Evaluation of Autonomous Vehicle Policies Using Adaptive Search

Peter Du and Katherine Driggs-Campbell

***Abstract*— Discovering potential failure conditions of an autonomous vehicle policy is an important step in validating and evaluating the performance of any self driving system. The use of physical vehicles to find failure cases by logging miles in the real world may be insufficient to uncover a large region of the failure space due to time and cost constraints. Using high fidelity simulation can greatly reduce the costs associated with running a fleet of real vehicles. However, even with simulation, the space of scenarios is too large for naive sampling approaches to explore. We believe that adaptive search techniques can address this problem by efficiently exploring the failure space and uncovering the most important failure scenarios.**

I. Introduction

Autonomous vehicles (AVs) are rapidly being developed with complex self driving algorithms that aim to replace or minimize the need for a human driver. As these autonomous systems enter the public domain and become tangible technologies that other road users interact with, ensuring they operate in a safe and reliable manner is of great importance. However, validation of self driving algorithms remains a difficult task for system designers due to the broad range of traffic situations and interactions a vehicle may face in the real world. Physical vehicle-level testing is often employed by AV designers to validate their systems and gain confidence in a self driving agent's ability to operate reliably. However, the space of failures are often comprised of rare events which require significant time and monetary resources to properly explore using real world testing [1]. Simulation can be used to complement physical vehicle-level testing by allowing AV designers to run a greater number vehicles for a prolonged period of time. Despite this, the space of scenarios an AV may face remains too large for naive sampling methods to adequately explore [2].

We believe that adaptive failure generation techniques can address this problem and aid in the evaluation of self driving algorithms. These techniques allow for the generation of safety critical scenarios using simulation that exposes a self driving agent's vulnerabilities. Deep generative models have been proposed to produce safety critical driving scenarios that can be used to evaluate the performance of an autonomous policy [3], [4]. These methods aim to learn a parametric distribution of safety critical states by optimizing the likelihood of collected vehicle traces. Meanwhile, various adaptive sampling approaches aim to efficiently search for critical scenarios that may lead to AV failure by using a simulation of the system-under-test (SUT) [5]–[7].

P. Du and K. Driggs-Campbell are with the Department of Electrical and Computer Engineering at the University of Illinois at Urbana-Champaign. email: {peterdu2,krdc}@illinois.edu

Adaptive Stress Testing (AST) is one such adaptive searching method that is used to evaluate the behaviour of autonomous policies by identifying high risk failure scenarios. Failure search of the SUT is formulated as a Markov decision process which is then solved by training a reinforcement learning (RL) policy through simulation [8], [9].

II. Adaptive Stress Testing

AST performs adaptive search of failures through the use of a simulator $\mathscr{S}$, reinforcement learning solver $\mathcal{S}$, and RL reward function $\mathcal{R}$. The framework supports black-box simulators that may (partially) hide its internal state. The use of an RL paradigm allows the solver to learn a policy for failure generation without full observability of the internal state of $\mathscr{S}$ as long as the simulator is able to indicate that the SUT has reached a failure state.

The solver $\mathcal{S}$ is trained to output a policy generating failure trajectories by iteratively exploring the environment and collecting reward signals. At each training step, $\mathcal{S}$ generates an environment action which is used to step the simulation forward in time. Environment actions contain commands that affect the other agents and conditions within the simulator. For example, if the SUT is an autonomous vehicle driving down a road with pedestrians and other vehicles, the environment actions may include: the acceleration/braking of the other vehicles, the motion of pedestrians, the conditions of the road (wet,dry), etc. Once the simulation is updated with the new environment action, a reward signal is generated by $\mathcal{R}$ and used by $\mathcal{S}$ to update its policy.

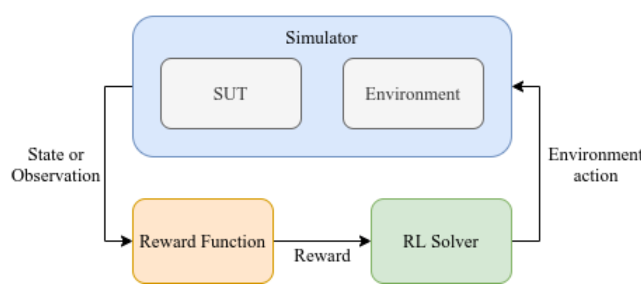

Fig. 1: Adaptive Stress Testing architecture.

