning	Random & Greedy

Reward	Road/Task	Weather
Safety: R_{TTC}, R_{DTO} Comfort: R_{Jerk}	R1, R2 R3, R4	Rainy day, Sunny day Rainy night, Sunny night

Lu, Chengjie, et al. "Learning configurations of operating environment of autonomous vehicles to maximize their collisions." ***IEEE Transactions on Software Engineering*** 49.1 (2022): 384-402.

TTC: Time-To-Collision
DTO: Distance-To-Obstacle
Jerk: The changing rate of acceleration



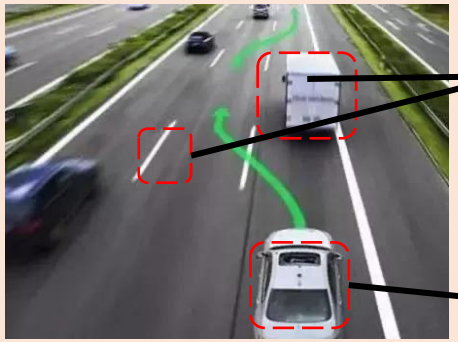
(a) R1 (1.0km) (b) R2 (1.2km) (c) R3 (1.0km) (d) R4 (1.5km)

Methodology

We design DeepScenario toolset to facilitate recording and executing driving scenarios

Scenario definition

$S =< scene_1, scene_2, \ldots, scene_n, T\}$



Lane changing

Environment
Weather
Road user
……

Ego Vehicle
Acceleration
Speed
……

ScenarioCollector

```
<DeepScenario timestep="0.5">
    <FileHeader author="Greedy_Strategy" date="2021-11-27T10:00:00" description="DeepScenario Format" simulatorVersion="2021.01"/>
    <Environment>
        <HDMap city="SanFrancisco"/>
        <EnvironmentInitialization dateTime="2021-07-08 08:00:00" fixTime="False" unixTimeStamp="1625702400"/>
        <OpenWeatherDatabase filepath="SanFrancisco_2021-7-8-rain(RD)"/>
    </Environment>
    <Entities>
        <ScenarioObject model="8c954dd5-e44d-4bea-ac1c-8d103a49e719" name="Ego0" uid="cb095417-fef3-4cff-b52b-22a464fd831f">
            <ObjectParameter ObjectColor="Default" ObjectColorVector="Vector(-1, -1, -1)" ObjectType="Ego"/>
        </ScenarioObject>
        <ScenarioObject model="BoxTruck" name="NPC0" uid="BoxTruck(Clone)48bbebb0-b126-4978-b2eb-2acb7184b7d0">
            <ObjectParameter ObjectColor="Default" ObjectColorVector="Vector(1, 0, 0)" ObjectType="NPC"/>
        </ScenarioObject>
    </Entities>
    <StoryBoard>
        <Initialization>
            <ObjectInitialization objectRef="Ego0">
                <ObjectInitialPosition positionX="-201.879" positionY="10.27" positionZ="217.72"/>
                <ObjectInitialRotation rotationX="0.01" rotationY="180" rotationZ="-0.004"/>
                <ObjectInitialVelocity velocityX="-0.0" velocityY="-0.25" velocityZ="-0.0"/>
                <ObjectInitialAngularVelocity angularVelocityX="-0.001" angularVelocityY="-0.0" angularVelocityZ="0.001"/>
            </ObjectInitialization>
            <ObjectInitialization objectRef="NPC0">
                <ObjectInitialPosition positionX="-220.176" positionY="10.208" positionZ="236.698"/>
                <ObjectInitialRotation rotationX="-0.0" rotationY="85.842" rotationZ="-0.0"/>
                <ObjectInitialVelocity velocityX="0" velocityY="0" velocityZ="0"/>
                <ObjectInitialAngularVelocity angularVelocityX="0" angularVelocityY="0" angularVelocityZ="0"/>
            </ObjectInitialization>
        </Initialization>
```

Scenario file

ScenarioRunner

Scenario info

Replaying scenario

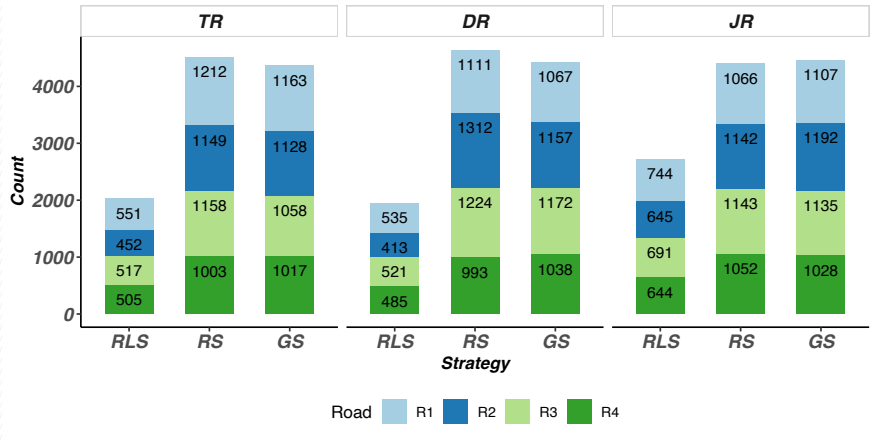
```
"NPC0": {
    "model": "Sedan",
    "color": "white",
    "type": "NPC",
    "uid": "Sedan(Clone)362986ab-a2a8-49d1-aa74-6ec130ebe17a",
    "initialization": {
        "position": {
            "x": -40.673,
            "y": -1.036,
            "z": -11.588
        },
        "rotation": {
            "x": 0.0,
            "y": 104.823,
            "z": 0.0
        },
```

