SBFT/ICST TOOL COMPETITION: SELF-DRIVING CAR TESTING

Aryan Prakash
University of Bern
Switzerland

Christian Birchler
University of Bern
Switzerland

Tommaso Fulcini
Politecnico di Torino
Italy

Luigi Sterace
Università degli Studi di Napoli Federico II
Italy

Sebastiano Panichella
University of Bern
Switzerland
AI4I - The Italian Institute of Artificial Intelligence for Industry
Italy

SELF-DRIVING CAR COMPETITION

SELF-DRIVING CAR COMPETITION

CONTEXT


AXA crashtests
AXA.ch


AXA crashtests
AXA.ch

- Cost

- Cost
- Replicability

- Cost
- Replicability

- Cost
- Replicability

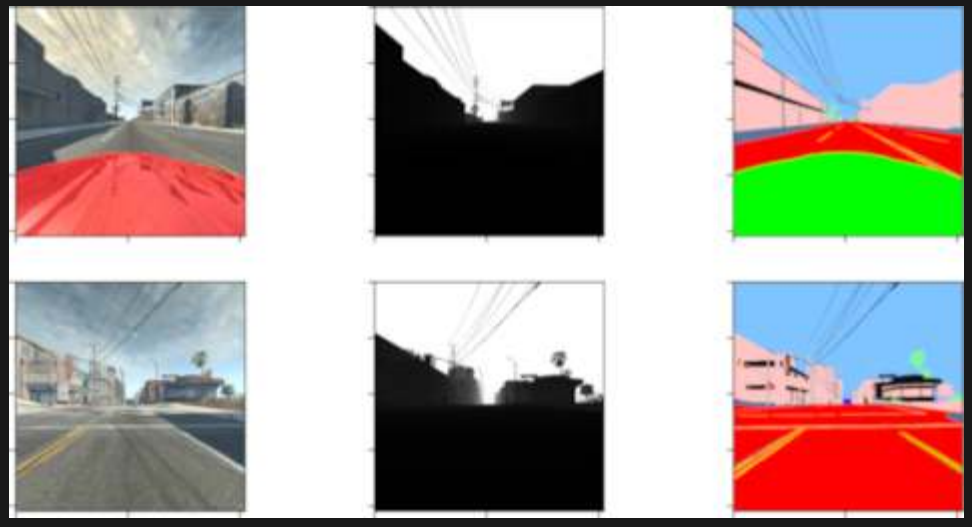
- Cost
- Replicability

- Cost
- Replicability

- Cost
- Replicability

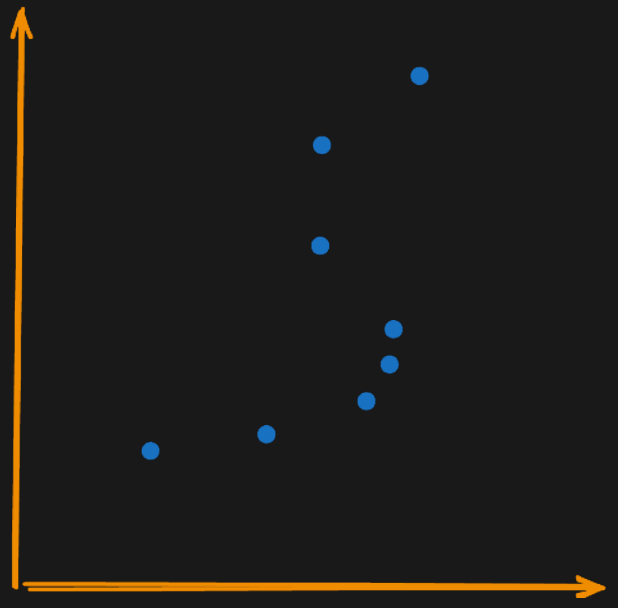
BeamNG.tech Simulator

```
(beamng) C:\dev\BeamNGpy>python
Python 3.9.6 (default, Aug 18 2021, 15:44:49) [MSC v.1916 64 bit (AMD64)]
:: Anaconda, Inc. on win32
Type "help", "copyright", "credits" or "license" for more information.
>>> from beamngpy import BeamNGpy, Scenario, Vehicle
>>> beamng = BeamNGpy('localhost', 64256)
>>> beamng.open()
<beamngpy.beamng.beamng.BeamNGpy object at 0x000001AA501AB1F0>
>>> scenario = Scenario('west_coast_usa', 'demo')
>>> vehicle = Vehicle('demo_vehicle', model='etk800',
...                   color='Green', license='PYTHON')
>>> scenario.add_vehicle(vehicle, pos=(-717, 101, 118.5),
...                      rot_quat=(0, 0, 0.92, -0.38))
>>> scenario.make(beamng)
>>> beamng.scenario.load(scenario)
>>> beamng.scenario.start()
>>> vehicle.control(throttle=0.075)
>>> vehicle.control(steering=0.5)
```

