Research Statement: Scalable and Generalizable Value Function-based Frameworks for Safe Control and Decision Making of Autonomous Agents

Jason J. Choi jason.choi@berkeley.edu
Advisors: Koushil Sreenath, Claire J. Tomlin

February 14, 2022

Overview

In vast literature in controls, reinforcement learning, and robotics, a specifically designed scalar function is often used to verify whether an autonomous agent can achieve a given task with its actions. The utility of this function stems from how it is constructed–by incorporating the task information, the agent's dynamics, as well as the action constraint into a single scalar measure. These "certificate" functions go by different names depending on their various definitions and purposes, for instance, Lyapunov functions for stability analysis and "value functions" in the optimal control literature.

Some of the fundamental works on safety verification and control synthesis [1, 2] indicate that verifying safety is equivalent to solving a reachability problem, a specific type of an optimal control problem. Thus, in principle, once the value function of this problem is solved, one can claim that the corresponding safety control problem is deciphered. However, until these days, theories and constructive methods for these value functions are very limited to fairly simple and low-dimensional systems, wherein the discrepancy to more complicated real-world systems remains as an open research question. Moreover, in practice, there still exist many fundamental challenges. For example, the value function is prone to errors in the model that was used for its construction, or the safety specification varies in time as the agent interacts with the real-world dynamic environments, which requires the value function to be updated online.

Towards this end, my PhD research centers around pushing the boundaries of the value function-based approaches for safety control problems so that they can be used effectively to ensure safety of the real-world autonomous robot systems. From a holistic view, most of the value function-based frameworks are composed of two key steps: 1) its construction based on the task and system specifications, and 2) the derivation of safe control policies from the constructed value function (Fig. 1). My main research thrust is to verify novel ideas to tackle crucial challenges in each of the components as well as the entire pipeline itself. In the following sections, my three ongoing main projects are briefly introduced; their progress so far and their future work directions.

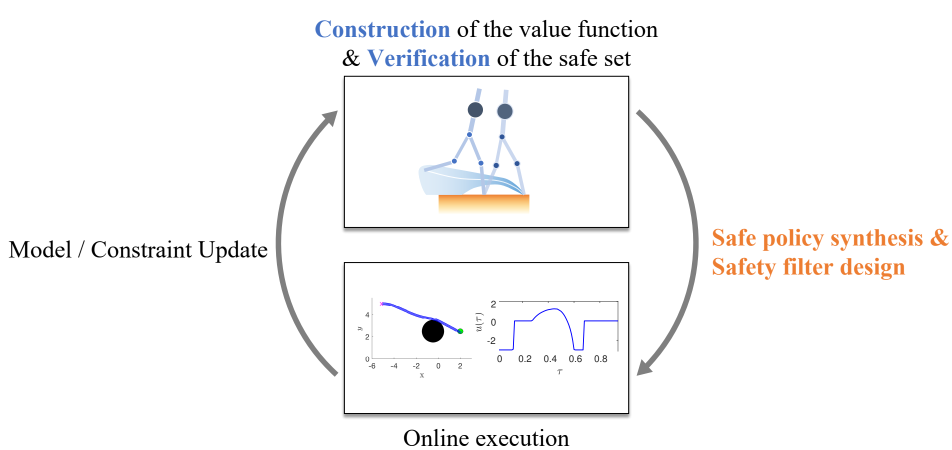

Figure 1: General pipeline of the value-function based approaches for safety control

Merging various notions of the value functions.

The main theoretical axis I am investigating is how different notions of the value functions are related to each other and how we can merge them in complementary ways. For safety control problems, three of the most popular value function concepts are Hamilton-Jacobi (HJ) Reachability [3], Control Barrier Functions (CBFs) [4], and the value functions in the safe RL literature [5, 6]. In one of my works, we brought in one of the main characteristics of the CBF into the HJ reachability formulation, and proposed a new concept of Control Barrier-Value Functions for finite-time horizon safety problems [7]. My ongoing work is extending this formulation to an infinite-horizon setup. Finally, I am also investigating how ideas in the reachability literature can be used in RL to learn safer policies, or vice versa–to learn the reachability-based value functions more efficiently with RL.