III. AST for Evaluation of AV Policies

The use of a reinforcement learning agent to perform adaptive failure search enables one to encode domain specific insights into the reward function of AST and adapt the method for different AV policies. Reward augmentation allows for the comparison of failure trajectories of different policies using a standard metric. For example, the Responsibility-Sensitive Safety (RSS) is a set of rules designed to capture human driving intuitions [10]. RSS can be incorporated into the failure search and used to assess the reward value of trajectories by weighting the reward function $\mathcal{R}$ with the proportion of timesteps violating the ruleset [11]. Other metrics incorporated in the reward function may include: the likelihood of an environment action, the risk/criticality of the current simulator state, or heuristic measures to guide failure search.

IV. Conclusion

Simulation can be used to find critical scenarios that both inform autonomous vehicle designers of system vulnerabilities and allow for the evaluation of AV performance under rare and dangerous situations. Due to the large space of possible conditions a vehicle may face, efficient generation of critical scenarios in simulation is an important area to investigate. Adaptive searching techniques such as AST allows for AV designers to uncover failure trajectories of their systems in a far more targeted approach than naive random sampling. The reward metrics for which failures are compared against can be tuned specifically for the task of AV validation and thus allows for flexibility in generating a large class of failure scenarios.

References

[1] P. Koopman and M. Wagner, "Challenges in Autonomous Vehicle Testing and Validation," in *SAE World Congress*, 2016.

[2] N. Kalra and S. M. Paddock, "Driving to Safety: How Many Miles of Driving Would It Take to Demonstrate Autonomous Vehicle Reliability?" RAND Corporation, 2016.

[3] W. Ding, W. Wang, and D. Zhao, "A multi-vehicle trajectories generator to simulate vehicle-to-vehicle encountering scenarios," in *International Conference on Robotics and Automation (ICRA)*, 2019.

[4] W. Ding, B. Chen, B. Li, K. J. Eun, and D. Zhao, "Multimodal safety-critical scenarios generation for decision-making algorithms evaluation," 2020.

[5] G. E. Mullins, P. G. Stankiewicz, and S. K. Gupta, "Automated generation of diverse and challenging scenarios for test and evaluation of autonomous vehicles," in *IEEE International Conference on Robotics and Automation (ICRA)*, 2017.

[6] M. O'Kelly, A. Sinha, H. Namkoong, J. Duchi, and R. Tedrake, "Scalable end-to-end autonomous vehicle testing via rare-event simulation," in *Conference on Neural Information Processing Systems (NeurIPS)*, 2018.

[7] C. E. Tuncali, G. Fainekos, H. Ito, and J. Kapinski, "Simulation-based adversarial test generation for autonomous vehicles with machine learning components," in *IEEE Intelligent Vehicles Symposium (IV)*, 2018.

[8] R. Lee, M. J. Kochenderfer, O. J. Mengshoel, G. P. Brat, and M. P. Owen, "Adaptive stress testing of airborne collision avoidance systems," in *Digital Avionics Systems Conference (DASC)*, 2015.

[9] M. Koren, S. Alsaif, R. Lee, and M. J. Kochenderfer, "Adaptive Stress Testing for Autonomous Vehicles," in *IEEE Intelligent Vehicles Symposium (IV)*, 2018.

[10] S. Shalev-Shwartz, S. Shammah, and A. Shashua, "On a Formal Model of Safe and Scalable Self-Driving Cars," *ArXiv*, no. 1708.06374, 2017.

[11] A. Corso, P. Du, K. Driggs-Campbell, and M. J. Kochenderfer, "Adaptive stress testing with reward augmentation for autonomous vehicle validation," in *IEEE Intelligent Transportation Systems Conference (ITSC)*, 2019.