Observation, simulation and estimation: common patterns in domain models

Simon Cox

CSIRO Exploration and Mining

30 November 2006

www.csiro.au

Science relies on observations

Evidence & validation

Involves sampling

A cross-domain terminology and information-model

What is “an Observation”

Observation act involves a *procedure* applied at a specific *time and place*

Result of an observation is an estimate of some *property value*

The property is associated with the observation domain or *feature of interest*

The location of the procedure may not be the location of interest for spatial analysis of results

Observed property

Sensible phenomenon or property-type

- Length, mass, temperature, shape
- location, event-time
- colour, chemical concentration
- count/frequency, presence
- species or kind

Expressed using a reference system or scale

- Scale may also be ordinal or categorical
- May require a complex structure

“Sensible”, but not necessarily physical …

Feature-of-interest

The observed property is associated with something

- Location does not have properties,
  the substance or object at a location does
- The property must be logically consistent with the feature-type,
  as defined in the application domain
  - E.g. rock-density, pixel-colour, city-population, ocean-surface-temperature

The proximate feature-of-interest may *sample* a more meaningful domain-feature

- Rock-specimen samples an ore-body
- Well samples an aquifer
- Sounding samples an ocean/atmosphere column
- Cross-section samples a rock-unit

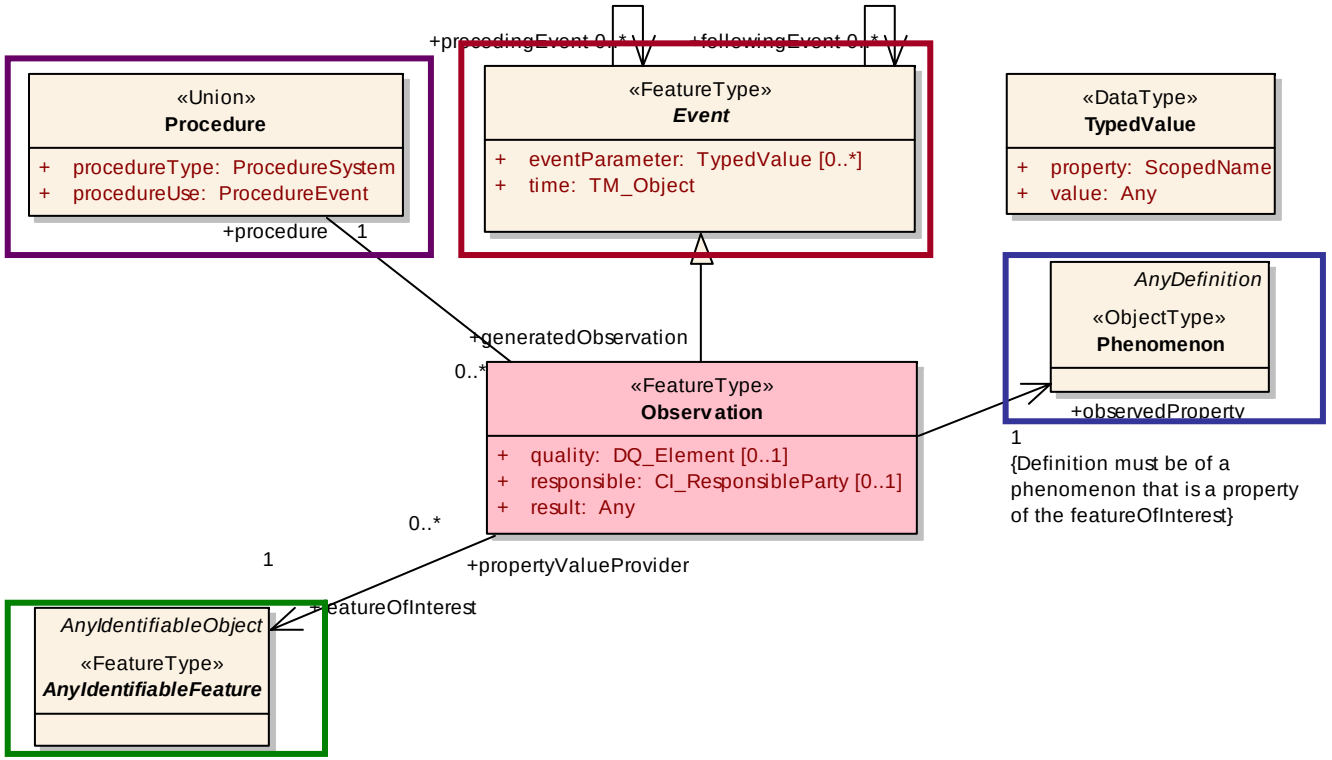
… Observation-target

Procedures

Instruments & Sensors

- Respond to a stimulus from local physics or chemistry
- Intention may concern local or remote source
- Sample may be in situ or re-located

A common pattern: the observation model



An **Observation** is an **Event** whose **result** is an **estimate** of the **value** of some **Property** of the **Feature-of-interest**, obtained using a specified **Procedure**

Procedures are usually process chains

Procedure often includes data processing, to transform "raw" data to semantically meaningful values

- Voltage ➔ orientation
- count ➔ radiance ➔ NDVI
- Position + orientation ➔ scene-location
- Mercury meniscus level ➔ temperature
- Shape/colour/behaviour ➔ species assignment

