

Tyler Sorensen – CV

Department of Computer Science
Computer Science Building, 35 Olden St.
Princeton, NJ 08540

www.cs.princeton.edu/~ts20
ts20@cs.princeton.edu
+1 (385) 271-3112

Research Summary

My research interests are in enabling efficient and reliable applications to be developed and executed on current and near-future systems (e.g. heterogeneous architectures). I am interested in a three step approach: *modeling* the semantics of underspecified areas of the system (e.g. operational/axiomatic semantic modeling), *validating* the semantics (e.g. empirical testing, model checking), and *developing* more efficient and reliable applications using the new understanding of the system (e.g. auto-tuning, domain-specific languages).

Current Position

July 2018 - Present Princeton University: Department of Computer Science
Postdoctoral research associate
Supervisor: Margaret Martonosi

Education

Fall 2014- Fall 2018 PhD in Computer Science - Imperial College London, UK
Thesis: *Device-wide Barrier Synchronisation on Graphics Processing Units*
Supervisor: Alastair F. Donaldson

Fall 2013 - Spring 2014 MS in Computer Science - University of Utah, USA
Thesis: *Testing and Exposing Weak GPU Memory Models*
Supervisor: Ganesh Gopalakrishnan

Fall 2010 - Spring 2012 BS in Computer Science - University of Utah, USA
Thesis: *Towards Shared Memory Consistency Models for GPUs*
Supervisor: Ganesh Gopalakrishnan

Fall 2010 - Spring 2012 BS in Applied Mathematics - University of Utah, USA

Fall 2008 - Spring 2010 AA with engineering emphasis - Snow College (community college), USA

Publications

Key: ★ Lead (or co-lead) author. A+, A, B, C: The venue CORE ranking (where available).

Conference Publications

- [1] Tyler Sorensen, Hugues Evrard, Alastair F. Donaldson. “GPU Schedulers: How Fair is Fair Enough?”. In: *Int. Conf. on Concurrency Theory (CONCUR)*. 2018. ★ **CONCUR A** 2018
- [2] Nathan Chong, Tyler Sorensen, John Wickerson. “The Semantics of Transactions and Weak Memory in x86, Power, ARM, and C++”. In: *ACM Conf. on Programming Language Design and Implementation (PLDI)*. 2018. **PLDI A+** 2018
Artifact evaluated and accepted.
Industry collaboration (Nathan Chong at ARM). Distinguished paper award.

- [3] Tyler Sorensen, Hugues Evrard, Alastair F. Donaldson. “Cooperative Kernels: GPU Multitasking for Blocking Algorithms”. In: *ACM Symp. on The Foundations of Software Engineering*. 2017. **Distinguished paper award**. ★ **FSE** A+ 2017
- [4] John Wickerson, Mark Batty, Tyler Sorensen, George A. Constantinides. “Automatically Comparing Memory Consistency Models”. In: *ACM Symp. on Principles of Programming Languages (POPL)*. 2017. **Artifact evaluated and accepted**. **Methodology used by Industry in defining the Vulkan GPU memory consistency model**. **POPL** A+ 2017
- [5] Tyler Sorensen, Alastair F. Donaldson, Mark Batty, Ganesh Gopalakrishnan, Zvonimir Rakamarić. “Portable Inter-workgroup Barrier Synchronisation for GPUs”. In: *ACM Conf. on Object-Oriented Programming, Systems, Languages, and Applications (OOPSLA)*. 2016. **Artifact evaluated and accepted**. ★ **OOPSLA** A+ 2016
- [6] Tyler Sorensen, Alastair F. Donaldson. “Exposing Errors Related to Weak Memory in GPU Applications”. In: *ACM Conf. on Programming Language Design and Implementation (PLDI)*. 2016. ★ **PLDI** A+ 2016
- [7] Jade Alglave, Mark Batty, Alastair F. Donaldson, Ganesh Gopalakrishnan, Jeroen Ketema, Daniel Poetzl, Tyler Sorensen, John Wickerson. “GPU Concurrency: Weak Behaviours and Programming Assumptions”. In: *Int. Conf. on Architectural Support for Programming Languages and Operating Systems (ASPLOS)*. 2015. **HiPEAC Paper Award**. **Invited for fast-track submission to ACM Transactions on Computer Systems**. ★ **ASPLOS** A+ 2015

Workshop/Other Publications (Peer reviewed)

- [8] Tyler Sorensen, Alastair F. Donaldson. “The Hitchhiker’s Guide to Cross-platform OpenCL Application Development”. In *Int. Workshop on OpenCL (IWOCL)*. 2016. ★ **IWOCL** 2016
- [9] Jade Alglave, Luc Maranget, Daniel Poetzl, Tyler Sorensen. “I Compute, Therefore I am (Buggy): Methodic Doubt Meets Multiprocessors”. In *Tiny Transactions on Computer Science Volume 3 (Tiny ToCS)*. 2015. ★ **TinyToCS** 2015
- [10] Tyler Sorensen, Ganesh Gopalakrishnan, Vinod Grover. “Towards Shared Memory Consistency Models for GPUs”. In *Int. Conf. on Supercomputing (ICS)*. 2013. **1st place ACM undergrad student research competition (SRC)**. ★ **ICS** A 2013

