AI-guided Symbolic Execution

Mentor: Semyon Grigorev

Maxim Nigmatulin

Saint-Petersburg Research Institute

August 30, 2023

Path Selector For Symbolic Execution

Symbolic execution is a software analysis technique for understanding what data causes each part of a program to execute

Challenges:

- The problem of choosing the best path in the execution graph is an undecidable
- Path explosion problem: a symbolic executor can fork on each branch, causing the total number of states to explode

Project goal: implement AI-guided path selector for Universal Symbolic Virtual Machine (USVM)

Existing Solutions

Solutions	Disadvantages
Classic approaches (fuzzing, path merging, path prioritization, interleaved SE)	Each method is good only at specific task
Artificial Intelligence based (RL-guided SE[1], SyML[2], LEARCH[3])	Lack of flexibility

[1] Jie Wu, Chengyu Zhang, and Geguang Pu. “Reinforcement Learning Guided Symbolic Execution”

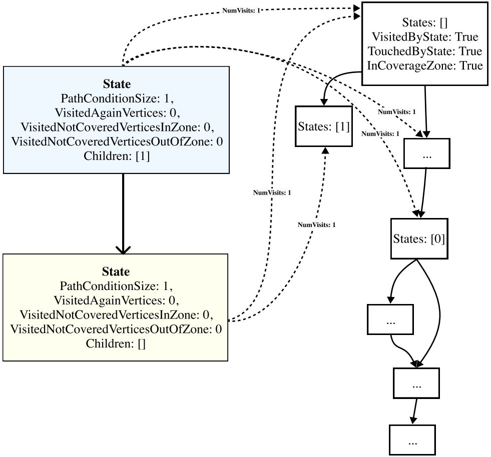
[2] Nicola Ruaro et al. “SyML: Guiding Symbolic Execution Toward Vulnerable States Through Pattern Learning”

[3] Jingxuan He et al. “Learning to Explore Paths for Symbolic Execution”

Proposed Approach

Graph Neural Network + Machine Learning

Proposed Approach: Graph Neural Network



State
PathConditionSize: 1,
VisitedAgainVertices: 0,
VisitedNotCoveredVerticesInZone: 0,
VisitedNotCoveredVerticesOutOfZone: 0
Children: [1]

State
PathConditionSize: 1,
VisitedAgainVertices: 0,
VisitedNotCoveredVerticesInZone: 0,
VisitedNotCoveredVerticesOutOfZone: 0
Children: []

States: []
VisitedByState: True
TouchedByState: True
InCoverageZone: True

States: [1]

States: [0]

Figure: Control Flow Graph

Proposed Approach: Genetic & Supervised Learning

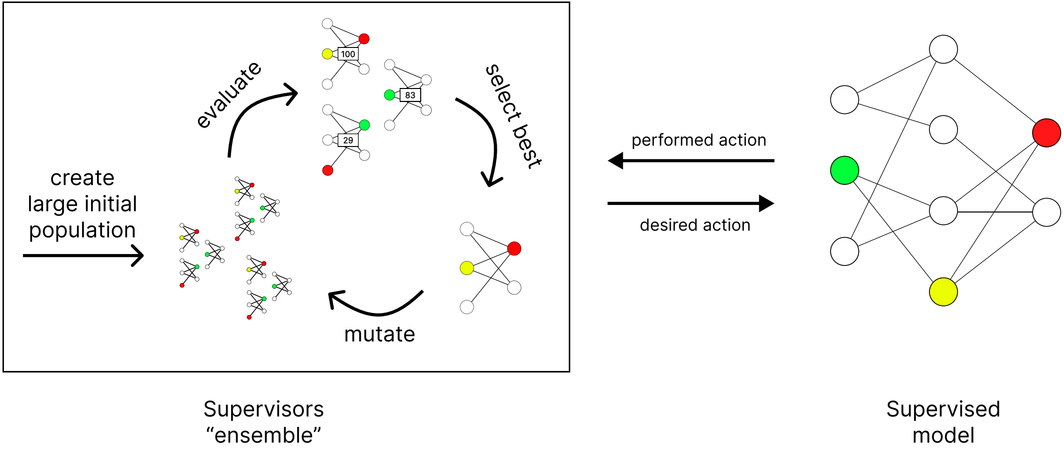
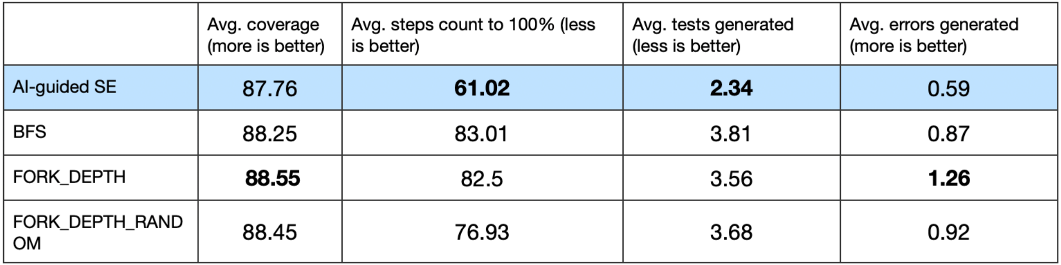


Figure: Learning pipeline

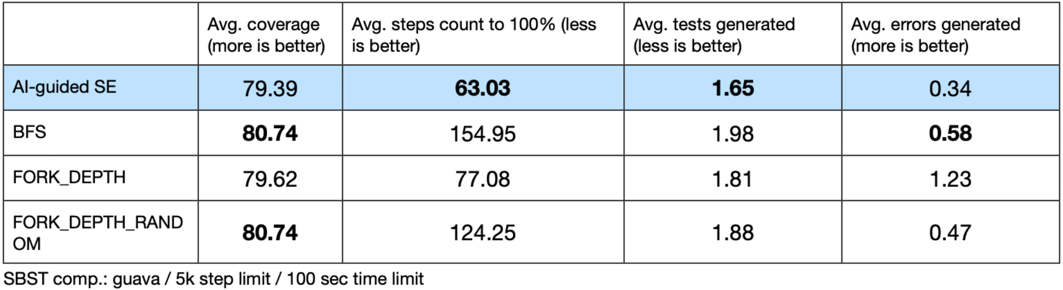
Benchmark Results: Simple Tests



	Avg. coverage (more is better)	Avg. steps count to 100% (less is better)	Avg. tests generated (less is better)	Avg. errors generated (more is better)
AI-guided SE	87.76	**61.02**	**2.34**	0.59
BFS	88.25	83.01	3.81	0.87
FORK_DEPTH	**88.55**	82.5	3.56	**1.26**
FORK_DEPTH_RANDOM	88.45	76.93	3.68	0.92

UTBot test suite / 5k step limit / 100 sec time limit

Benchmark Results: Complex Tests



	Avg. coverage (more is better)	Avg. steps count to 100% (less is better)	Avg. tests generated (less is better)	Avg. errors generated (more is better)
AI-guided SE	79.39	**63.03**	**1.65**	0.34
BFS	**80.74**	154.95	1.98	**0.58**
FORK_DEPTH	79.62	77.08	1.81	1.23
FORK_DEPTH_RANDOM	**80.74**	124.25	1.88	0.47

SBST comp.: guava / 5k step limit / 100 sec time limit

Future Work

- Experiment with Genetic Learning paremeters
- Select best performing model architecture
- Add new program graphs to dataset
- PR to USVM

Thank you for your attention!

Max Nigmatulin
email: mvnigma@gmail.com
tg: @mvnigma

link