Overview. Assured autonomy is essential for safety-critical systems in high-stakes applications, such as autonomous vehicles, medical devices, and industrial control systems. This CAREER project aims to establish the PI as a leader in the field of assured autonomy by advancing both design-time and runtime assurance. The proposal focuses on revolutionizing design-time assurance by investigating joint falsification and simulation optimization, which will bridge the gap between high-fidelity and low-fidelity simulators. Additionally, it aims to elevate runtime assurance through joint learning of policy with unknown specification, empowering autonomous systems to adapt and navigate safely in the face of unknown constraints. The project integrates research and education through the development of an educational game designed to incorporate research outcomes and make the principles of assured autonomy accessible to a broader audience. The project will also partner with the National Institute for the Deaf and Rochester Prep High School for outreach activities, promoting inclusivity and ensuring a diverse pool of future researchers in the field of AI safety.

Intellectual Merit. The intellectual merit of this proposal lies in the significant advancements it aims to make in algorithmic techniques for developing and deploying safe and reliable autonomous systems. The proposed research focuses on joint learning with simulation parameters and joint learning with specification to enhance the accuracy and efficiency of simulation-based validation and learning of safety constraints and formal specifications, ultimately improving the safety-critical autonomy of autonomous systems in real-world operation. Learning optimal fidelity settings through joint falsification with simulation parameters can address existing challenges in simulation-based validation methods, such as manual tuning of simulator parameters, limited ability to model complex system behaviors, and inadequate generalization to new and unseen scenarios. Additionally, joint learning with specification represents a largely unexplored research area, despite significant recent advancements in learning constraints or specifications from data. Defining safety constraints and requirements for autonomous systems is a complex task, further complicated by the subjective nature of safety, and the unavailability of valid constraints for real-world safety-critical applications. This research proposal aims to tackle these challenges by investigating novel algorithmic approaches to joint learning of policy and constraint, specifically targeting safety-critical applications. The potential impact of this research includes advancing algorithmic techniques for developing safe, learning-enabled systems, which will have broad societal implications. By addressing the challenges associated with joint falsification with simulation parameters and specification learning, this research has the potential to revolutionize the field of assured autonomy and create safer, more reliable autonomous systems.

Broader Impacts. The broader impacts of this research span across societal, professional, and educational domains. On a societal level, the proposed research will substantially enhance the safety and reliability of autonomous systems, which could lead to fewer accidents and increased public trust. The research has potential applications in diverse fields such as aerospace, defense, and healthcare, aiming to revolutionize safety-critical systems. By tackling the challenge of defining safety constraints and requirements for autonomous systems, this work will contribute to the safe deployment of these technologies. This will lead to societal benefits like improved transportation systems, reduced environmental impacts, and a better quality of life. Professionally, in the robotics and control community, this research aims to enable the learning of optimal simulator configurations. This would not only save computational resources but also make simulation results more interpretable for human experts, speeding up the development and testing process. Furthermore, the scalability of low-fidelity simulators will allow more efficient large-scale safety evaluations. The joint learning of policies and constraints could address existing challenges in control and robotics, improving the safety, reliability, and efficiency of real-world operations. On the educational front, this project will have a significant impact on developing future researchers in the field of assured autonomy. By creating educational game that incorporate research outcomes, the principles of assured autonomy will become more accessible, fostering interest and understanding. Outreach activities partnering with the National Technical Institute for the Deaf and Rochester Prep High School will promote inclusivity and ensure a diverse set of future researchers engage in the field. This will strengthen the connection between research and education and prepare the next generation to tackle the challenges of assured autonomy in safety-critical systems.