Massachusetts Institute of Technology
Department of Aeronautics and Astronautics

Thesis Proposal
Doctor of Philosophy

Large-scale, full-stack robot safety verification using program analysis and automated inference

Date of Submission:

March 24, 2023

Author: Charles Dawson
PhD Candidate

Committee: Chuchu Fan (Chair)
Russ Tedrake
Sertaç Karaman

External Evaluator: Vikash Mansinghka

Contents

Abstract

Abstract should be no more than **300 words in 1 page**.

State the significance of the proposed research. Include long-term objectives and specific aims. Describe concisely the research design and methods for achieving these objectives. Highlight the specific hypotheses to be tested, goals to be reached, or technology to be developed, which are intended to be your original contributions. Avoid summaries of past accomplishments.

1 Introduction

1.1 Thesis Objectives

This thesis aims to do X. In support of this goal, I will:

1. Goal 1
2. Goal 2
3. Goal 3
4. Goal 4

1.2 Impact

2 Background and Significance

My thesis aims to build on prior work in three distinct but related areas: differentiable and probabilistic programming, safety verification (both model-based and black-box), and robust and adversarial optimization. This section will review each of these fields with an eye towards framing the significance of my planned thesis contributions.

2.1 Programs as mathematical models

2.1.1 Differentiable programming

2.1.2 Probabilistic programming

2.2 Safety verification

2.2.1 Model-based safety verification

2.2.2 Black-box safety verification

2.3 Robust and adversarial optimization

2.4 Significance of planned contributions

2.4.1 Programs-as-models: a middle-ground between model-free and model-based

2.4.2 Dynamic adversarial optimization

3 Local methods for design & verification

3.1 Problem statement

3.2 Variance-regularized design optimization

3.2.1 Case study: multi-robot manipulation

3.3 Local adversarial testing

3.3.1 Case study: robust planning from formal specifications

3.4 Discussion & Limitations

4 Global methods for design and verification

4.1 From optimization to inference

4.1.1 Gradient-accelerated automated inference

4.2 Sampling diverse failure modes

4.2.1 Case study: predicting transmission outages in electrical power networks

4.3 Repairing failure modes

4.3.1 Case study: robust generation dispatch for secure power networks

5 Future work

5.1 Pushing the limits of differentiability: vision, manipulation, and beyond

5.2 Theoretical analysis

5.3 Probabilistic programming

5.4 Static program analysis

6 Milestones and Program Logistics

6.1 Classes and Degree Milestones

Table 6.1 shows my competed coursework, and Table 6.2 shows completed and anticipated degree milestones.

Table 6.1 | My completed coursework, satisfying all academic requirements for the doctoral program. Major: autonomy. Minor: controls.

Semester	Class	Req.	Status
Fall 2019	16.413 Principles of Autonomy & Decision Making	major	completed
Fall 2019	6.255 Optimization Methods	major/math	completed
Spring 2020	16.412 Cognitive Robotics	major	completed
Spring 2020	6.832 Underactuated Robotics	minor	completed
Fall 2020	18.385 Nonlinear Dynamics and Chaos	minor/math	completed
Fall 2020	2.160 Identification, Estimation, and Learning	minor	completed
Spring 2021	16.S398 Formal Methods in Autonomy	major	completed
Fall 2021	6.843: Robotic Manipulation	major	completed
Fall 2021	16.995 Doctoral Research & Communication Seminar	RPC	completed

Table 6.2 | Milestones towards my completion of the doctoral degree. Italicized milestones are anticipated.

Fall 2019 (September)	Began studies at MIT
Fall 2020 (December)	Field evaluation complete
Spring 2021 (May)	Masters thesis submitted
Fall 2022 (September)	Committee meeting #1
Summer 2023 (July)	Commitee meeting #2 and thesis proposal
Fall 2023	*Committee meeting #3*
Spring 2024	*Committee meeting #4*
Spring 2024	*Thesis defense*

6.2 Research Schedule

My thesis research will proceed in stages, as outlined below.

Already completed:

Spring 2022

1. Certifiable robot design optimization using differentiable programming
   (a) Develop design optimization tool using automatic differentiation
   (b) Develop statistical robustness certification tool based on extremal types theorem
   (c) Hardware deployment
   (d) Accepted to RSS 2022
2. Robust counterexample-guided optimization with temporal logic specifications

(a) Define two-player zero-sum game between the designer and the verifier

(b) Incorporate counterexamples from the verifier to guide robust design optimization

(c) Use differentiable signal temporal logic for complex task specification

(d) Submitted to IROS 2022

Fall/Winter 2022

1. Improving design optimization and verification through automated failure mode discovery

   (a) Use differentiable programming and sequential MCMC to discover a diverse set of possible failure modes.

   (b) Use counterexamples from all failure modes to guide robust design optimization

Spring 2023

1. Extend failure mode prediction and mitigation framework to challenging new problem domains, including:

   (a) Perception-in-the-loop controllers, through the use of differentiable rendering,

   (b) Deep neural network controllers.

Fall 2023

1. Develop MCMC-based algorithm for safety verification and optimization of systems with discrete structure (e.g. involutive MCMC for robot manipulators with variaible morphologies).

2. Write thesis

Spring 2024

1. Write thesis

2. Defend thesis and graduate

References