This requires consideration of "sensor"-models and calibrations

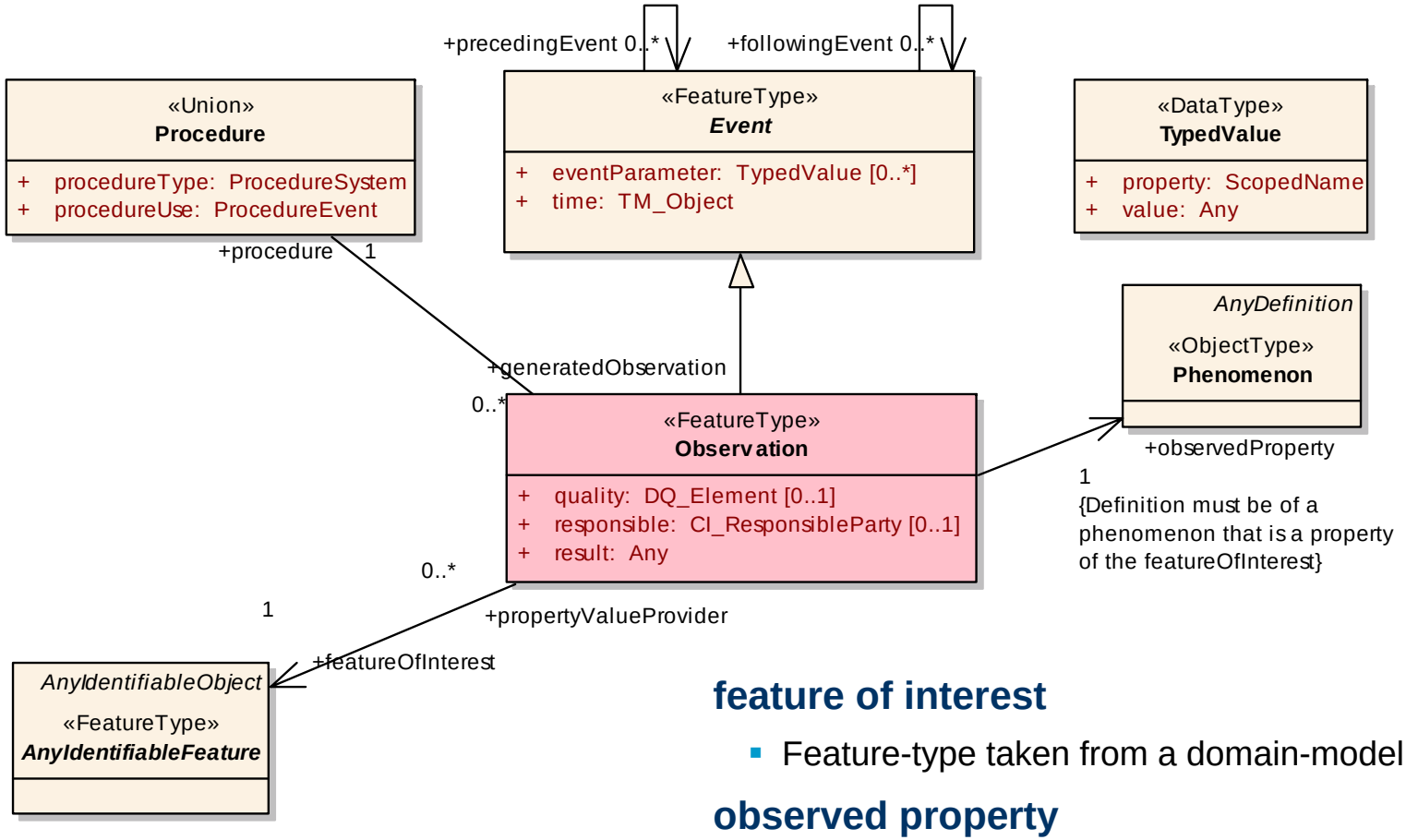
Advanced procedures

Modelling, simulation, classification are procedures

- “raw” data == modeling constraints (sensor-outputs, process-inputs)
- “processed” data == simulation results (outputs)
- “interpreted” data == classification results (outputs)

SensorML provides a model and syntax for describing process-chains

Application to a domain



feature of interest

- Feature-type taken from a domain-model

observed property

- Member of feature-of-interest-type

procedure

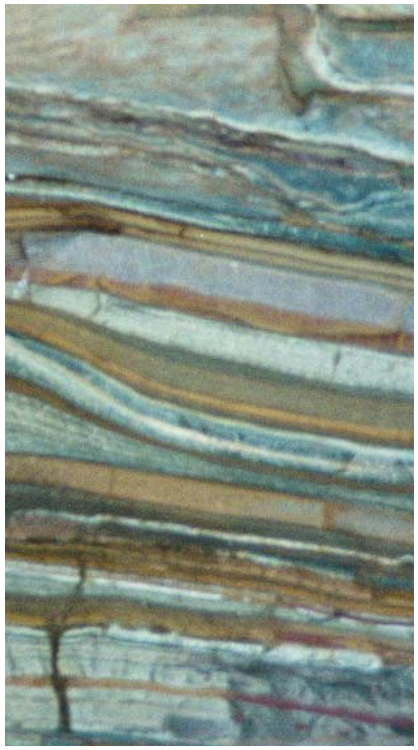
- Suitable for property-type

Conceptual object model: features

Digital object corresponding with identifiable, typed, object in the real world

- mountain, road, specimen, event, tract, catchment, wetland, farm, bore, reach, property, license-area, station

Feature-type **is characterised by a specific set of** ***properties***



Specimen

- ID (name)
- description
- mass
- processing details
- sampling location
- sampling time
- related observation
- material
- …

Geology domain model - feature type catalogue

Borehole

- **collar location**
- **shape**
- collar diameter
- length
- operator

Fault

- **shape**
- **surface trace**
- **displacement**
- age
- ...

License area

- issuer
- holder
- interestedParty
- **shape(t)**
- right(t)
- ...

Ore-body

- commodity
- deposit type
- host formation
- **shape**
- resource estimate
- ...

