Title

Compositional Falsification of Cyber-Physical Systems with Machine Learning Components.

Authors

Dreossi, Tommaso
Donzé, Alexandre
Seshia, Sanjit A

Publication Date

2019

DOI

10.1007/s10817-018-09509-5

Peer reviewed

Compositional Falsification of Cyber-Physical Systems with Machine Learning Components

Tommaso Dreossi, Alexandre Donzé, and Sanjit A. Seshia. **Compositional Falsification of Cyber-Physical Systems with Machine Learning Components**. *Journal of Automated Reasoning*, 63(4):1031–1053, 2019.

Download

[pdf]

Abstract

Cyber-physical systems (CPS), such as automotive systems, are starting to include sophisticated machine learning (ML) components. Their correctness, therefore, depends on properties of the inner ML modules. While learning algorithms aim to generalize from examples, they are only as good as the examples provided, and recent efforts have shown that they can produce inconsistent output under small adversarial perturbations. This raises the question: can the output from learning components lead to a failure of the entire CPS? In this work, we address this question by formulating it as a problem of falsifying signal temporal logic specifications for CPS with ML components. We propose a compositional falsification framework where a temporal logic falsifier and a machine learning analyzer cooperate with the aim of finding falsifying executions of the considered model. The efficacy of the proposed technique is shown on an automatic emergency braking system model with a perception component based on deep neural networks.

BibTeX

```
@article{dreossi-jar19,
  author    = {Tommaso Dreossi and
               Alexandre Donz{\'{e}} and
               Sanjit A. Seshia},
  title     = {Compositional Falsification of Cyber-Physical Systems with Machine
               Learning Components},
  journal   = {Journal of Automated Reasoning},
  volume    = {63},
  number    = {4},
  pages     = {1031--1053},
  year      = {2019},
  abstract  = {Cyber-physical systems (CPS), such as automotive systems, are starting to include sophisticated machine learning (ML) comp
}
```