Scenario diagnoses

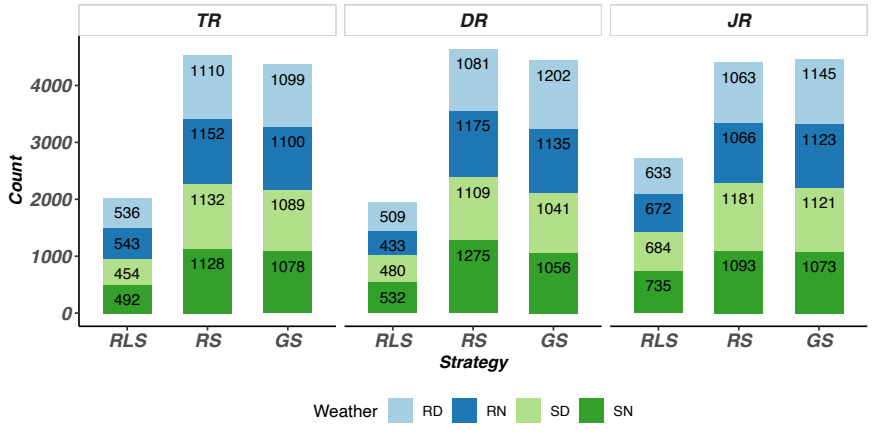
Dataset Analysis

We show Scenario distribution across roads and weather conditions

Scenario distribution across roads



Scenario distribution across weather conditions



Road 4 has the least number of scenarios generated.

There is not much difference observed among the four weather conditions.

R4 has more complicated road structure, which is more prone to collisions. Therefore, test ends earlier.

One plausible explanation is that the ADS is advanced enough to handle different weather conditions.

Dataset Analysis

We show scenario attributes and their statistics

Reward attributes

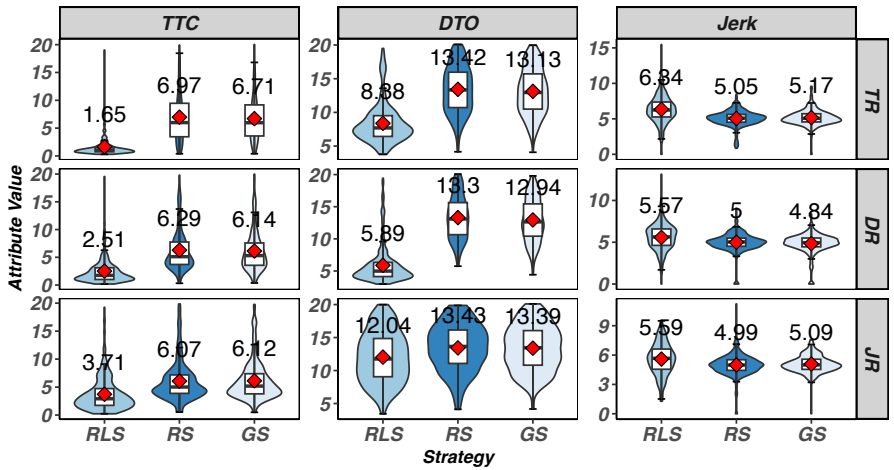
Time-To-Collision (TTC)
Distance-To-Obstacle (DTO)
Jerk

Collision attributes

Collision (COL)
Collision-Type (COLT)
Speed-At-Collision (SAC)

Scenario attributes are used to characterize driving scenarios in the dataset.

Statistics of reward attributes



RLS outperforms RS and GS in terms of all three reward attributes.