Conceptual classification

Multiple geometries

Water resources feature type catalogue

Aquifer

Storage

Stream

Well

Entitlement

Observation

…

Meteorology feature type catalogue

Front

Jetstream

Tropical cyclone

Lightning strike

Pressure field

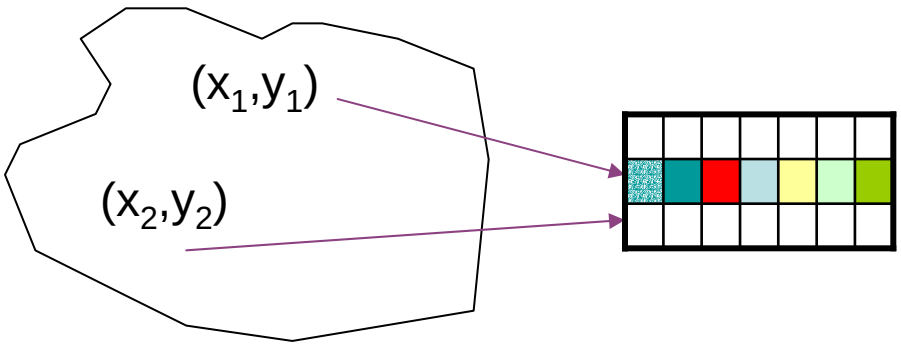
Rainfall distribution

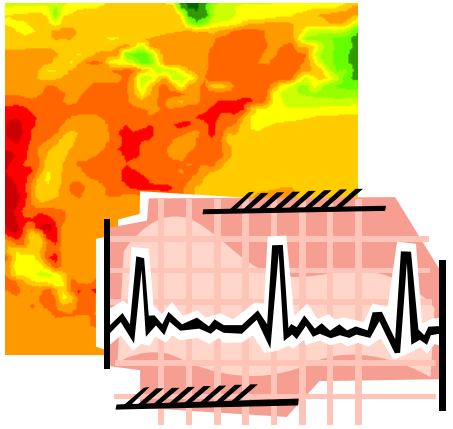
…

Bottom two are a different kind of feature

Spatial function: coverage

Variation of a property across the domain of interest





- For each element in a spatio-temporal *domain*, a value from the *range* can be determined
- Used to analyse patterns and anomalies, i.e. to detect features (e.g. storms, fronts, jetstreams)

Discrete or continuous domain

- Domain is often a grid
- Time-series are coverages over time

Features vs Coverages

Feature

- object-centric
- heterogeneous collection of properties
- “summary-view”

Coverage

- property-centric
- variation of homogeneous property
- patterns & anomalies

Both needed; transformations required

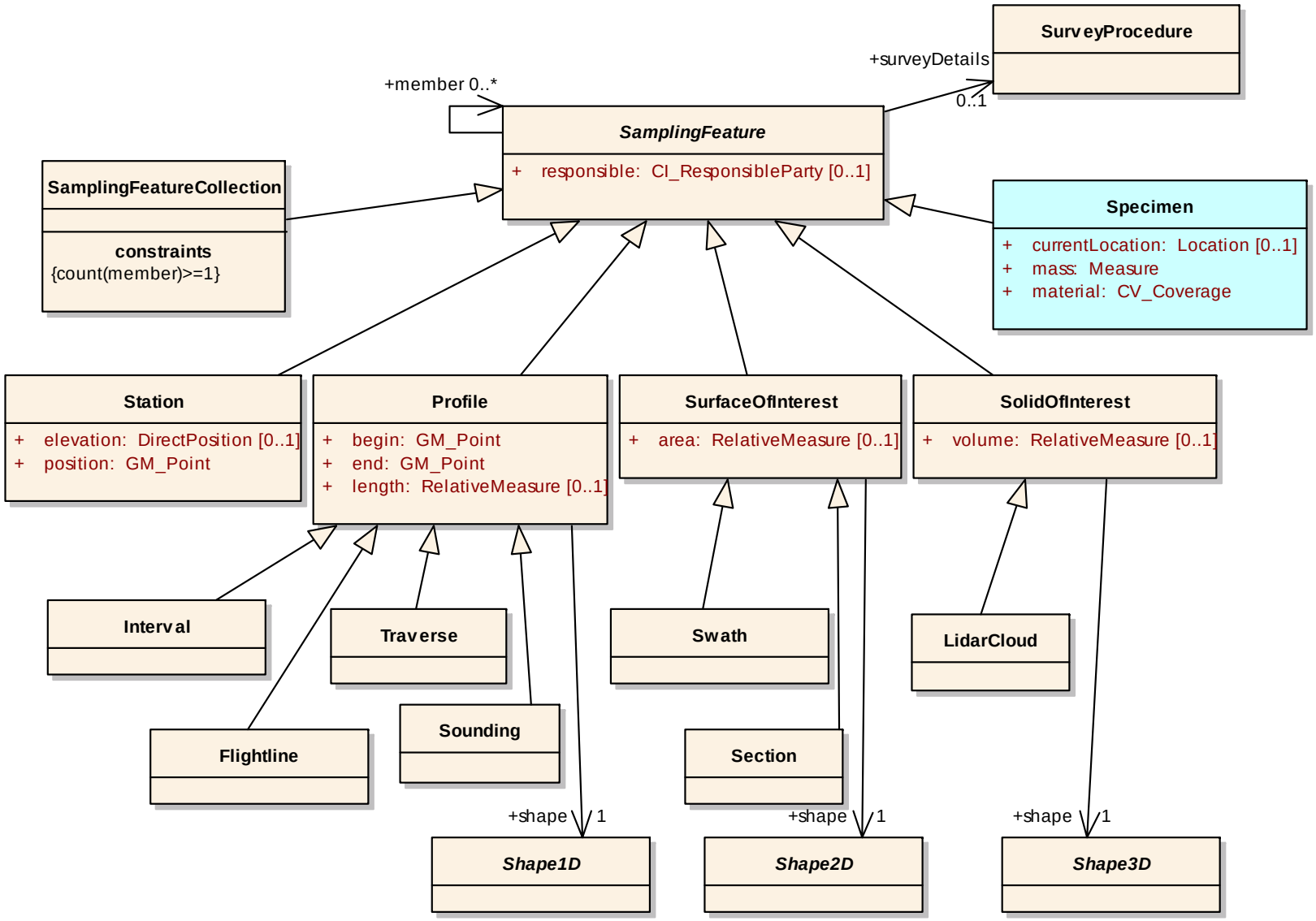
“Cross-sections” through collections

Specimen	Au (ppm)	Cu-a (%)	Cu-b (%)	As (ppm)	Sb (ppm)
ABC-123	1.23	3.45	4.23	0.5	0.34

A Row gives properties of one feature

A Column = variation of a single property across a domain (i.e. set of locations)

