World Academy of Science, Engineering and Technology
International Journal of Computer and Information Engineering
Vol:16, No:04, 2022

Safety Validation of Black-Box Autonomous Systems: A Multi-Fidelity Reinforcement Learning Approach

Authors : Jared Beard, Ali Baheri

Abstract : As autonomous systems become more prominent in society, ensuring their safe application becomes increasingly important. This is clearly demonstrated with autonomous cars traveling through a crowded city or robots traversing a warehouse with heavy equipment. Human environments can be complex, having high dimensional state and action spaces. This gives rise to two problems. One being that analytic solutions may not be possible. The other is that in simulation based approaches, searching the entirety of the problem space could be computationally intractable, ruling out formal methods. To overcome this, approximate solutions may seek to find failures or estimate their likelihood of occurrence. One such approach is adaptive stress testing (AST) which uses reinforcement learning to induce failures in the system. The premise of which is that a learned model can be used to help find new failure scenarios, making better use of simulations. In spite of these failures AST fails to find particularly sparse failures and can be inclined to find similar solutions to those found previously. To help overcome this, multi-fidelity learning can be used to alleviate this overuse of information. That is, information in lower fidelity can simulations can be used to build up samples less expensively, and more effectively cover the solution space to find a broader set of failures. Recent work in multi-fidelity learning has passed information bidirectionally using “knows what it knows” (KWIK) reinforcement learners to minimize the number of samples in high fidelity simulators (thereby reducing computation time and load). The contribution of this work, then, is development of the bidirectional multi-fidelity AST framework. Such an algorithm, uses multi-fidelity KWIK learners in an adversarial context to find failure modes. Thus far, a KWIK learner has been used to train an adversary in a grid world to prevent an agent from reaching its goal; thus demonstrating the utility of KWIK learners in an AST framework. The next step is implementation of the bidirectional multi-fidelity AST framework described. Testing will be conducted in a grid world containing an agent attempting to reach a goal position and adversary tasked with intercepting the agent as demonstrated previously. Fidelities will be modified by adjusting the size of a time-step, with higher-fidelity effectively allowing for more responsive closed loop feedback. Results will compare the single KWIK AST learner with the multi-fidelity algorithm with respect to number of samples, distinct failure modes found, and relative effect of learning after a number of trials.

Keywords : multi-fidelity reinforcement learning, multi-fidelity simulation, safety validation, falsification
Conference Title : ICMLA 2022 : International Conference on Machine Learning and Applications
Conference Location : Boston, United States
Conference Dates : April 21-22, 2022