CF Chapter 9

This chapter provides representations for **discrete sampling geometries,** such as time series, vertical profiles and trajectories

John Caron

5/12/2011

Discrete Sampling Convention

- Originally called *Point Observation Convention*
- Encoding standard for netCDF classic files
  - Challenge: represent ragged arrays efficiently
- Classifies data according to connectedness of time/space coordinates
- Defines netCDF data structures that represent *features*
- Make it easy / efficient to
  - Store collections of features in one file
  - Read a Feature from a file
  - Subset the collection by space and time

Feature Types

- **point**: a single data point (having no implied coordinate relationship to other points)
- **timeSeries**: a series of data points at the same spatial location with monotonically increasing times
- **trajectory**: a series of data points along a path through space with monotonically increasing times
- **profile**: an ordered set of data points along a vertical line at a fixed horizontal position and fixed time
- **timeSeriesProfile**: a series of profile features at the same horizontal position with monotonically increasing times
- **trajectoryProfile**: a series of profile features located at points ordered along a trajectory

Feature Instances

```
float temp(station, time);
float humidity(station, time);
  float lon(station);
  float lat(station);
  float alt(station);
  char station_name(station, name_strlen);
  char desc(station, desc_strlen);
  float stuff( station);
  double time(time);
```

instance dimension = station

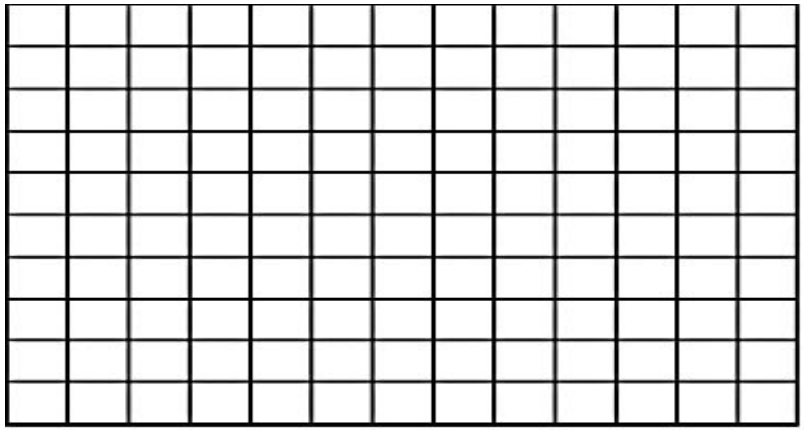
instance variables = lat, lon, alt, station_name, desc, stuff

- Instance variables provide the metadata that differentiates individual features.

Representations

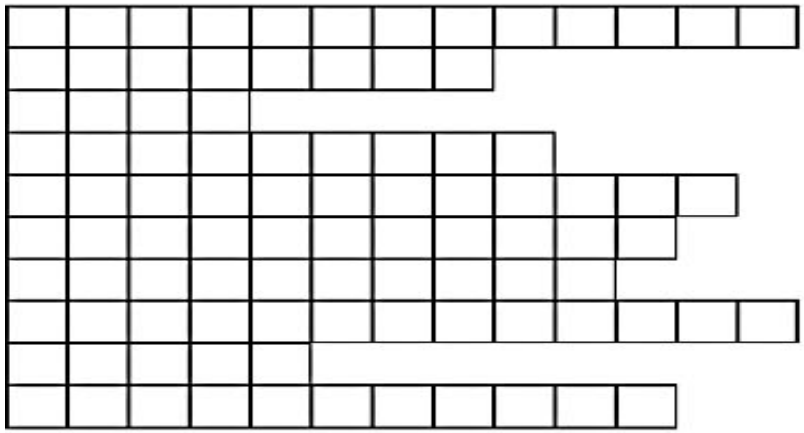
- 2 variants of multidimensional arrays
- 2 variants of ragged arrays

Rectangular Array

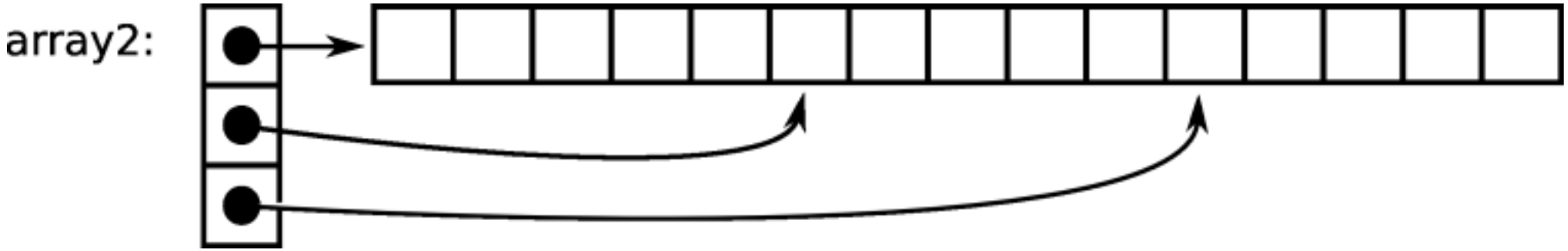


A standard two-dimensional array is a rectangle.