Some feature types only exist to support observations



Assignment of property values

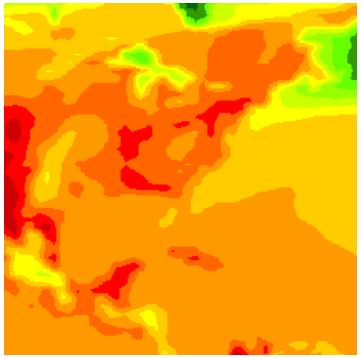
For each property of a feature, the value is either

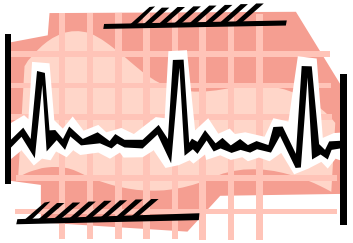
i. asserted
   - name, owner, price, boundary (cadastral feature types)

ii. estimated
   - colour, mass, shape (natural feature types)
   - *i.e. error in the value is of interest*

Variable property values

Some property values are not constant





- colour of a Scene or Swath varies with position
- shape of a Glacier varies with time
- temperature at a Station varies with time
- rock density varies along a Borehole

Variable values may be described as a *Coverage* over some axis of the feature

Observations, features and coverages

Same property on multiple samples is a another kind of coverage

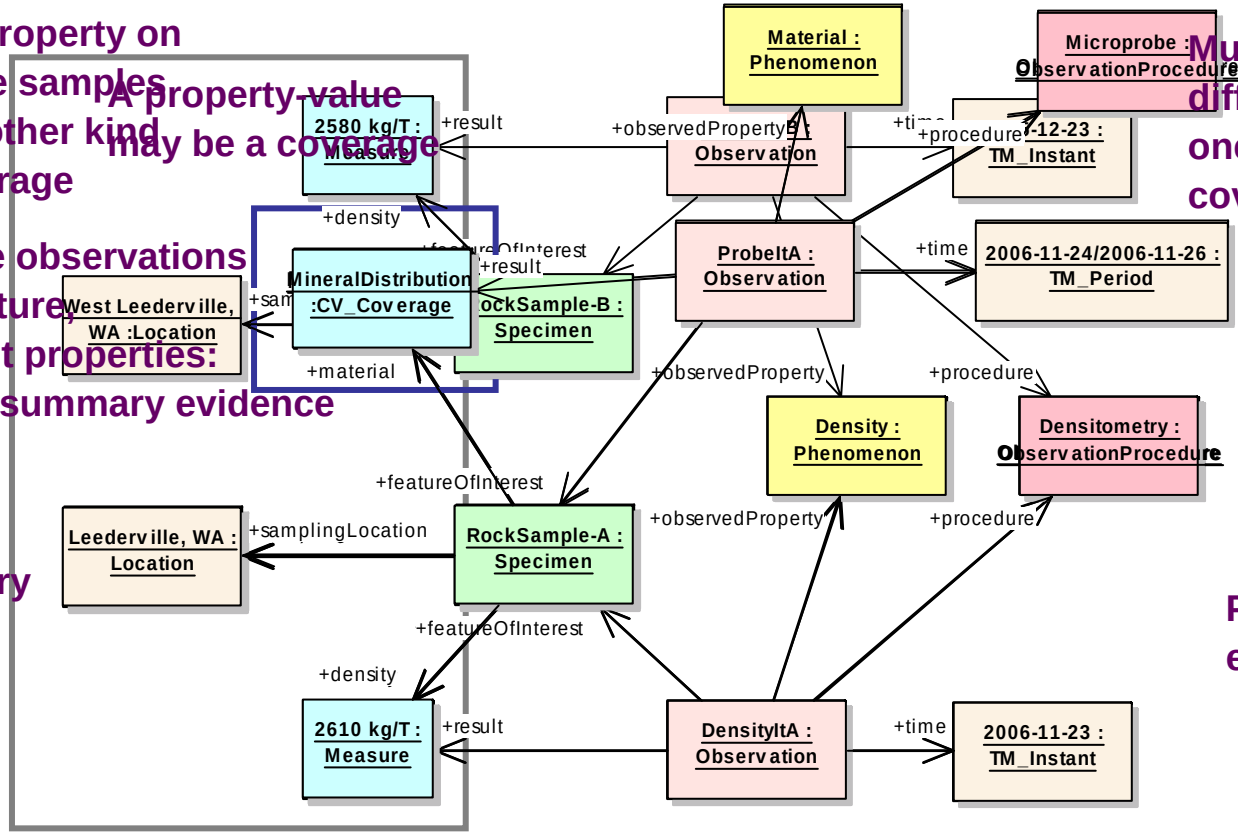
A property value may be a coverage

Multiple observations one feature, different properties: feature summary evidence

Feature summary

Multiple observations different features, one property: coverage evidence

Property-value evidence



Development and validation

O&M conceptual model and XML encoding

Developed in the context of

- XMML Geochemistry/Assay data
- OGC Sensor Web Enablement – environmental and remote sensing

Subsequently applied in

- Water resources/water quality (WQDP, AWDIP, WRON)
- Oceans & Atmospheres (UK CLRC, UK Met Office)
- Natural resources (NRML)
- Taxonomic data (TDWG)
- Geology field data (GeoSciML)

I could have put dozens of logos down here

Status

OGC Best Practice paper, r4 – 2006

OGC RFC expecting to lead to Adopted Specification – 2007

ISO specification – 2008-9?

