AWKWARD-CPP AND PYBIND11

BY CHARLES ESCOTT

MENTORS: JIM PIVARSKI AND DAVID LANGE
THE PROBLEM

• Particle physics data is too complicated for normal coding libraries

• Awkward Arrays supports less conventional types/functions using Python
  • Easier to work with
  • JIT (just in time) compilation => slow to run

Jagged Array Structure

<table>
<thead>
<tr>
<th>Rows</th>
<th>Columns</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>0 1 2 3</td>
</tr>
<tr>
<td>1</td>
<td>4 5</td>
</tr>
<tr>
<td>2</td>
<td>6 7 8</td>
</tr>
</tbody>
</table>
THE SOLUTION

• Implement Awkward Arrays in C++!
  • Pre-compiled => usually much faster speeds
• Pybind11 allows C++ code to compile into Python binaries
• End product: a Python library that’s pre-compiled, written in C++
NEW PROJECT
ARCHITECTURE

Made by Jim Pivarski (thanks Jim, I'm stealing your diagram)
IMPLEMENTING C >> C++ >> PYTHON

• C is a bit more difficult than C++
  • No templates
  • No overloaded functions
  • No classes
  • No exceptions
  • No booleans

• Keeping C methods clean of memory management

• Ultimately, not too bad
“FINAL” PRODUCT

• A solid basis for Awkward 1.0 (the new architecture)

• Many hours and 5000+ lines of code worth of troubleshooting, research, and experience

• A JaggedArray class in C++/C which functions independently from awkward-array and contain ZERO Python code
THANK YOU!