Under Submission

- [11] Tyler Sorensen, Sreepathi Pai, Alastair F. Donaldson. “When One Size Doesn’t Fit All: Quantifying Performance Portability of Graph Applications on GPUs”. 2018. ★ 2018
- [12] Dan Iorga, Tyler Sorensen, Alastair F. Donaldson. “Do Your Cores Play Nicely? A Portable Framework for Multi-core Interference Tuning and Analysis”. 2018. 2018

Professional Service

Conference and workshops

Year	Venue	Role
2018	PLDI (A+)	External review committee (ERC)
2018	ECOOP (A)	Program committee for doctoral symposium (PC for DS)
2016	ASPLOS (A+)	Sub-reviewer
2016	TinyToCS	Program committee (PC)
2016	POPL (A+)	Artifact evaluation committee (AEC)
2014	POPP (A+)	Sub-reviewer

Journals

Year	Venue	Role
2018	Formal Methods in System Design (FMSD)	Article reviewer
2018	ACM Trans. on Arch. and Code Optimization (TACO)	Article reviewer
2014	ACM Trans. on Prog. Lang. and Systems (TOPLAS)	Sub-reviewer

Invited Talks

Industry

- Cooperative Kernels: GPU Multitasking for Blocking Algorithms (extended FSE'17 talk)
 - Feb., 2017: ARM Cambridge, UK
- Portable Inter-workgroup Barrier Synchronisation for GPUs (extended OOPSLA'16 talk)
 - May, 2016: AMD Bellevue, WA, USA

Academic

- GPU Concurrency: The Wild West of Programming (PhD Talk):
 - June, 2018: University of Utah, USA
 - Sept., 2018: TU Darmstadt, Germany
 - March, 2018: University of Glasgow, UK
- Exposing Errors Related to Weak Memory in GPU Applications (extended PLDI'16 talk)
 - Nov., 2016: University College London, UK
 - Oct., 2015: University of Kent, UK

Awards, Press and Recognition

Summer 2018	Distinguished paper award – Programming Language Design and Implementation (PLDI) 2018
Summer 2017	Distinguished paper award – Foundations of Software Engineering (FSE) 2017
Fall 2016	Art of research staff pick award – Imperial College http://multicore.doc.ic.ac.uk/projects/artofresearch
Fall 2015	GPU testing led to errata entry for textbook: <i>CUDA by Example</i> https://developer.nvidia.com/cuda-example-errata-page

Fall 2015	GPU testing contributions mentioned in LWN article <i>Axiomatic validation of memory barriers and atomic instructions</i> https://lwn.net/Articles/608550/
Spring 2014	Outstanding graduating senior award – University of Utah
Spring 2014	List of 40 outstanding alumni under 40 – Snow College
Fall 2013	Top 5 in undergraduate researcher competition – Computing Research Association (CRA)
Summer 2013	1st place ACM undergrad student research competition – Int. Conf. on Supercomputing (ICS)
Fall 2012	1st place team super computing student cluster competition - Super Computing (SC)
Fall 2012	2nd place at local ACM fall programming competition – Brigham Young University (BYU)
Spring 2012	1st place at local ACM spring programming competition – Brigham Young University (BYU)
Spring 2011	Honorable mention in the Microsoft Imagine Cup game design competition

Industry Impact Summary

The work I have been involved in has had significant impact in industry. In particular:

- **Memory Consistency Models for GPUs:** Over the last few years, GPU standards have provided more precise memory model documentation: OpenCL (in 2014), HSA (in 2016), Nvidia PTX (in 2017), Vulkan (in 2018). Our ASPLOS'15 and PLDI'16 work revealed significant misunderstandings in the GPU programming community and is considered a key motivator of this industrial trend. My BS thesis (2013) is one of the first works on formalizing GPU memory consistency models.
- **Vulkan Memory Consistency:** The cross-vendor GPU programming language Vulkan added a formal memory consistency model in 2018 using the exact methodology of our POPL'17 paper. The influence of our work was directly acknowledged in communication with members of the standards committee.
- **ARM Transactional Memory:** The framework of our PLDI'18 work revealed a subtle bug in an internal ARM proposal for transactional memory support. ARM acknowledged the issue internally and used our methodology to provide confidence in subsequent proposals.