Conclusions

Different viewpoints of same information for different purposes

- Summary vs. analysis

Some values are determined by observation

- Sometimes the description of the estimation process is necessary

Transformation between views important

Management of observation evidence can be integrated

(Bryan Lawrence issues)

For rich data processing, rich data models are needed

- Explicit or implicit

Data models (types, features) are important constraints on service specification

CSIRO Exploration and Mining

Name Simon Cox

Title Research Scientist

Phone +61 8 6436 8639

Email Simon.Cox@csiro.au

Web www.seegrid.csiro.au

Thank You

Contact CSIRO

Phone 1300 363 400

+61 3 9545 2176

Email enquiries@csiro.au

Web www.csiro.au

Features, Coverages & Observations (1)

Observations and Features

- An observation provides evidence for estimation of a property value for the feature-of-interest

Features and Coverages (1)

- The value of a property that varies on a feature defines a coverage whose domain is the feature

Observations and Coverages (1)

- An observation of a property sampled at different times/positions on a feature-of-interest estimates a discrete coverage whose domain is the feature-of-interest

- *feature-of-interest is **one big feature** – property value varies within it*

Features, Coverages & Observations (2)

Observations and Features

- An observation provides evidence for estimation of a property value for the feature-of-interest

Features and Coverages (2)

- The values of the same property from a set of features constitutes a discrete coverage over a domain defined by the set of features

Observations and Coverages (2)

- A set of observations of the same property on different features provides an estimate of the range-values of a discrete coverage whose domain is defined by the set of features-of-interest

- *feature-of-interest is **lots of little features** – property value constant on each one*

Sensor service

premises:

O&M is the high-level information model

SOS is the primary information-access interface

SOS can serve:

an *Observation* (Feature)

- getObservation == “getFeature” (WFS/Obs) operation

a *feature of interest* (Feature)

- getFeatureOfInterest == getFeature (WFS) operation

or *Observation/result* (often a time-series == discrete Coverage)

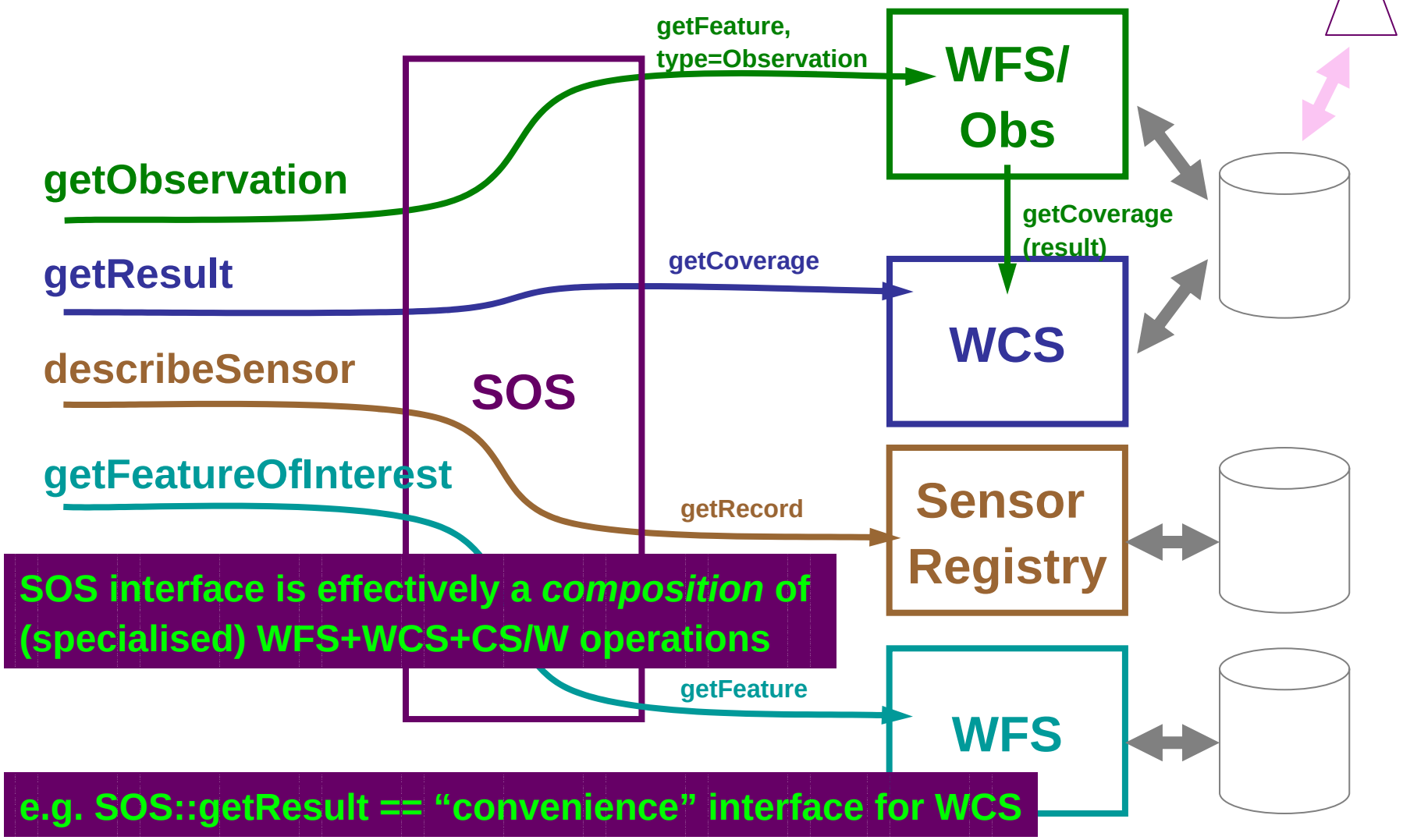
- getResult == “getCoverage” (WCS) operation

or *Sensor* == *Observation/procedure* (SensorML document)

