

Victor A. Ying

(301) 337-8428 • victory@csail.mit.edu • victorying.com

RESEARCH INTERESTS	Computer architecture, parallelizing compilers, parallel programming models, programmable accelerators, locality-aware execution, speculative execution, distributed systems, parallel algorithms, parallel runtimes	
EDUCATION	Massachusetts Institute of Technology , Cambridge, Massachusetts	
	Ph.D. in Electrical Engineering and Computer Science	<i>anticipated 2022</i>
	S.M. in Electrical Engineering and Computer Science	June 2019
	<ul style="list-style-type: none">• Cumulative GPA: 4.93 / 5.00• Thesis title: Scaling Sequential Code with Hardware–Software Co-Design for Fine-Grain Speculative Parallelization• Thesis advisor: Daniel Sanchez	
	Princeton University , Princeton, New Jersey	
	B.S.E. <i>summa cum laude</i> in Electrical Engineering	May 2016
	<ul style="list-style-type: none">• Cumulative GPA: 3.95 / 4.00• Thesis title: Analyzing Decision Heuristic Effectiveness in Boolean Satisfiability Solvers• Thesis advisor: Sharad Malik	
	Selected coursework: Computer architecture, operating systems, computer networks, algorithms, functional programming, program analysis, logic design, machine learning	
RESEARCH & INDUSTRY EMPLOYMENT	Research Assistant and Edwin Webster Fellow	September 2016 – Present
	MIT Computer Science and Artificial Intelligence Lab, Cambridge, Massachusetts	
	<ul style="list-style-type: none">• Supervisor: Daniel Sanchez• Design and evaluate enhancements to the Swarm multicore architecture, using microarchitectural simulation.• Lead development of LLVM/Clang-based compilers targeting new hardware for massive parallelism.• Implement new language extensions and domain-specific languages for high-performance graph processing.	
	Research Intern	June – September 2021
	Microsoft Research Lab - Redmond, Washington	
	<ul style="list-style-type: none">• Build MLIR-based compiler and prototype code transformations to co-optimize communication and computation in distributed GPU workloads such as training and inference for enormous ML models.	
	Research Intern	May – August 2018
	NVIDIA Research, Westford, Massachusetts	
	<ul style="list-style-type: none">• Develop analytical modeling tool for design space exploration and code optimization for efficient execution of linear algebra and machine learning workloads on a range of future hardware architectures.	
	Hardware Engineering Intern	May – August 2015
	Pure Storage, Mountain View, California	
	<ul style="list-style-type: none">• Developed firmware (C) and created tools (Python) for debugging prototype embedded hardware through a serial connection. Implemented a command line interface, GDB server, resource monitoring tools, and a checksummed packet protocol.	
	Software Engineering Intern	May – August 2014
	Pure Storage, Mountain View, California	
	<ul style="list-style-type: none">• Developed and deployed the first driver enabling integration of Pure Storage FlashArrays and OpenStack, an open-source cloud platform. Transferred ownership of this sales-driving feature to full-time engineers.• Wrote and open-sourced a Python library for managing FlashArrays, used for automated testing.	
	Technical Intern	June – August 2013
	Northrop Grumman Electronic Systems, Baltimore, Maryland	
	<ul style="list-style-type: none">• Optimized designs of RF electronics in radar systems using CAD and simulation tools.• Characterized prototypes to identify suspect connections and components to be redesigned.	

Student Technician

June 2012 – June 2016

National Institute of Standards and Technology, Gaithersburg, Maryland

- Supervisor: Heather J. Patrick
- Developed precise positioning software for robotic arms to enable repeatable reflectance measurements.
- Modeled distortions in optical scattering measurements and automated post-processing correction factors.

**REFEREED
CONFERENCE
PAPERS**

A. Brahmakshatriya, E. Furst, **V. A. Ying**, C. Hsu, C. Hong, M. Ruttenberg, Y. Zhang, T. Jung, D. Richmond, M. Taylor, J. Shun, M. Oskin, D. Sanchez, and S. Amarasinghe, “Taming the Zoo: A Unified Graph Compiler Framework for Novel Architectures”, in *48th Intl. Symposium on Computer Architecture (ISCA)*, 2021. Acceptance rate: 76/407 (19%)

V. A. Ying, M. C. Jeffrey, and D. Sanchez, “T4: Compiling Sequential Code for Effective Speculative Parallelization in Hardware”, in *47th Intl. Symposium on Computer Architecture (ISCA)*, 2020. Acceptance rate: 77/428 (18%)

A. Parashar, P. Raina, Y. S. Shao, Y.-H. Chen, **V. A. Ying**, A. Mukkara, R. Venkatesan, B. Khailany, S. W. Keckler, and J. Emer, “Timeloop: A Systematic Approach to DNN Accelerator Evaluation”, in *Intl. Symposium on Perf. Analysis of Systems and Software (ISPASS)*, 2019. Acceptance rate: 26/88 (30%)