Coding Projects

Primary Author/Contributor

- **OpenCL Irgl:** DSL/compiler for portable GPU graph algorithms
Code available on request; public release planned soon
Portable state-of-the-art fundamental graph algorithms (BFS, SSSP, etc.) in portable OpenCL. Tested on Nvidia, AMD, ARM and Intel GPUs. *Currently the fastest implementations we are aware of for AMD, ARM and Intel GPUs.*
- **Discovery Barrier:** Portable device-wide GPU synchronization barrier
https://github.com/mc-imperial/gpu_discovery_barrier
Portable library for a GPU device-wide synchronization barrier using theory from the OOPSLA'16 paper. Tested on Nvidia, AMD, ARM and Intel GPUs. *Currently the only implementation that we are aware of that is capable of providing this synchronization in a portable way.*
- **PBDD:** Binary decision diagram (BDD) web application
<https://github.com/tyler-utah/PBDD>
Interactive web application for constructing BDDs from logic formulas, intended for pedagogical use.

- Average of 98 visits a month (as of Sept. 2018)
- Used in courses taught at Cambridge University (UK), University of Utah (USA), University of Nebraska (USA), Washington University in St. Louis (USA)

Contributor

- **MemAlloy**: Relational modeling and checking of memory consistency semantics

<https://github.com/johnwickerson/memalloy>

A framework based on the Alloy relational constraint solver for performing bounded model checking of many interesting memory consistency queries. My contributions examined exhaustively enumerating conformance tests (up to a bound) for a given memory consistency model. Used in POPL'17 and PLDI'18 publications.

- **CLBlast**: Tuned OpenCL BLAS library

<https://github.com/CNugteren/CLBlast>

Added low-level Nvidia-specific instructions to achieve up to a *1.5x speedup* for some GEMM (General matrix-matrix multiplication) configurations.

Teaching

Supervision

PhD students to whom I currently provide provide regular (at least weekly) technical advising

- Dan Iorga (Imperial College London)
Primarily supervised by Alastair F. Donaldson.
Closely supervised work leading to a top conference submission (Do Your Cores Play Nicely? A Portable Framework for Multi-core Interference Tuning and Analysis) and continuing weekly advising.
- Aninda Manocha (Princeton University)
Primarily supervised by Margaret Martonosi.
Technical supervision on a large DARPA project in which we aim to accelerate applications on new heterogeneous architectures. Supervision topics include: rigorous benchmarking practices, presentation of performance results, and proper usage of a variety of profiling tools.

Fall 2014 Microsoft research student workshop: *Think Computer Science*
<https://www.microsoft.com/en-us/research/event/think-computer-science-2014/>
Tutored lab sessions for middle school students using Python on Raspberry Pis

Fall 2012 Guest lecture - *University of Utah: CS3100 Models of Computation*
Gave two guest lectures on using pedagogical tools developed during the summer (see below)

Summer 2012 Summer project developing pedagogical material
Created (and continue to maintain) interactive tools for propositional logic courses.
Used in courses taught at Cambridge University (UK), University of Utah (USA),
University of Nebraska (USA), Washington University in St. Louis (USA)

Teaching Assistant

Responsibilities included: holding office hours, running lab sessions, grading, and exam invigilation.

Term	Course	Professor
Fall 2012	University of Utah: CS3100 Models of Computation	Ganesh Gopalakrishnan
Fall 2011	University of Utah: CS1410 Object Orient Programming	Joe Zachary

Grants

I contributed significantly to the writing of the following grant:

- **Multi-Core Microprocessor Test Harness (R1000115750)**

Defence Science and Technology Laboratory (DSTL)

Principal Investigator: Alastair F. Donaldson

Amount: 73,854 GBP

We proposed a methodology for rigorously testing the runtime of sequential applications in the presence of independent adversarial processes running on distinct cores of a multi-core system.

Internships

- Spring 2017** Microsoft Research – *RiSE group*
Mentors: Todd Mytkowicz, Madan Musuvathi, Saeed Maleki
Project: a parallel algorithm for DNNs that preserves sequential semantics
- Spring 2016** Microsoft – *Speech decoding product group*
Mentors: Veljko Miljanic, Hosam Khalil, Madan Musuvathi
Project: practical parallel speech decoding for product (Cortana backend)
Impact: several contributions accepted to main production branch
- Summer 2014** Nvidia – *Compiler group*
and Mentor: Vinod Grover
- Summer 2013** Project: Contributed to internal shared memory consistency model

References

- **Alastair F. Donaldson**
Reader, Department of Computing, Imperial College London, UK
Software Engineer, Android Platform Graphics, Google, London, UK
alastair.donaldson@imperial.ac.uk
- **Margaret Martonosi**
Professor, Department of Computer Science, Princeton University, NJ, USA
mrm@princeton.edu
- **Ganesh Gopalakrishnan**
Professor, University of Utah, UT, USA
ganesh@cs.utah.edu
- **Zvonimir Rakamarić**
Associate Professor, University of Utah, UT, USA
zvonimir@cs.utah.edu
- **Sreepathi Pai**
Assistant Professor, Department of Computer Science, University of Rochester, NY, USA
sree@cs.rochester.edu
- **John Wickerson**
Lecturer, Department of Electrical and Electronic Engineering, Imperial College London, UK
j.wickerson@imperial.ac.uk
- **Hugues Evrard**
Software Engineer, Android Platform Graphics, Google, London, UK
hevrard@google.com