UNIVERSITY MICROFILMS INC.

D
PAO

MIT Doctoral Dissertation
PUBLISH ABSTRACT ONLY
INFORMATION

______________ Abstract No.

DO NOT WRITE IN THIS SPACE

Vol/Issue ______________
School Code ______________
Advisor ______________

PLEASE TYPE OR PRINT

PERSONAL DATA

1. Full name (as it appears on dissertation title page)

Gross	Jason	S.
(last)	(first)	(middle)

2. Year of birth (optional) 1991

3. Present mailing address 258 Prospect Street
Apt 1L
Cambridge, MA 02139

Future mailing address c/o Terri Ackerman
126 Hayrick Lane
Commack, NY 11725

Effective date of future mailing address February 1, 2021
Home telephone ______________ Business telephone ______________

DOCTORAL DEGREE DATA

4. Full name of university conferring degree Massachusetts Institute of Technology

5. Degree awarded (check one) ☑ Ph.D. ☐ Sc.D.

6. Year degree awarded 2021

---

7. IMPORTANT: Attach a copy of your dissertation title page and abstract to this form. Please be certain that the name of your dissertation supervisor is included on both.

8. Subject categories for your dissertation. Enter 4-digit code from 'Subject Categories' list found on the opposite side of this form, and write in the category selected. You may enter two additional codes and categories on the lines provided.

Code 0984 Category Computer science
Code ______________ Category ______________
Code ______________ Category ______________

(Optional) List up to five additional words from your dissertation not already found in ***either*** your title ***or*** abstract which would be useful for database access.

a. ______________ b. ______________ c. ______________

d. ______________ e. ______________

[**TODO:** Make sure to exclude no files (not rewriting,rewriting-appendix)]

Performance Engineering of Proof-Based Software Systems at Scale

by

Jason S. Gross

B.S., Massachusetts Institute of Technology (2013)
S.M., Massachusetts Institute of Technology (2015)

Submitted to the Department of Electrical Engineering and Computer Science
in partial fulfillment of the requirements for the degree of

Doctor of Philosophy

at the

MASSACHUSETTS INSTITUTE OF TECHNOLOGY

February 2021

Author . . . .
Department of Electrical Engineering and Computer Science
January 27, 2021

Certified by . . . .
Adam Chlipala
Associate Professor of Electrical Engineering and Computer Science
Thesis Supervisor

Accepted by . . . .
Leslie A. Kolodziejski
Professor of Electrical Engineering and Computer Science
Chair, Department Committee on Graduate Students

Performance Engineering of Proof-Based Software Systems at Scale

by

Jason S. Gross

Submitted to the Department of Electrical Engineering and Computer Science on January 27, 2021, in partial fulfillment of the requirements for the degree of Doctor of Philosophy

Abstract

Formal verification is increasingly valuable as our world comes to rely more on software for critical infrastructure. A significant and understudied cost of developing mechanized proofs, especially at scale, is the computer performance of proof generation. This dissertation aims to be a partial guide to identifying and resolving performance bottlenecks in dependently typed tactic-driven proof assistants like Coq.

We present a survey of the landscape of performance issues in Coq, with micro- and macro-benchmarks. We describe various metrics that allow prediction of performance, such as term size, goal size, and number of binders, and note the occasional surprising lack of a bottleneck for some factors, such as total proof term size. To our knowledge such a roadmap to performance bottlenecks is a new contribution of this dissertation.

The central new technical contribution presented by this dissertation is a reflective framework for partial evaluation and rewriting, already used to compile a code generator for field-arithmetic cryptographic primitives which generates code currently used in Google Chrome. We believe this prototype is the first scalably performant realization of an approach for code specialization which does not require adding to the trusted code base. Our extensible engine, which combines the traditional concepts of tailored term reduction and automatic rewriting from hint databases with on-the-fly generation of inductive codes for constants, is also of interest to replace these ingredients in proof assistants' proof checkers and tactic engines. Additionally, we use the development of this framework itself as a case study for the various performance issues that can arise when designing large proof libraries. We also present a novel method of simple and fast reification, developed and published during this PhD.

Finally, we present additional lessons drawn from the case studies of a category-theory library, a proof-producing parser generator, and cryptographic code generation.

Thesis Supervisor: Adam Chlipala
Title: Associate Professor of Electrical Engineering and Computer Science