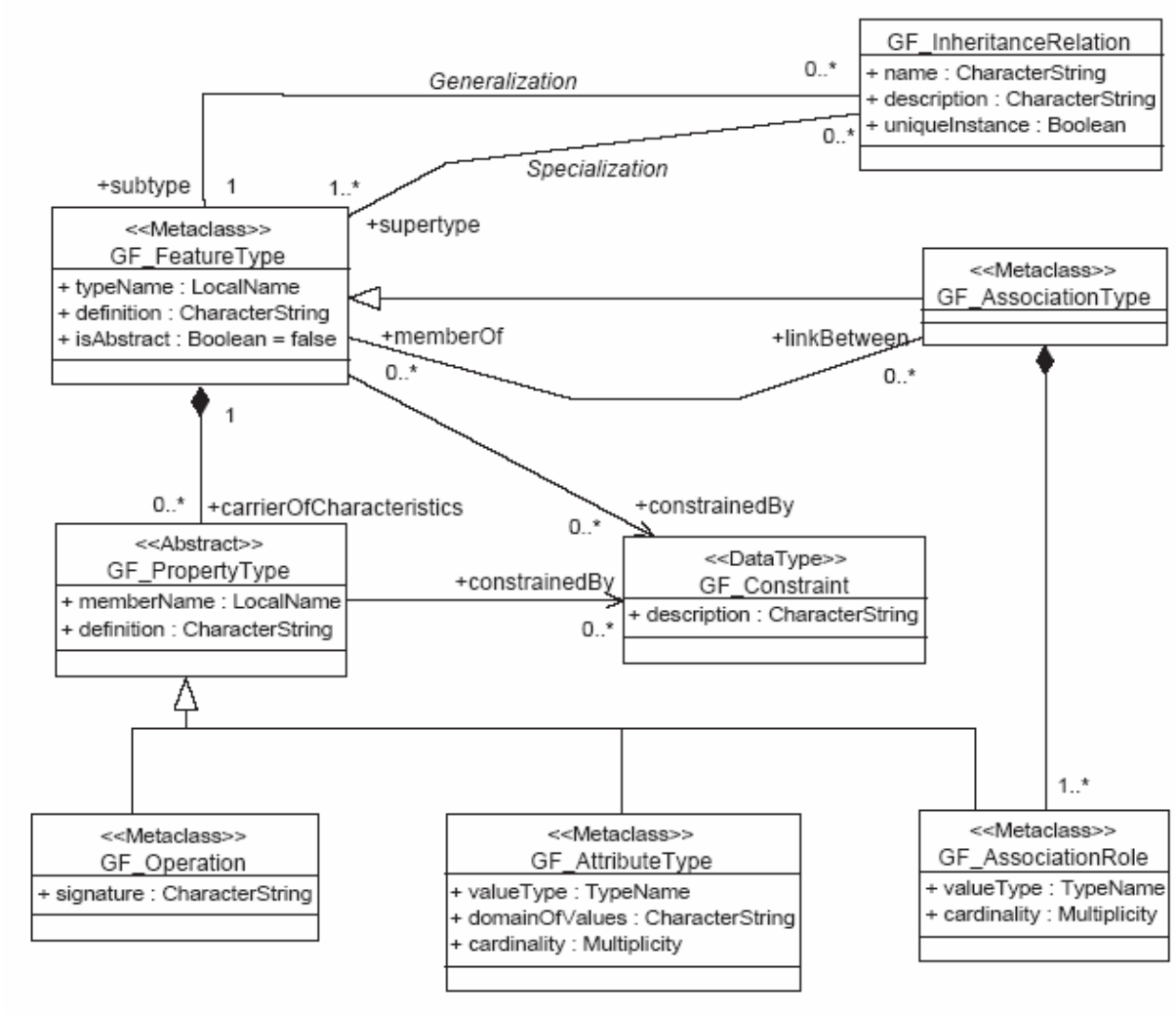
- describeSensor == “getFeature” (WFS) *or* “getRecord” (CSW) operation

optional – probably required for dynamic sensor use-cases

SOS vs WFS, WCS, CS/W?



