Contribution ID: **108**   Type: **Poster**

Awkward Family: expanding functionality through interrelated Python packages

Wednesday 13 March 2024 16:15 (30 minutes)

In the 5+ years since their inception, Uproot and Awkward Array have become cornerstones for particle physics analysis in Python, both as direct user interfaces and as base layers for physicist-facing frameworks. Although this means that the software is achieving its mission, it also puts the need for stability in conflict with new, experimental developments. Boundaries must be drawn between the code that must stay robust and the code that implements new ideas.

In this poster, I'll describe how we leverage Python's packaging infrastructures to separate stable components from experimental components at the package boundaries. The uproot and awkward packages were chosen to be long-term maintenance components, while new capabilities are provided in:

- dask-awkward: distributed computing
- awkward-pandas: DataFrame interface
- AwkwardArray.jl: Julia interface and reinterpretation
- kaitai_struct_awkward_runtime: arbitrary file format → Awkward Array generator
- odapt: high-level file operations: copying, converting, concatenating, skimming and slimming
- uproot-browser: high-level TUI interface to Uproot
- ragged: just the ragged arrays, but satisfying the Python Array API
- vector: Lorentz vector manipulation (in and out of Awkward Arrays)

I'll also describe some best practices (learned through mistakes!) in coordinating versions, managing deprecations, public/private API boundaries, and cross-package testing.

Significance

References

- Uproot: https://uproot.readthedocs.io/
- Awkward Array: https://awkward-array.org/
- dask-awkward: https://github.com/dask-contrib/dask-awkward
- awkward-pandas: https://github.com/intake/awkward-pandas
- AwkwardArray.jl: https://github.com/JuliaHEP/AwkwardArray.jl
- kaitai_struct_awkward_runtime: https://github.com/ManasviGoyal/kaitai_struct_awkward_runtime
- odapt: https://github.com/zbilodea/odapt
- uproot-browser: https://github.com/scikit-hep/uproot-browser
- https://github.com/jpivarski/ragged

- Vector: https://github.com/scikit-hep/vector
- Python Array API: https://data-apis.org/array-api/latest/index.html

Experiment context, if any

Primary author: PIVARSKI, Jim (Princeton University)

Presenter: PIVARSKI, Jim (Princeton University)

Session Classification: Poster session with coffee break

Track Classification: Track 1: Computing Technology for Physics Research