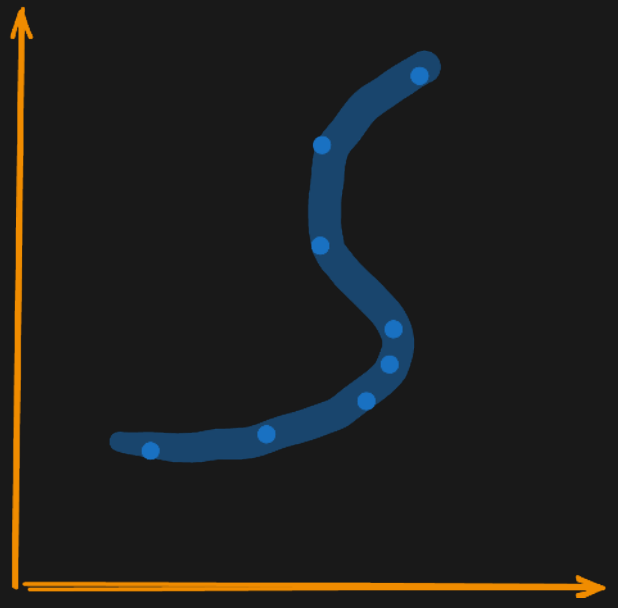


HOW IS A TEST DEFINED?

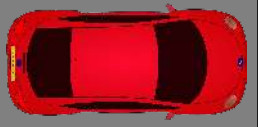
HOW IS A TEST DEFINED?



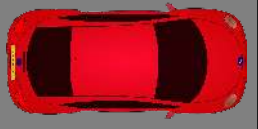
HOW IS A TEST DEFINED?

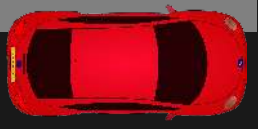


WHEN IS A TEST FAILING OR PASSING?