ISO 19101, 19109 General Feature Model

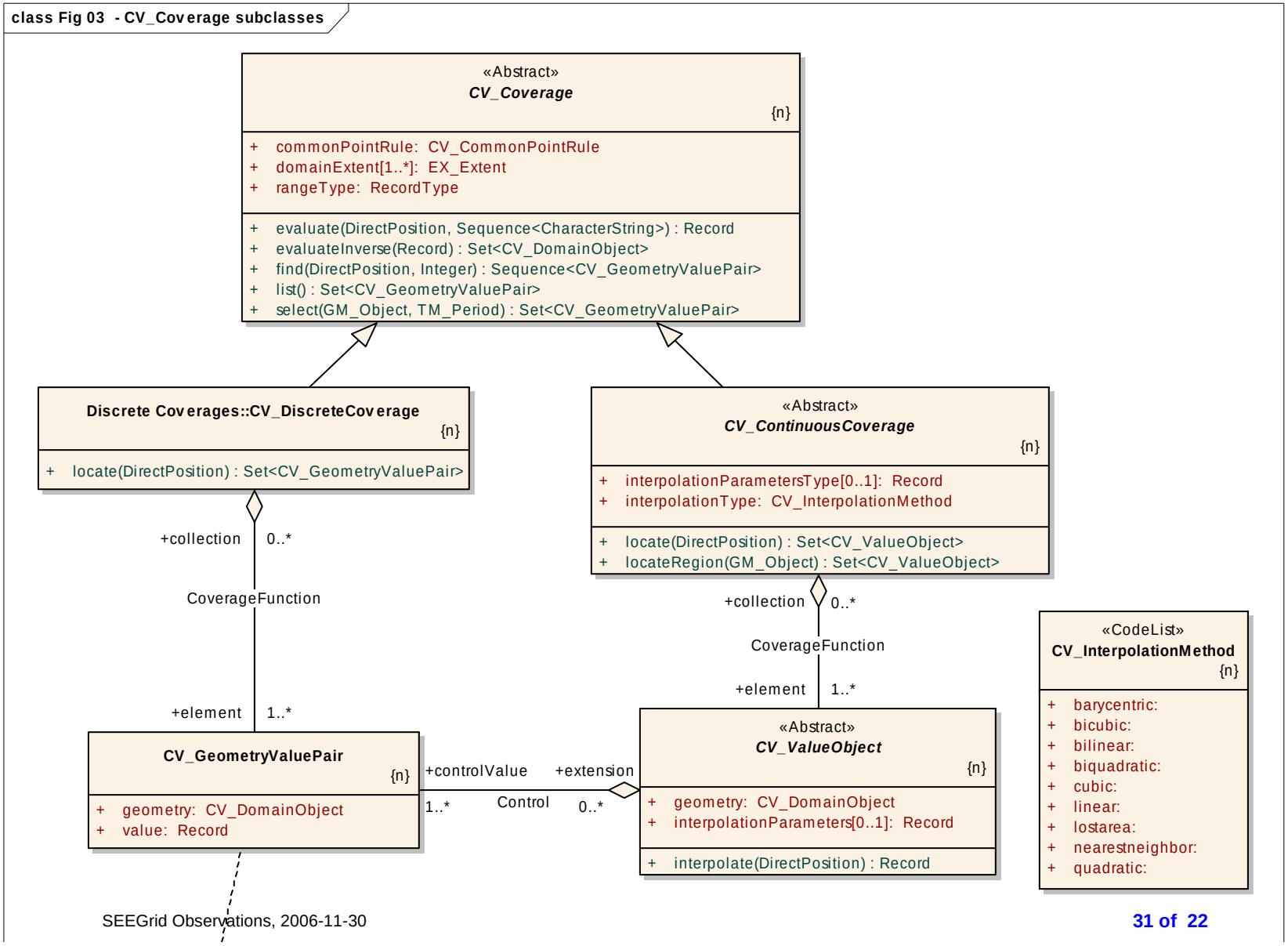


Properties include

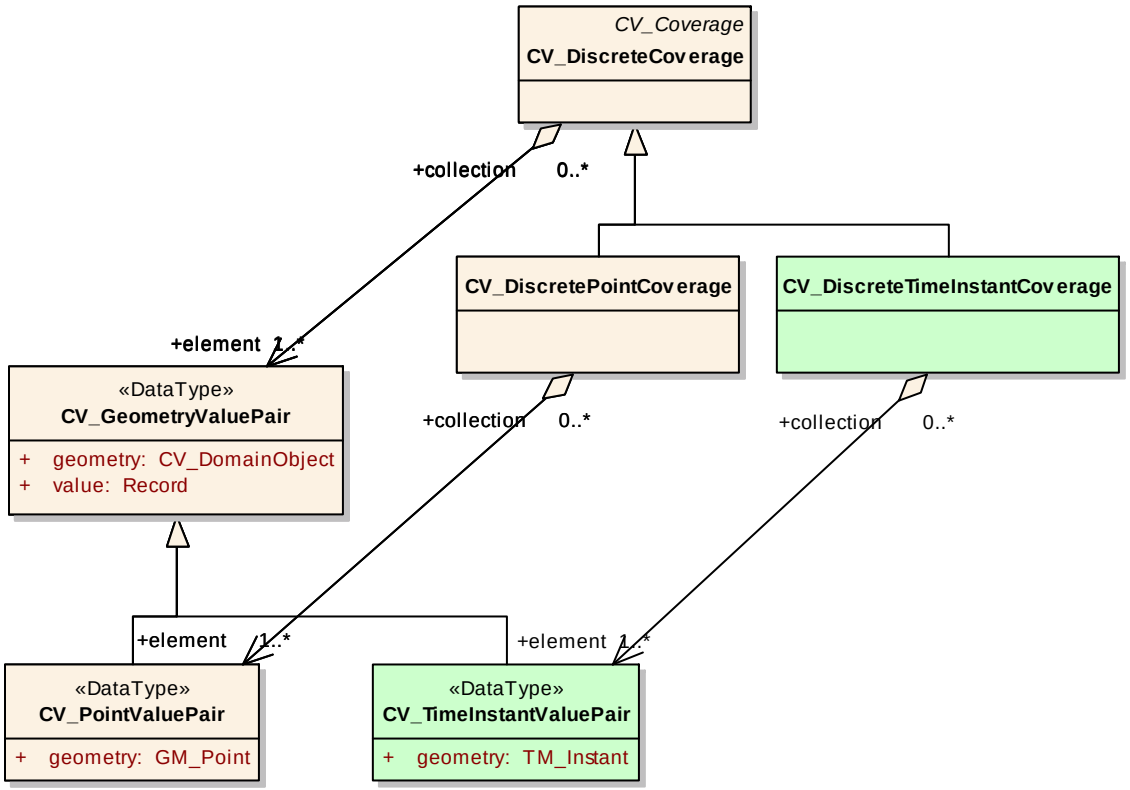
- *attributes*
- *associations* between objects
  - value may be object with identity
- *operations*