Construction of the value function for more complicated systems.

The majority of the optimal control literature and the relevant numerical methods are dedicated to continuous dynamical systems, however, real-world robot systems are often subjected to discrete evolution due to, for instance, contacts or impacts. Thus, I am investigating how the reachability formulations and numerical methods can be extended to hybrid systems. In [8], we verified that a simple value remapping technique derived from the Bellman principle can reason about discrete state jumps defined by a reset map. This technique can be used to verify regions of attractions for hybrid limit cycles as well as the corresponding stabilizing controllers that can leverage the reset maps. We found out that this can be extremely useful for designing walking robots' controllers stabilizing to their nominal gaits when they deviate far from the gaits. We tested the tabular version of this method on a simple two-link walker, and currently are extending the method to using deep neural network-based approximation for high-dimensional systems.

Addressing model uncertainties in value function-based safety filters.

Model uncertainty, which is an error of the controller designer's model from the true plant, is often the primary source of why the model-based safety guarantee fails when it is executed on the true plant. Safety filters that are based on the value functions also suffer from this problem because their safety constraints are often model-dependent. However, by observing the real-world data, we can learn how the uncertainty affects these safety constraints and adequately update them to recover safety guarantees for the true plant. In my past projects, we have applied various data-driven methods, from RL [9] to Gaussian Process regression [10, 11], to achieve this recovery. Recently, we are extending our GP-based framework to enhance its scalability to high-dimensional systems.

References

[1] J. Lygeros, "On reachability and minimum cost optimal control," *Automatica*, 2004.

[2] I. M. Mitchell, A. M. Bayen, and C. J. Tomlin, "A time-dependent Hamilton-Jacobi formulation of reachable sets for continuous dynamic games," *IEEE TAC*, July 2005.

[3] S. Bansal, M. Chen, S. Herbert, and C. J. Tomlin, "Hamilton-Jacobi reachability: A brief overview and recent advances," in *IEEE CDC*, 2017.

[4] A. D. Ames, S. Coogan, M. Egerstedt, G. Notomista, K. Sreenath, and P. Tabuada, "Control barrier functions: Theory and applications," in *European Control Conference*, 2019.

[5] K. Srinivasan, B. Eysenbach, S. Ha, J. Tan, and C. Finn, "Learning to be safe: Deep rl with a safety critic," *arXiv preprint arXiv:2010.14603*, 2020.

[6] H. Bharadhwaj, A. Kumar, N. Rhinehart, S. Levine, F. Shkurti, and A. Garg, "Conservative safety critics for exploration," *arXiv preprint arXiv:2010.14497*, 2020.

[7] **J. J. Choi**, D. Lee, K. Sreenath, C. J. Tomlin, and S. L. Herbert, "Robust control barrier-value functions for safety-critical control," in *IEEE CDC*, 2021.

[8] **J. J. Choi**, A. Agrawal, K. Sreenath, C. J. Tomlin, and S. Bansal, "Computation of regions of attraction for hybrid limit cycles using reachability: An application to walking robots," in *IEEE RA-L (under review)*, 2021.

[9] **J. J. Choi***, F. Castañeda*, C. Tomlin, and K. Sreenath, "Reinforcement Learning for Safety-Critical Control under Model Uncertainty, using Control Lyapunov Functions and Control Barrier Functions," in *Robotics: Science and Systems*, 2020.

[10] F. Castañeda*, **J. J. Choi***, B. Zhang, C. J. Tomlin, and K. Sreenath, "Gaussian process-based min-norm stabilizing controller for control-affine systems with uncertain input effects and dynamics," in *American Control Conference*, 2021.

[11] ——, "Pointwise feasibility of gaussian process-based safety-critical control under model uncertainty," in *IEEE CDC*, 2021.