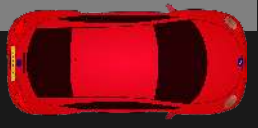


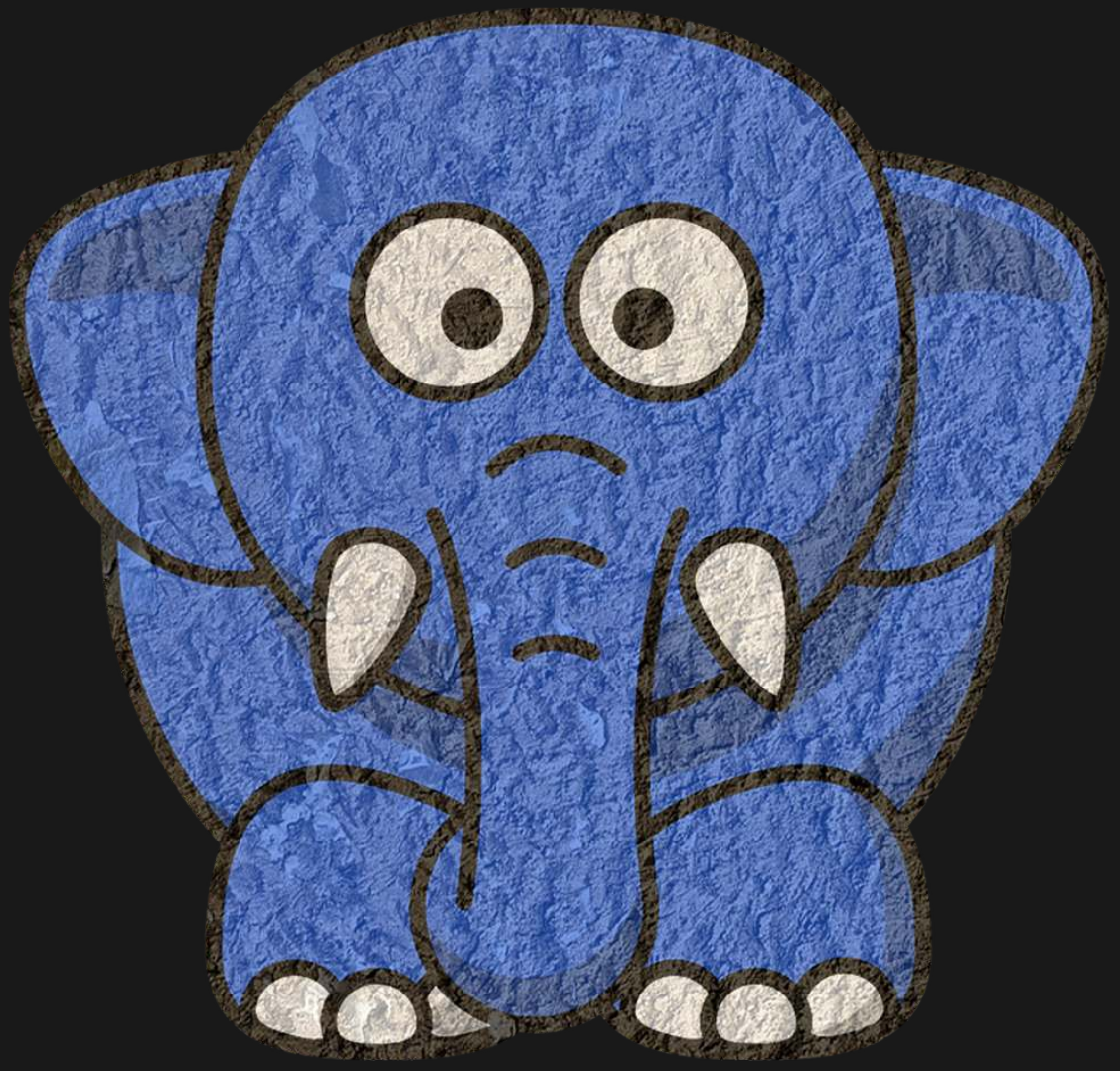
PASSED



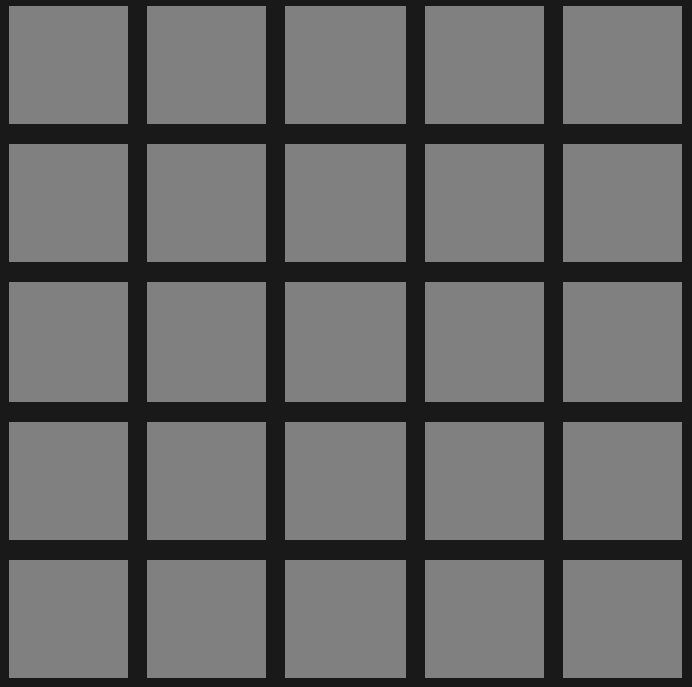


FAILED

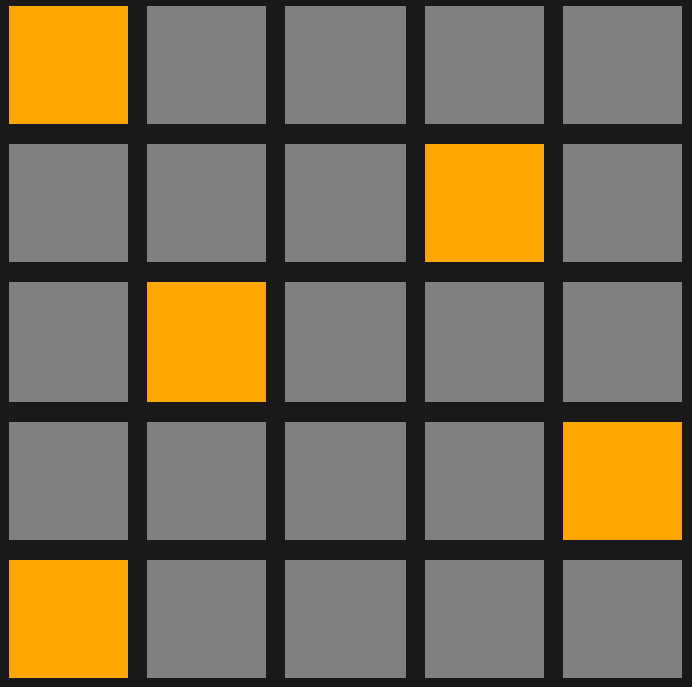




