Simple Geometries in NetCDF

Tim Whiteaker – The University of Texas at Austin
2017 Advancing NetCDF Workshop
Boulder, CO
# Geometry Types

<table>
<thead>
<tr>
<th>Included</th>
<th>Not Included</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Type</strong></td>
<td><strong>Type</strong></td>
</tr>
<tr>
<td>Point</td>
<td>MultiPoint</td>
</tr>
<tr>
<td>LineString</td>
<td>MultiLineString</td>
</tr>
<tr>
<td>Polygon</td>
<td>MultiPolygon</td>
</tr>
</tbody>
</table>

https://en.wikipedia.org/wiki/Well-known_text
Geometry Types

Included

<table>
<thead>
<tr>
<th>Type</th>
<th>MultiPoint</th>
<th>MultiLineString</th>
<th>MultiPolygon</th>
</tr>
</thead>
<tbody>
<tr>
<td>Point</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>LineString</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Polygon</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Compatible With

- Well-known Text geometry primitives
- OGC Simple Features
- GeoJSON
- Shapefile
- Various geospatial databases

https://en.wikipedia.org/wiki/Well-known_text
Some examples

Two locations where temperature is measured:

- (0, 90) – at north pole
- (0, 0) – at equator
How it works -- a point example

NetCDF File

- time
- lat
- lon
- temperature

geometry_container
  geometry_type = point
  node_coordinates = x y

- x: 0, 0
- y: 0, 90

Legend
- Old Stuff
- Geometry Stuff
Two Line Features

A

B

A
How to deal with different node counts

X Variable (2 by 5 array)

<table>
<thead>
<tr>
<th>Feature A</th>
<th>Feature B</th>
</tr>
</thead>
<tbody>
<tr>
<td>a1</td>
<td>b1</td>
</tr>
<tr>
<td>a2</td>
<td>b2</td>
</tr>
<tr>
<td>a3</td>
<td></td>
</tr>
<tr>
<td>a4</td>
<td></td>
</tr>
<tr>
<td>a5</td>
<td></td>
</tr>
</tbody>
</table>

This is inefficient
Contiguous Ragged Arrays

Values

<table>
<thead>
<tr>
<th>X</th>
</tr>
</thead>
<tbody>
<tr>
<td>a1</td>
</tr>
<tr>
<td>a2</td>
</tr>
<tr>
<td>a3</td>
</tr>
<tr>
<td>a4</td>
</tr>
<tr>
<td>a5</td>
</tr>
<tr>
<td>b1</td>
</tr>
<tr>
<td>b2</td>
</tr>
</tbody>
</table>

Counts

<table>
<thead>
<tr>
<th>Node Count</th>
</tr>
</thead>
<tbody>
<tr>
<td>5</td>
</tr>
<tr>
<td>2</td>
</tr>
</tbody>
</table>
NetCDF File

- geometry_container
  - geometry_type = line
  - node_coordinates = x y
  - node_count = node_count

- node_count = 5, 2

- time
- lat
- lon

- temperature
  - geometry = geometry_container

- x = 0, 1, 2, 3, 4, 0, 2
- y = 2, 3, 1, 3, 2, 8, 7
Two Multilines

B

A

one part

two parts
Multiline example

NetCDF File

geometry_container
  geometry_type = multiline
  node_coordinates = x y
    node_count = node_count
    part_node_count = part_node_count

x

y

node_count

5, 4

part_node_count

5, 2, 2

time
lat
lon

temperature
  geometry = geometry_container

5, 4
A Polygon

The hole is actually a second part to the polygon

exterior ring (part #1)

interior ring (part #2)
Polygon example (with holes)

NetCDF File

**geometry_container**
- geometry_type = polygon
- node_coordinates = x y
  - node_count = node_count
  - part_node_count = part_node_count
- interior_ring = interior_ring

**node_count**

**part_node_count**

**interior_ring** (yes or no)

0, 1

**time**

**lat**

**lon**

**temperature**
- geometry = geometry_container
Wouldn’t it be cool if this was in the CF Conventions
Timeline for Getting into CF (1.8)

May, 2016
- GitHub repo for simple geoms in netCDF
  - Has wiki and Python example
  - Uses coordinate index like UGRID
  - A second repo for R also created

September
- Request for feedback sent to CF-Metadata mailing list
- AGU poster

February, 2017
- Proposal sent to CF-Metadata mailing list
- Major revisions in our GitHub repo based on feedback
- Dropped coordinate index variable

April
- Pull request to cf-conventions on GitHub
- Chapter 7 - cells
- EC All-hands poster

June
- 2nd pull request to accommodate GitHub workflow

August
- trac ticket #164

https://github.com/cf-convention/cf-conventions/pull/115
Next Steps

- trac
- Persevere
- Update
- Explore

Limbo screenshot from Jesse Radonski on Flickr; GitHub images courtesy of GitHub.
Geometry and the Enhanced Data Model

- **VLEN**
  - Eliminates need for contiguous ragged arrays
  - Eliminates node count variable

  \[
  \begin{align*}
  x &: \{0, 1, 2, 3, 4\}, \{0, 2\} \\
  y &: \{2, 3, 1, 3, 2\}, \{8, 7\} 
  \end{align*}
  \]

- **Groups**
  - Could store each feature in its own group
  - Could store parts as VLENs, eliminating node counts and part node counts
Want To Contribute?

• Wiki and Python implementation
  https://github.com/twhiteaker/netCDF-CF-simple-geometry

• R implementation
  https://github.com/dblodgett-usgs/NCDFSG

• Pull request to cf-conventions
  https://github.com/cf-convention/cf-conventions/pull/115

• tim@utexas.edu