

Pratyush Das

- *Email* das160@purdue.edu
- *GitHub* <https://github.com/reikdas>

Education

Purdue University

August, 2021 -

PhD in Computer Science (Advisor - Professor Tiark Rompf). GPA: 3.77/4.0

Research interests - Metaprogramming, Automatic parallelization, Static analysis

Institute of Engineering & Management, Kolkata (MAKAUT)

August, 2017 - May, 2021

Bachelor of Technology in Computer Science and Engineering. CGPA: 8.89/10

Awarded the Director's Award for Best Scientific Mind

Internships

Google Summer of Code - Student

June, 2021 - August, 2021

The LLVM Compiler Infrastructure Organization

Supervisors - William Moses (MIT), Dr. Johannes Doerfert (Argonne National Laboratory)

- **Enzyme**: LLVM Pass to perform automatic differentiation of statically analyzable LLVM IR
 - Integrated custom derivatives of BLAS functions into Enzyme.
 - Wrote an LLVM pass to inline function definitions from bitcode files into LLVM IR.

IRIS-HEP - Fellow

June, 2020 - September, 2020

Supervisor - Dr. Jim Pivarski (Princeton University)

- **Awkward Array**: Library for nested, variable-sized data using NumPy-like idioms
 - Created a source to source compiler to generate equivalent Python for a subset of C++.
 - Created a property based testing framework.
 - Created a source to source compiler to generate equivalent parallel CUDA from specification (Python and type info).

IRIS-HEP and DIANA-HEP - Fellow

June, 2018 - September, 2018; June, 2019 - September, 2019

Location: Fermi National Accelerator Laboratory, USA - LHC Physics Centre

Supervisor - Dr. Jim Pivarski (Princeton University)

- **Uproot**: Python implementation of ROOT I/O, an open source file format storing over an exabyte of HEP data
 - Enabled writing fundamental HEP data structures like TTrees and histograms to ROOT files.
 - Uproot has become one of the most widely used HEP libraries.

Open Source Contributions

Supervisors - Dr. Nikos Vasilakis (MIT), Konstantinos Kallas (UPenn)

February, 2022 -

- **PaSh**: A system for parallelizing POSIX shell scripts
 - Helped extend PaSh for distributed systems

Supervisor - Dr. Vassil Vassilev (Princeton University)

November, 2019 - May, 2021

- **ROOT**: An open-source data analysis framework storing over an exabyte of data
 - Improvements to interpreter (rootcling)
- **Cling**: Interactive C++ interpreter built on top of Clang
 - Maintained cpt.py installer and packager
- **Clang**: C language family frontend for LLVM
 - Several patches to print type information of C++ template arguments

Supervisor - Dr. Jim Pivarski (Princeton University)

January, 2021 - February, 2021

- **Awkward Array** - Library for nested, variable-sized data using NumPy-like idioms
 - Created a parser for Awkward Array's type system

Supervisors - Dr. Jim Pivarski (Princeton University), Dr. Viktor Khristenko (CERN)

June, 2017 - August, 2017

- **spark-root** - Apache Spark datasource for ROOT
 - Separated spark bindings from TTree reading code
- **root4j** - Java implementation of ROOT file reader
 - Optimized codebase to facilitate interoperability

Teaching Experience

• **CS 354: Operating Systems** - Purdue University

Fall 2022

• **CS 240: Programming in C** - Purdue University

Fall 2021, Spring 2022

Programming Languages and Tools	
Experienced: Python, C, CUDA	
Familiar: C++, Java, Scala, Coq, ROOT, Bash, L ^A T _E X, Swift	
Summer Schools	
Oregon Programming Languages Summer School - University of Oregon	2021
Computational and Data Science for High Energy Physics - Princeton University	2019
Publications	
<ul style="list-style-type: none"> J.Pivarski, I.Osborne, P.Das, D.Lange, P.Elmer, “AwkwardForth: accelerating Uproot with an internal DSL”, 25th International Conference on Computing in High-Energy and Nuclear Physics (vCHEP, 2021), DOI: 10.1051/epjconf/202125103002. 2021 J.Pivarski, I.Osborne, P.Das, A.Biswas, P.Elmer, “Awkward Array: JSON-like data, NumPy-like idioms”, Proceedings of the 19th Python in Science Conference (SciPy USA, 2020), Pages 68-74, DOI: 10.25080/Majora-342d178e-00b. 2020 E.Rodrigues, et al., “The Scikit HEP Project - overview and prospects“, Proceedings of the 24th International Conference on Computing in High Energy and Nuclear Physics (CHEP 2019), DOI: 10.1051/epjconf/202024506028. 2020 N.Saha, P.Das, H.N.Saha, “Authorship Attribution of Short Texts using a Multi Layer Perceptron”, International Journal of Applied Pattern Recognition, 2018 Vol. 5 No. 3, Pages 251-259, DOI: 10.1504/IJAPR.2018.10016100. 2018 	
Drafts	
<ul style="list-style-type: none"> P.Das, A.Xhebraj, S.Abeysinghe, T.Rompf, “Specializing Data Access in a Distributed File System”. T.Mustafa, K.Kallas, P.Das, N.Vasilakis, “DiSh: Dynamic Shell-Script Distribution”. 	
Invited talks at Conferences	
• GSoC Experience - Enzyme (LLVM Developers’ Meeting)	2021
• Python in High Energy Physics (SciPy India, PyCon USA)	2019, 2020
• Writing files with uproot (PyHEP)	2019
• Writing files with uproot (ROOT Users’ Workshop)	2018
Invited talks at External Research Group Meetings	
• Language Transformations for the Awkward Array library (IRIS-HEP Fellow Presentations)	2020
• CUDA backend for the Awkward Array project (Princeton University Liberty Research Group)	2020
• Writing TTrees with uproot (IRIS-HEP Topical Meeting: Summer student project presentations)	2019
• Writing files with uproot (DIANA Meeting: Updates on ROOT I/O)	2018
• Separation of Concerns - Refactoring code between ROOT4J and Spark-Root (CMS Big Data Science, DIANA-HEP)	2017