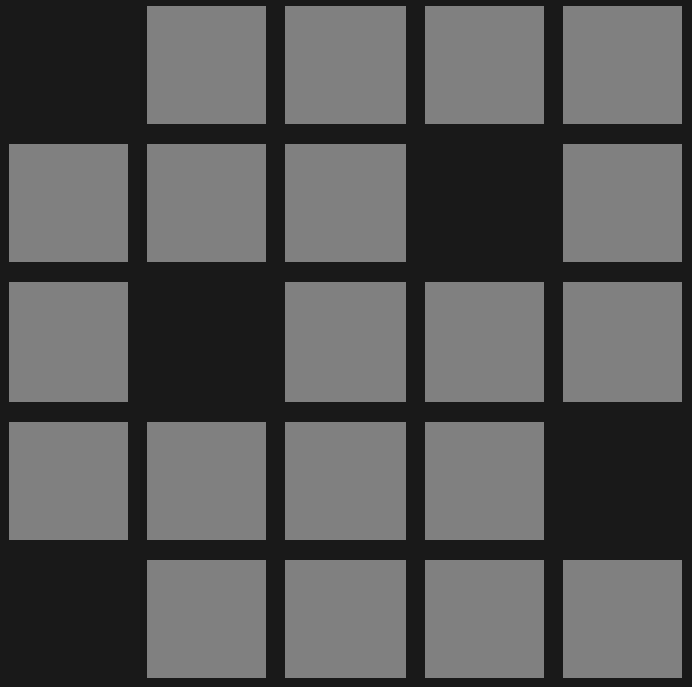
REGRESSION TESTING



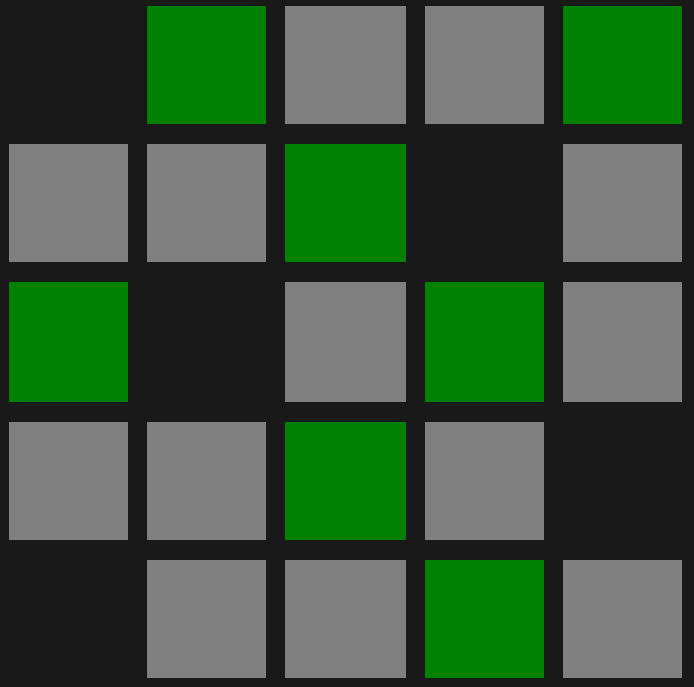
REGRESSION TESTING



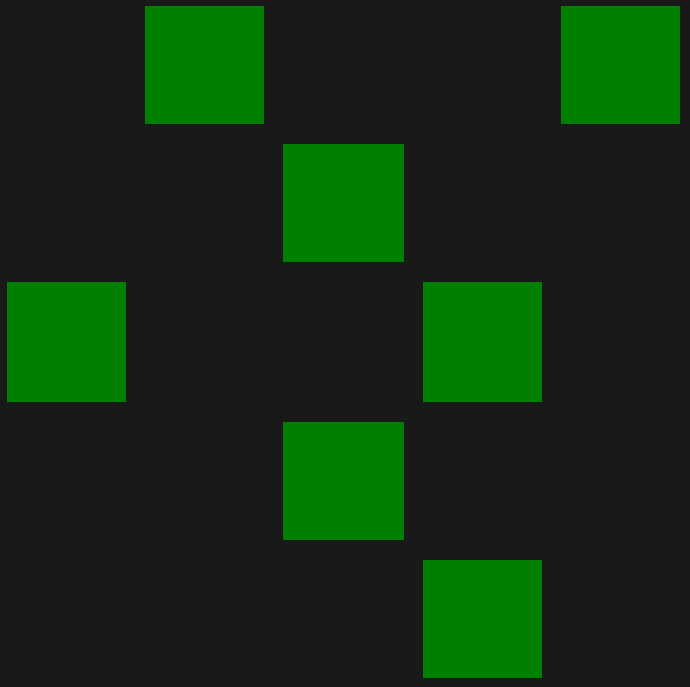
REGRESSION TESTING



REGRESSION TESTING