M. C. Jeffrey, **V. A. Ying**, S. Subramanian, H. R. Lee, J. Emer, and D. Sanchez, “Harmonizing Speculative and Non-Speculative Execution in Architectures for Ordered Parallelism”, in *51st Intl. Symposium on Microarchitecture (MICRO)*, 2018. Acceptance rate: 74/351 (21%)

S. Subramanian, M. C. Jeffrey, M. Abeydeera, H. R. Lee, **V. A. Ying**, J. Emer, and D. Sanchez, “Fractal: An Execution Model for Fine-Grain Nested Speculative Parallelism”, in *44th Intl. Symposium on Computer Architecture (ISCA)*, 2017. Acceptance rate: 54/322 (17%)

**OTHER
PUBLICATIONS**

S. Malik and **V. A. Ying**, “On the Efficiency of the VSIDS Decision Heuristic”, presented at *Theoretical Foundations of SAT Solving Workshop*, 2016.

H. J. Patrick, C. J. Zarobila, T. A. Germer, **V. A. Ying**, C. A. Cooksey, and B. K. Tsai, “Tunable supercontinuum fiber laser source for BRDF measurements in the STARR II gonireflectometer”, in *Proceedings of SPIE Volume 8495*, 2012.

TALKS

“Parallelizing Sequential Code with Compiler-Hardware Co-Design”, at UC Santa Cruz (Languages, Systems, and Data Seminar), February 2021.

“T4: Parallelizing Sequential Code with Compiler-Hardware Co-Design”, at Facebook, June 2020.

“T4: Compiling Sequential Code for Effective Speculative Parallelization in Hardware”, at *47th Intl. Symposium on Computer Architecture (ISCA)*, June 2020.

“SCC: Compiling Sequential Code for Effective Speculative Parallelization in Hardware”, at *Boston Area Architecture Workshop (BARC)*, January 2020.

“Compiling Sequential Code for a Speculative Parallel Architecture”, selected from Student Research Competition to present in main session of *41st ACM SIGPLAN Conference on Programming Language Design and Implementation (PLDI)*, June 2019.

“Making Parallelism Pervasive with the Swarm Architecture”, guest lecture in MIT course 6.S898: *Advanced Performance Engineering for Multicore Applications*, 2017.

**HONORS
& AWARDS**

Best PhD Forum Poster, HPDC 2019

Second Place in Student Research Competition, PLDI 2019

Best Poster, Industry-Academia Partnership MIT Cloud Workshop 2018

Honorable Mention, NSF Graduate Research Fellowship Program 2018

Edwin Webster Fellowship, \$77,711 from MIT Dept. of EECS 2016–2017

Honorable Mention, Ford Foundation Predoctoral Fellowship Program 2016

Highest Honors , Princeton Dept. of Electrical Engineering	2016
Hisashi Kobayashi Prize , Princeton Dept. of Electrical Engineering	2016
Sigma Xi , Princeton Chapter	2016
Phi Beta Kappa , New Jersey Beta Chapter	2015
Tau Beta Pi , New Jersey Delta Chapter	2014
Shapiro Prize for Academic Excellence , Princeton University	2014

TEACHING & MENTORSHIP

Chief Operating Officer	June 2012 – Present
Kids Are Scientists Too, a national 501(c)(3) nonprofit	
<ul style="list-style-type: none"> • Expand after-school science programs for underprivileged elementary school students to nine states. • Mentor high school branch leaders and volunteers, who recruit peers, fundraise, and run science activities. • Manage finances, tax filings, nonprofit status, and KAST’s website and shared online resources for branches. 	
Teaching Assistant	Spring 2020
6.823: Computer System Architecture, MIT	
<ul style="list-style-type: none"> • Held discussion sessions, review sessions, and office hours on graduate-level computer architecture. • Wrote, edited, and graded lab assignments and quizzes to teach principles of architecture research. 	
Lab Teaching Assistant	Fall 2014, Fall 2015
ELE 206: Contemporary Logic Design, Princeton University	
<ul style="list-style-type: none"> • Held lab sessions and taught digital logic, RTL design, and FPGA synthesis. • Rewrote assignments to define and use a subset of Verilog and new cross-platform simulation software. • Overhauled the general-purpose processor design project with a new ISA and software testing tools. 	
Peer Academic Advisor and Peer Tutor	2015–2016
Office of the Dean of Undergraduate Students, Princeton University	
<ul style="list-style-type: none"> • Engage first years in planning their academic paths, enrolling in courses, and adjusting to college academics. • Tutor students in introductory mathematics, physics, and engineering classes. 	

SKILLS

Computer architecture research, analytical and simulation-based modeling, compiler optimizations.
 Proficient in C++, C, Python, LLVM, x86/64 assembly, and Unix tools.
 Experience with PyTorch, TensorFlow, MATLAB, Java, Verilog, OCaml, Haskell, Z3, and Coq.

[Curriculum vitae compiled on 2021-08-27]