Statistics of collision attributes

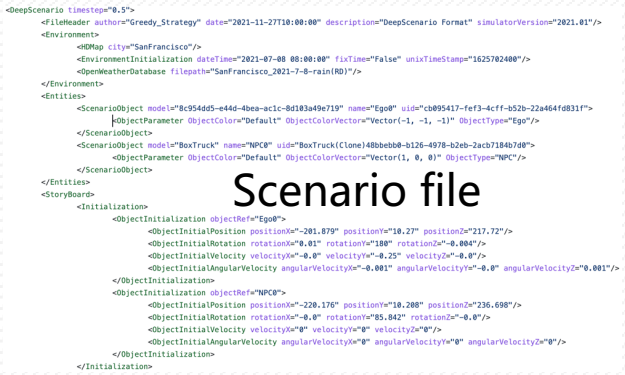
ST	COLT	R_{TTC}		R_{DTO}		R_{Jerk}		$\#C_{ST}$	SAC_{ST}
		#C	SAC	#C	SAC	#C	SAC		
RLS	NPC	250	4.97	222	3.76	153	5.83	**625**	**4.75**
	PED	2	6.27	21	5.04	27	5.27	**50**	**5.21**
	STA	4	6.08	5	3.30	16	3.39	**25**	**3.80**
	SUM	**256**	**5.00**	**248**	**3.86**	**196**	**5.55**	**700**	**4.75**
RS	NPC	36	3.31	47	3.17	53	2.87	**136**	**3.09**
	PED	5	3.49	8	4.54	9	3.88	**22**	**4.03**
	STA	0	N/A	0	N/A	0	N/A	**0**	**N/A**
	SUM	**41**	**3.34**	**55**	**3.37**	**62**	**3.02**	**158**	**3.22**
GS	NPC	60	3.81	44	3.87	52	3.24	**156**	**3.64**
	PED	13	3.63	11	3.83	9	4.56	**33**	**3.95**
	STA	0	N/A	3	0.63	0	N/A	**3**	**0.63**
	SUM	**73**	**3.78**	**58**	**3.70**	**61**	**3.43**	**192**	**3.64**
Total		**370**	**4.57**	**361**	**3.76**	**319**	**4.65**	**1050**	**4.32**

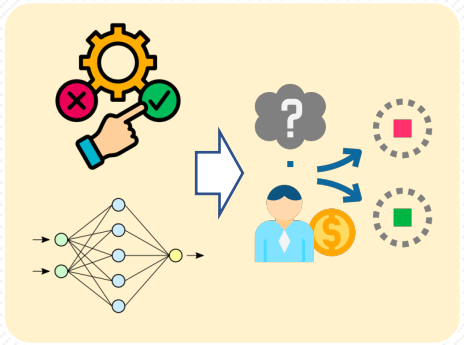
RLS achieved the best performance in generating collision scenarios.

Dataset Usages

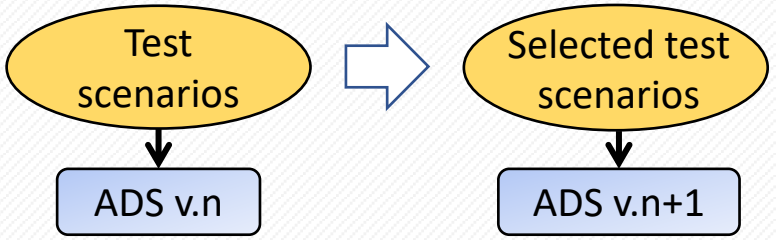
DeepScenario can be used in various ADS development contexts

Analyze scenarios for further diagnoses

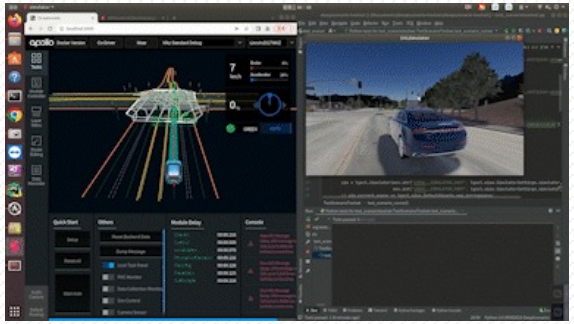




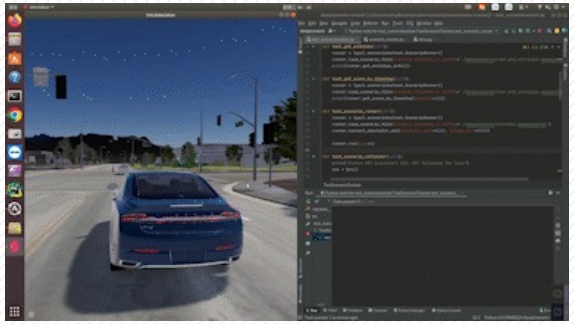
Select and prioritize test scenarios for regression testing



Replay scenarios to detect system-level failures



With ADS enabled



Without ADS enabled

Further resources

Usage example 1: https://www.youtube.com/watch?v=M1c0zP_cKz8
Usage example 2: https://www.youtube.com/watch?v=M1c0zP_cKz8
Usage example 3: https://www.youtube.com/watch?v=M1c0zP_cKz8
Usage example 4: https://www.youtube.com/watch?v=M1c0zP_cKz8

Questions?

Lu, Chengjie, et al. "Search-based selection and prioritization of test scenarios for autonomous driving systems." *Search-Based Software Engineering: 13th International Symposium, SSBSE 2021*