REGRESSION TESTING



Last year's competition:

Last year's competition:

What tests should be selected?

Last year's competition:

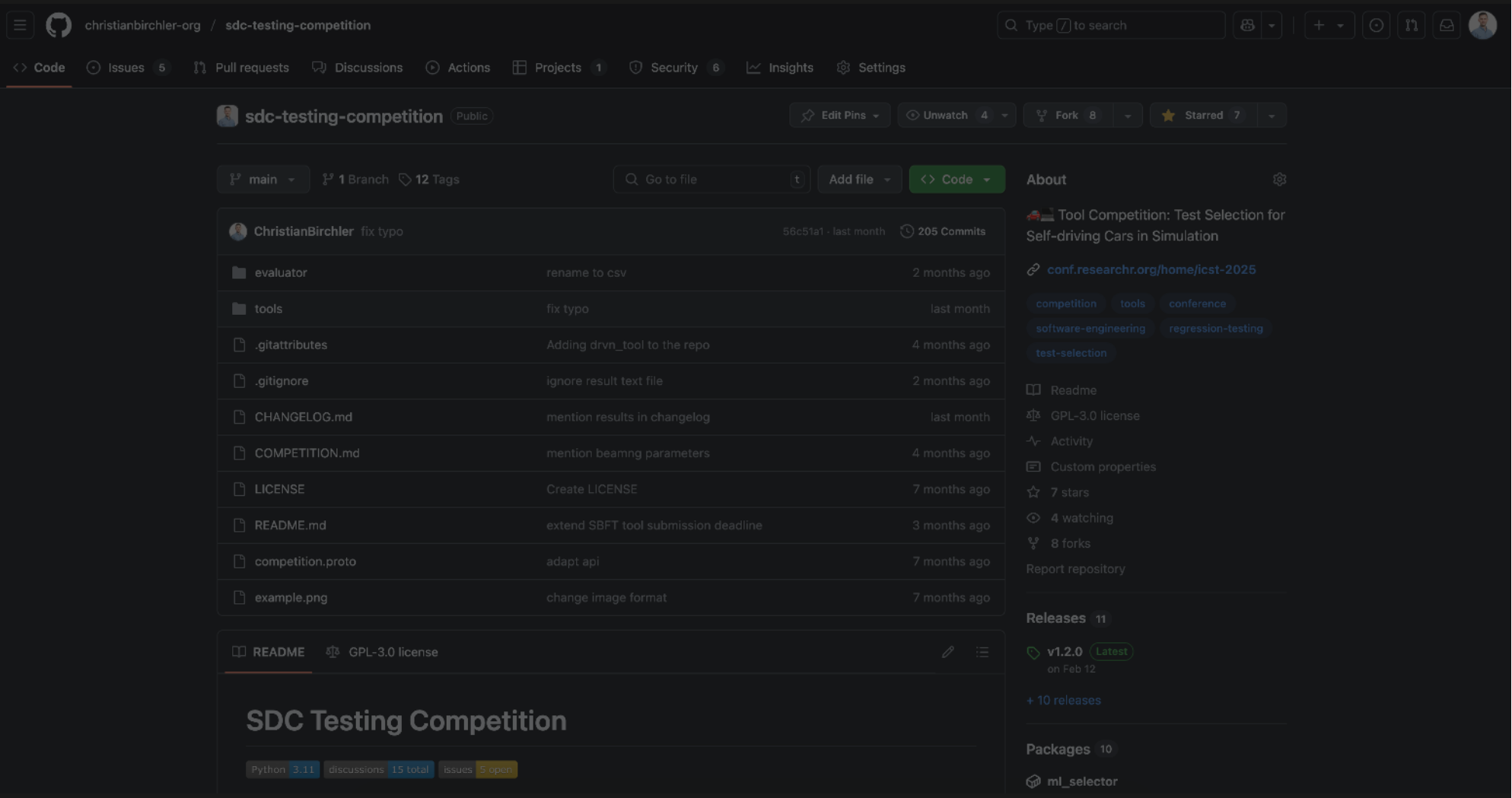
What tests should be selected?