Metaclass diagram

ISO 19123 Coverage model



Discrete coverage model

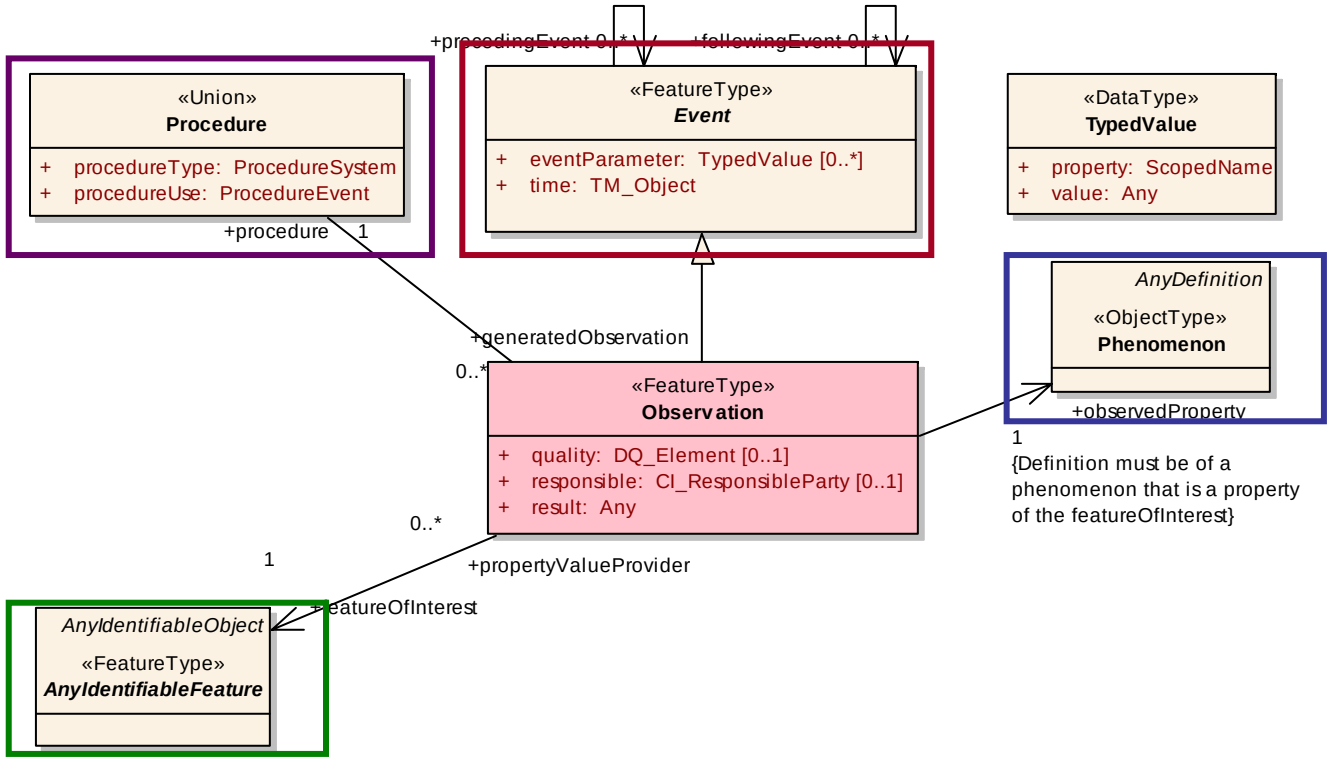


Value estimation process: observation

An Observation is a kind of “Event Feature type”, whose result is a value estimate,

and whose other properties provide metadata concerning the estimation process

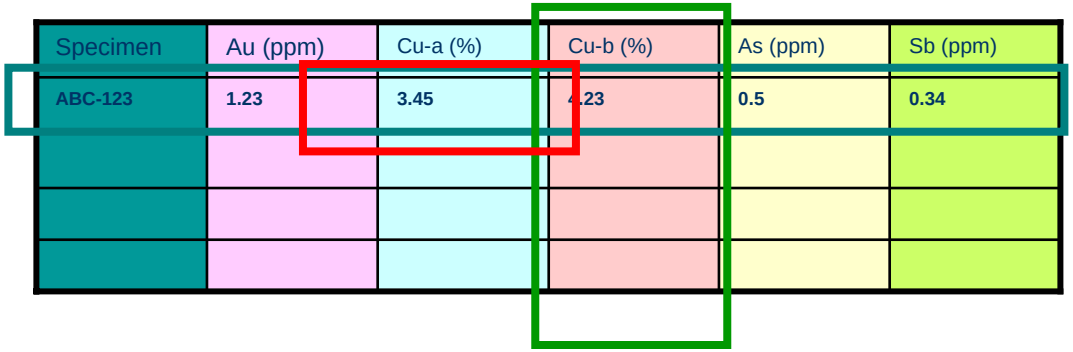
Observation model – Value-capture-centric view



An **Observation** is an **Event** whose **result** is an **estimate** of the **value** of some **Property** of the **Feature-of-interest**, obtained using a specified **Procedure**

“Cross-sections” through collections

A Cell describes the value of a single property on a feature,
often obtained by observation or measurement



A Row gives properties of one feature

A Column = variation of a single property across a domain (i.e. set of features)

Feature of interest

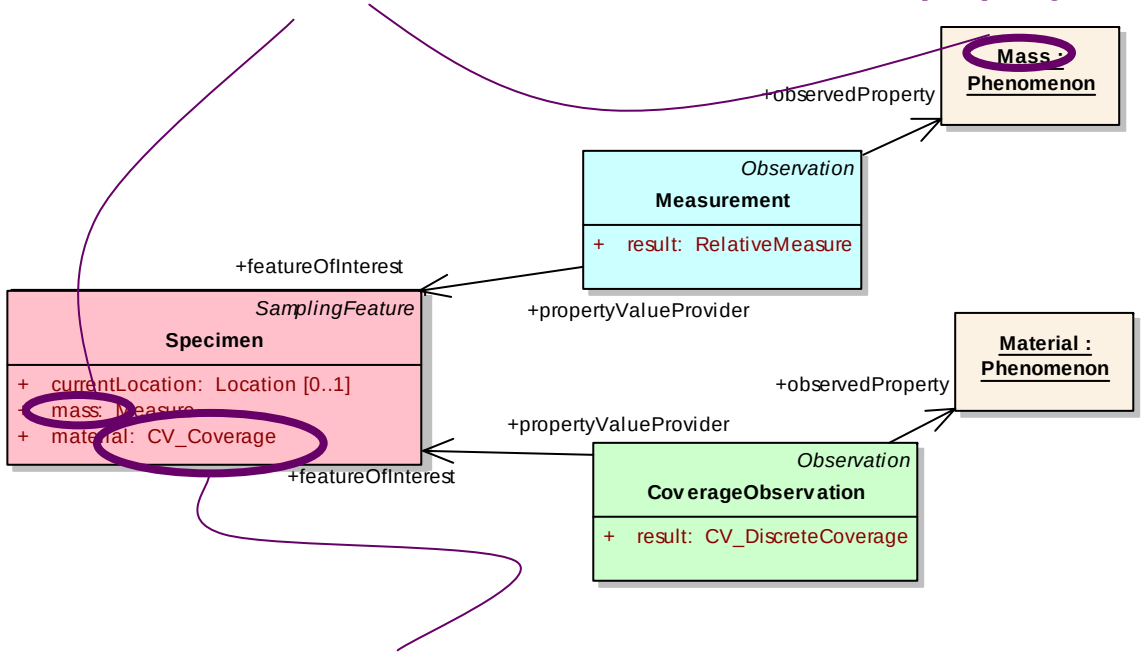
may be any feature type from any domain-model …

observations provide values for properties whose values are not asserted

i.e. the application-domain supplies the feature types

Observations support property assignment

These must match if the observation is coherent with the feature property



Some properties have interesting types …

Observations and coverages

If the *property value* is not constant across the *feature-of-interest*

- varies by location, in time

the corresponding observation *result* is a *coverage*

individual samples must be tied to the location within the domain, so result is set of e.g.

- *time-value*
- *position-value*
- *(stationID-value ?)*

Time-series observations are a particularly common use-case