Testing for Fault Diversity in Reinforcement Learning

Paper doi: https://doi.org/10.1145/3644032.3644458

Team members: Ciama Andreea Elena, Abrudan Rebeca Rafaela, Crișan Dragoș

The paper "Testing for Fault Diversity in Reinforcement Learning" by Quentin Mazouni et al. aims to address a significant gap in current software validation practices for reinforcement learning (RL) models. While traditional methods emphasize maximizing the number of detected faults, they often overlook the importance of fault diversity. Diverse faults can provide more informative insights into the model's weaknesses, leading to better improvements and understanding of the RL policy. The study explores the use of Quality Diversity (QD) optimization as a novel approach to enhance policy testing by not only detecting faults but ensuring these faults are diverse and behaviorally informative.

The methodology involves reformulating policy testing as a Quality Diversity optimization problem, leveraging the capabilities of QD to find both high-quality and diverse solutions. The researchers implemented two QD optimizers, MAP-Elites and Novelty Search, to test their effectiveness in detecting diverse faults in RL models. They compared these methods against MDPFuzz, a state-of-the-art policy testing framework. Experiments were conducted on three benchmark environments: Lunar Lander, Bipedal Walker, and Taxi. The evaluation metrics included the number of distinct faults detected, behavior space coverage (the number of different behaviors tested), and the diversity of final states (how varied the failure states were). These metrics provided a comprehensive assessment of both the efficiency and the diversity of the faults detected by each method.

The results revealed that QD optimization, particularly with the MAP-Elites algorithm, significantly outperformed traditional random testing and MDPFuzz in terms of both the number and diversity of faults detected. MAP-Elites consistently found more faults and exhibited greater stability and efficiency across different environments. Novelty Search also showed strong performance in some cases but was more variable and sensitive to initial conditions. The study found that while MDPFuzz could compete in certain environments, it generally did not match the fault detection efficiency or diversity of the QD-based methods. The experiments highlighted the importance of behavior space definitions, showing that the choice of descriptors can substantially impact the performance of QD-based testing.

The implications of this research are significant for both academia and industry. For researchers, the study introduces Quality Diversity optimization as a powerful tool for enhancing the fault detection process in RL models, encouraging further exploration of adaptive behavior space definitions and methods to handle stochastic environments. For practitioners, particularly those working with RL in safety-critical applications like autonomous driving or robotics, QD-based policy testing offers a method to build more reliable and trustworthy models by ensuring that detected faults are both numerous and diverse. This approach can lead to more robust RL policies, ultimately improving the safety and effectiveness of real-world applications. The study's findings suggest a shift in focus from merely maximizing fault counts to achieving a more nuanced understanding of model weaknesses through diverse fault detection.