How should tests be selected?

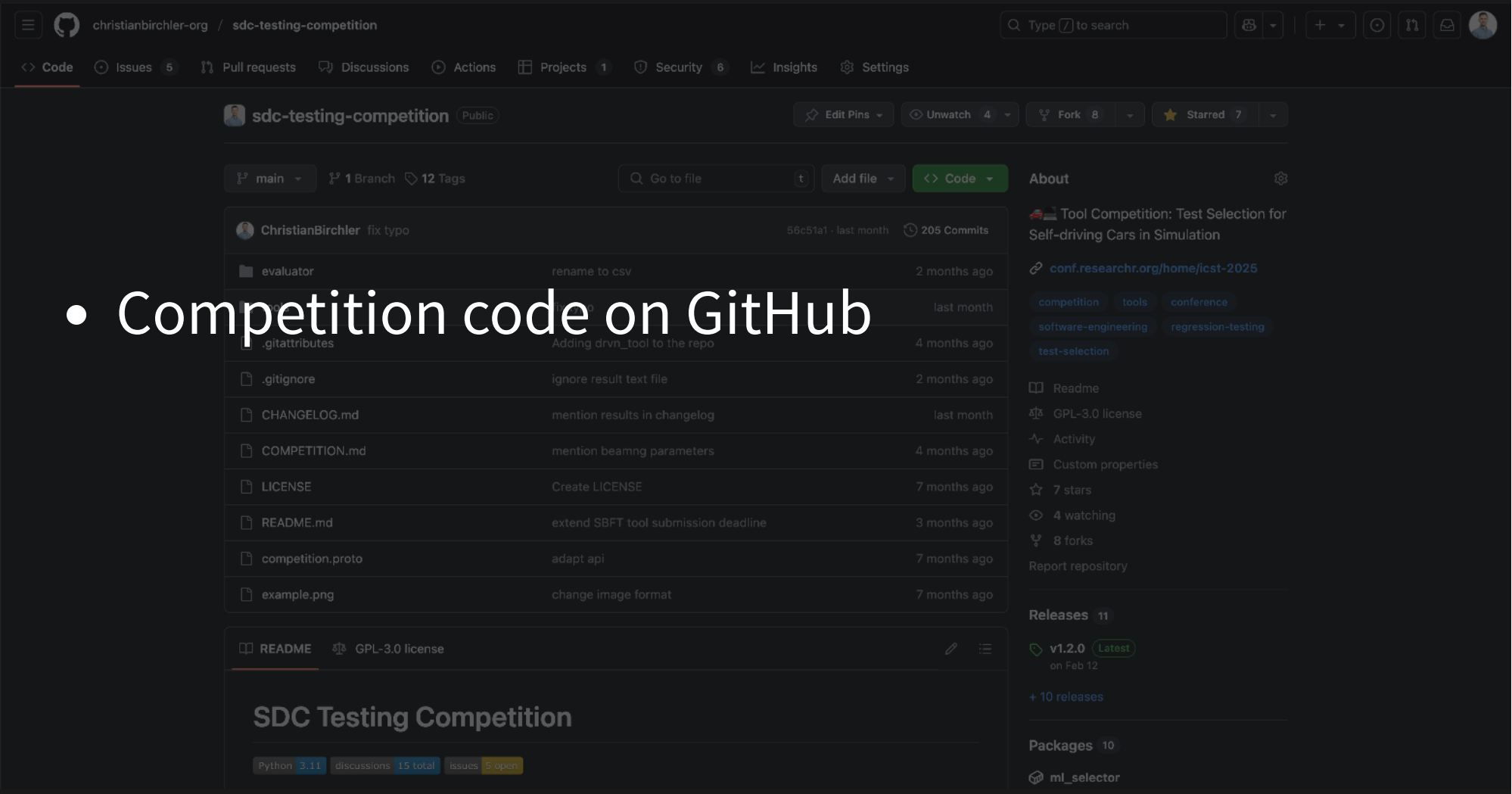
This year:

This year:

How should tests be prioritized?


christianbirchler-org / sdc-testing-competition
Type / to search
Code
Issues 5
Pull requests
Discussions
Actions
Projects 1
Security 6
Insights
Settings
sdc-testing-competition
Public
Edit Pins
Unwatch 4
Fork 8
Starred 7
main
1 Branch
12 Tags
Go to file
t
Add file
Code
ChristianBirchler fix typo
56c51a1 · last month
205 Commits
evaluator rename to csv 2 months ago
tools fix typo last month
.gitattributes Adding drvn_tool to the repo 4 months ago
.gitignore ignore result text file 2 months ago
CHANGELOG.md mention results in changelog last month
COMPETITION.md mention beamng parameters 4 months ago
LICENSE Create LICENSE 7 months ago
README.md extend SBFT tool submission deadline 3 months ago
competition.proto adapt api 7 months ago
example.png change image format 7 months ago
README
GPL-3.0 license
SDC Testing Competition
Python 3.11
discussions 15 total
issues 5 open
About
Tool Competition: Test Selection for Self-driving Cars in Simulation
conf.researchr.org/home/icst-2025
competition
tools
conference
software-engineering
regression-testing
test-selection
Readme
GPL-3.0 license
Activity
Custom properties
7 stars
4 watching
8 forks
Report repository
Releases 11
v1.2.0 Latest
on Feb 12
+ 10 releases
Packages 10
ml_selector

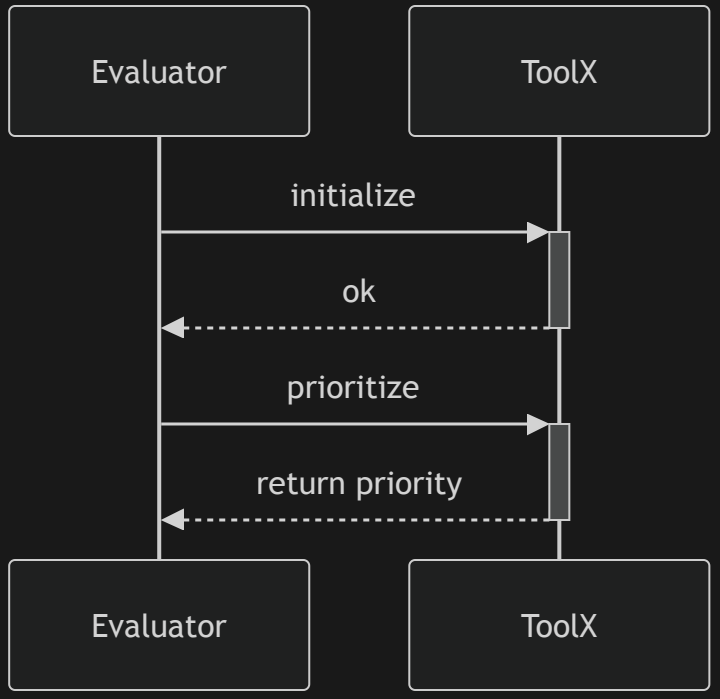
- Competition code on GitHub

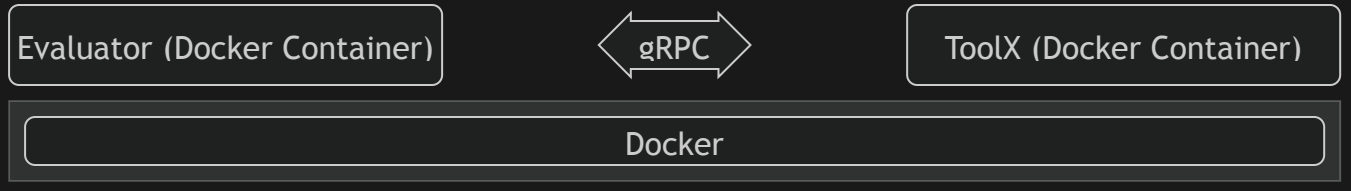


- Competition code on GitHub
- Troubleshooting with Issues and Discussion Forum

- Competition code on GitHub
- Troubleshooting with Issues and Discussion Forum
- Docker Images of all Tools!

