EnnCore: Safety Verification of Deep Reinforcement Learning

Yi Dong, Xingyu Zhao, and Xiaowei Huang

Department of Computer Science, University of Liverpool
Objective of this special session

- Black-box
- White-box
Objective of this special session: DRL verification vs CNN verification

DRL verification vs CNN verification

CNN Verification

DRL Verification
Black-box Verification

DRL Training → Abstraction

DRL Policy

Assessment

Properties:
- Robustness
- Resilience
- Detection
- Recovery
- Safety
- ...

Model Checking

DTMC Failure Model
Black-box Verification

Technical Problem:
• How to synthesize the Markov model (DTMC, MDP)?
• How to evaluate the dependability of the DRL policy?
• How do properties perform in DRL algorithms?

Technical Solution:
• Defined different properties
• Construct DTMC model
• Probabilistic model checking

Technical Observation:
• The dependability analysis are insensitive to the sample size
• trade-offs between different properties
Safety Properties

- Safety
- Robustness
- Resilience
- Detection
- Recovery
- ...
Black-box Verification
White-box Verification

Actor Networks

Environment Models

[Images and icons related to the concepts mentioned in the text]
Two-level verification

Low-level: for a given actor network, calculate reachable set of actions.

High-level: similar to black-box verification methods.

Connection: Do high-level verification based on low-level verification results.
White-box Verification

Temporal path

Sensor Signal
Real Position Range
Next Step Range

Range

Steps

0 1 2 3 4 ...

n
White-box Verification

A

B

C

D

Real Path

Observation Path

Safe

Unsafe

Real Path

Sensor Path

Polar Path

One-step Safe

Unsafe

Safety Threshold
White-box Verification
Challenges

• DTMC models
  – build accurate DTMC with less samples

• Verification of DRL Models
  – A key problem remains on the **scalability and time-efficiency**: trade-off between **soundness** and **completeness**

• Unknown Environment
  – Verify of Neural network model working in an **unknown environment**
<table>
<thead>
<tr>
<th>Time (UTC)</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>14:00-14:10</td>
<td>Welcome, overview, Lucas Cordeiro (University of Manchester, UK)</td>
</tr>
<tr>
<td>14:10-14:25</td>
<td>Verifying Quantized Neural Networks using SMT-Based Model Checking, Edoardo Manino (University of Manchester, UK)</td>
</tr>
<tr>
<td>14:25-14:40</td>
<td>Explainability and Inference Controls, André Freitas (University of Manchester UK &amp; Idiap Research Institute, Switzerland)</td>
</tr>
<tr>
<td>14:40-14:55</td>
<td>Safety Verification of Deep Reinforcement Learning, Yi Dong (University of Liverpool, UK)</td>
</tr>
<tr>
<td>14:55-15:10</td>
<td>Privacy Friendly Energy Consumption Prediction: Real Case-Studies, Mustafa A. Mustafa (University of Manchester, UK / KU Leuven, Belgium)</td>
</tr>
<tr>
<td>15:10-15:30</td>
<td>Closed-loop Safety of Bayesian Neural Networks and Stochastic Control Systems, Mathias Lechner, IST Austria</td>
</tr>
</tbody>
</table>