Ragged Array



With Variant arrays, you need not waste space.



Orthogonal multidimensional array

```
float humidity(station, time);
  float lon(station);
  float lat(station);
  float alt(station);
  float time(time);
```

The timeseries feature is connected in time, but not in space (unlike a grid), because theres no relationship between station i and i+1

Incomplete multidimensional array

```
float humidity(station, time) ;
  float lon(station);
  float lat(station);
  float alt(station);
  float time(station, time);
```

(i1, o1)	(i2, o1)	(i3, o1)	(i4, o1)
(i1, o2)	(i2, o2)	(i3, o2)	(i4, o2)
	(i2, o3)	(i3, o3)	(i4, o3)
	(i2, o4)		(i4, o4)
			(i4, o5)
			(i4, o6)

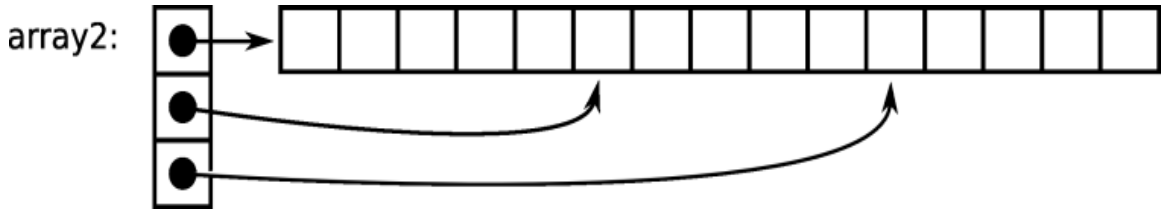
Contiguous ragged array

```
int stationId(station);
float lon(station);
float lat(station);
float alt(station);
float row_size(station);

float humidity(sample) ;
float time(sample);
```

Time series i comprises the data elements from

rowStart(i) to rowStart(i) + row_size(i) – 1, where
rowStart(i) = 0 if i = 0
rowStart(i) = rowStart(i-1)+row_size(i-1) if i > 0



Indexed ragged array

```
int stationId(station);
float lon(station);
float lat(station);
float alt(station);

float station_index(sample);
float humidity(sample) ;
float time(sample);
```

Time series i comprises the data elements from
with **station_index** = i

Status

- Now approved 42-page standard
- Implemented in netCDF-Java (CDM) library
  - “Nested Tables” configuration
  - “Point Feature Dataset” APIs
  - Feature Collections in TDS (aggregations version 2)
  - Creating data query services in TDS (alpha)
  - Starting to replace old APIs in IDV
- Well Sort of:
  - CDM based on older draft
  - Convention is more general than CDM
  - Implementation notes

Table Configurer Plugins

- BUFR
- CF Conventions
- Cosmic
- GEMPAK Point
- IRIDL station (IRI/LDEO Climate Data Library)
- Jason (NASA Ocean Surface Topography Mission)
- FSL Wind Profiler
- MADIS ACARS
- MADIS surface observations
- NBDC (National Buoy Data Center)
- NCAR-RAF/nimbus
- NLDN (National Lightning Data Network)
- Suomi-Station
- Unidata Observation Dataset Conventions

Issues

- The amount of time it takes for a complex proposal to get through the process
- Relationship of a specification to an implementation
  - Are implementations necessary?
  - Does the CDM have a special role to play?
  - Generality vs. specificity

My observations

- There are only a few in CF who have the energy for complex proposals
- Innovation in the standard is irresistible
- We've pushed the classic model as far as it should go
  - Start using extended (netcdf-4) model

- *2007 First design / implementations*
- 2007/09/11 Point Observation Draft-2
- 2007/09/10 - 2007/10/16 cf-pointobsconvention email group
- 2008/03/17 CDM Point Feature Types doc
- *2008/04 – 2008/05 implementation (nested tables)*
  - *(5 months)*
- 2008/10/20 CF Proposal #37
- 2008/10/20 - 2009/01/07 CF TRAC discussion
- 2008/12/15 - 2009/02/25 *implementation*
  - *(7 months)*
- 2009/09/22 – 2010/01/05 TRAC discussion
- 2009/10/06 Status Report at GO-ESSP Hamburg
- 2009/10/12 - 2009/11/04 *implementation*
- 2010/01 – 2011/02 hankin, caron, gregory
  - *5-6 telecons, much email, many delays*
- 2011/02/28 Consensus proposal - draft submitted
  - *(3 months)*
- 2011/05/10 Approved

What is the fate of other complex proposals?

- RADAR/LIDAR Data Format (#59)
  - Mike Dixon (NCAR)
- CF-Satellite
  - Email group
  - Proposal with SSEC
- Unstructured Grids
  - Google UGRID group
  - Other efforts