INFRASTRUCTURE





PROTOCOL BUFFERS

```
syntax = "proto3";

service CompetitionTool {
  rpc Name(Empty) returns (NameReply) {}

  rpc Initialize (stream Oracle) returns (InitializationReply) {}

// bidirectional streaming for high flexibility
  rpc Prioritize (stream SDCTestCase) returns (stream PrioritizationReply) {}
}

message Empty {}

message NameReply {
  string name = 1;
}

message Oracle {
  SDCTestCase testCase = 1;
  bool hasFailed = 2;
}

message SDCTestCase {
  string testId = 1;
  repeated RoadPoint roadPoints = 2;
}

message RoadPoint {
  int64 sequenceNumber = 1;
  float x = 2;
  float y = 3;
}
```

```
syntax = "proto3";

service CompetitionTool {
  rpc Name(Empty) returns (NameReply) {}
  rpc Initialize (stream Oracle) returns (InitializationReply) {}
  rpc Prioritize (stream SDCTestCase) returns (stream PrioritizationRep
}
...
```

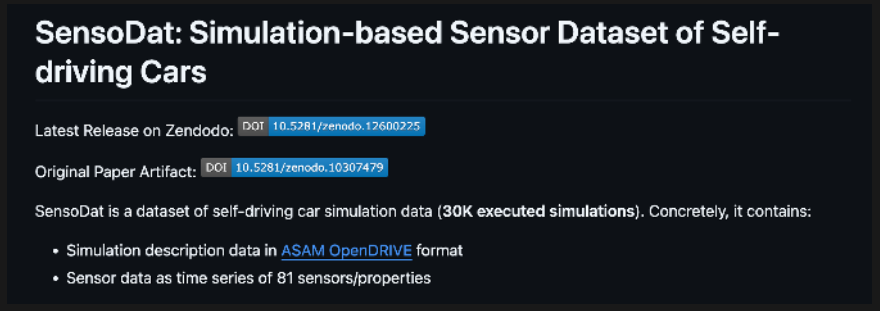
What are the evaluation metrics?

```python
@dataclass
class EvaluationReport:
    """ This class holds evaluation metrics of a tool."""

    test_suite_cnt: int
    benchmark: str
    time_to_initialize: float
    time_to_prioritize_tests: float
    tool_name: str
    time_to_first_fault: float | None
    time_to_last_fault: float | None
    apfd: float
    apfdc: float
```

BENCHMARK

SensoDat: Simulation-based Sensor Dataset of Self-driving Cars



SensoDat: Simulation-based Sensor Dataset of Self-driving Cars

Latest Release on Zendodo: DOI 10.5281/zenodo.12600225

Original Paper Artifact: DOI 10.5281/zenodo.10307479

SensoDat is a dataset of self-driving car simulation data (**30K executed simulations**). Concretely, it contains:

- Simulation description data in ASAM OpenDRIVE format
- Sensor data as time series of 81 sensors/properties

- 32,580 Test Cases
- Generated by three test generators

EXPERIMENTS

The experiments are conducted on a virtual machine (VM) with 16GB of RAM, eight virtual CPUs.

Results of SBFT Competition

Tool	Statistic	apfd	apfdc	t_prioritize_tests	t_first_fault	t_last_fault
ITEP4SDC	max	0.84	0.92	0.80	87.90	3,706.99
	mean	**0.79**	**0.83**	0.78	**54.77**	**3,016.25**
	std	0.03	0.05	0.02	28.31	553.80
	min	0.74	0.76	0.76	22.69	2,451.42
Random Baseline	max	0.62	0.67	0.52	668.30	6,613.83
	mean	0.50	0.52	**0.46**	145.75	5,781.33
	std	0.05	0.06	0.03	132.02	349.89
	min	0.38	0.39	0.37	6.65	4,898.84


WIN

ITEP4SDC

Ali ihsan Güllü, Faiz Ali Shah, Dietmar Pfahl
University of Tartu, Estonia

THANK YOU ALL!

THANK YOU ALL!

Any feedback and/or ideas are welcome for future editions.

github.com/christianbirchler-org/sdc-testing-competition