Report of UCD Working Group

Roy Williams

Caltech

Summary

- UCD is inherently fuzzy
  - does not contain all metadata
- UCD is a description, not a unique name
- UCD already works
  - easy to shoot at!
- UCD is mined from large amount of metadata
  - phenomenological, not opinion
- UCD will be eventually replaced by
  "*pointers into data model*"

Why is UCD useful

- Standard names for table attributes
  - Aladin can draw based on UCD info
  - Cone search demands UCD labelling of output
  - SQL queries based on UCD? "select PHOT* from catalog"
- Searching a registry that has tables
  - Relevance of a table based on its UCDs
  - Drilling-down searches by increasing specificity
  - Registry metadata for some data types defined with UCDs
    - please make space in VOResource for UCD

Why is UCD Useful?

- Resolving the UCD to get description
  - definition, examples, units, data type, related UCDs
  - Resolver service is prerequisite for new UCDs
- Partial hierarchy
  - improve understanding
  - provide object inheritance
  - there *may* be a parent
    - PHOT_FLUX_RADIO_1.6G
    - --> PHOT_FLUX_RADIO
    - --> PHOT_FLUX

Why is UCD useful?

- Which UCD to use
  - Assigneur service
- Extensibility
  - new UCDs that can be combined with core
  - Process for integration into core
- Automatic transformations, eg units

UCD has Sequence of Specificity

- UCD is a general guide to meaning -- not infinitely specific.
- In addition to semantic meaning of UCD, additional attributes may be necessary to understand its meaning
- 1. PHOT_FLUX in general
  - The UCD can deal with this very well
- 2. PHOT_FLUX centered at a specific wavelength
  - It can deal with this but not accurately: PHOT_FLUX_RADIO_1.6G
- 3. PHOT_FLUX taken through specified filter
  - eg. expressed as a curve with 1000 points
  - Use VOTable grouping -- connection to another data object.

UCD in VOTable context

- How about a semantic net of UCD associations?
  - But groups take us a long way!

```
<TABLE>
  <GROUP ucd="PHOT_FLUX">
    <FIELD name="fluxValue"/>
    <FIELD name="errFlux"/>
    <PARAM name="Frequency" value="8.62"/>
  </GROUP>
  <FIELD name="otherThings"/>
  …..
```

Core UCD
(IVOA namespace)

- Comes from original CDS set
  - Will be pruned and rebuilt before 10/03
- Benefits of using core UCD
  - INTEROPERABILITY!
  - Software support
  - VO Tools understand core UCD
- If everyone makes their own UCD set, then we have TOWER OF BABEL
  - see for example FITS keywords!

Every UCD can be *Resolved*

- For each UCD, should be an automatic way to get:
  - Definition
  - Description
  - Related UCDs
  - Examples
  - Units ....
- Resolver service with syntax
  - http://.....?ucd=POS_RA_EQ_MAIN
  - SOAP service

Extensibility Mechanism

- Use standard UCDs when applicable, or else
  - May build a new namespace
    - For new semantics
    - in prototype
  - If useful in general, can be absorbed into core
    - Not necessary
- Namespaces for extensibility
  - Step 1. personal extension -- new namespace
  - Step 2. with documentation (resolver service)
    - well-maintained, doesn't change
  - Step 3. submitted for approval to IVOA UCD WG
  - Step 4. absorption to core
- UCD Working Group
  - committee for consultation about porting to UCD core
    - drawn from ivoa member vo

Extending the UCD set means making a resolver service

- Suppose I want a new UCD

  `ucd="PHOT_FLUX,mynamespace:PHOT_MYFILTER"`

- BaseURL comes from
  - <PARAM
    ucd="UCD_resolver"
    name="mynamespace"
    value="http://blahblah.com/UCDresolver?" />

- Now can get description of the new UCD "PHOT_MYFILTER"

General Syntax of IVOA UCD

- Detailed draft is coming
  - Proper assessment
  - Draft will be presented at IVOA meeting 10/03
- character set is
  - A-Z, a-z, 0-9, hyphen, + 3 special
  - White space is always ignored
- Three special characters

  _ underscore for hierarchy

  , comma for modifiers

  : colon for namespaces

Hierarchy

- Partial hierarchy
  - Existing special character underscore _
  - can cut from beginning to a underscore
  - there *may* be meaning in the result

PHOT_FLUX_RADIO_1.6G
PHOT_FLUX_RADIO
PHOT_FLUX

Namespaces

- Namespaces
  - Introduce special character colon :
  - Temporary scaffolding while awaiting integration to core

mynamespace:PHOT_MYFILTER

Modifiers

- Semi-atomized UCD
  - Introduce special character comma ,

    class, modifier, modifier, ...
  - adding modifiers to existing UCD, eg
    - ERROR MODEL MIN MAX MEAN MAIN OTHER RELATIVE UNIT
  - if you say storm, thunder, hail then it is meaningful
  - if you say apple, orange then you are an idiot

UCD parsing

1. Split on comma ,
   - First is main UCD, others are modifier UCDs
2. For each token, split on colon :
   - if only one, it is in core (default namespace)
   - if two, first is namespace, second is UCD
   - else fail
3. We **may** find a parent UCD by removing last underscore _ and following characters
   - And again

Photometry

- All UCDs beginning PHOT have extra metadata about bandpass
  - Can see this with "Search by wavelength" service from CDS
- Example UCDs that fit from 2 um to 5 um

  PHOT_FLUX_IR_3.5, PHOT_FLUX_IR_L, PHOT_IR_3.4, PHOT_IR_4.2, PHOT_IR_L0, PHOT_JHN_K, PHOT_JHN_L, PHOT_JHN_M, PHOT_SB_JHN_K
- Question -- Should the PHOT UCDs be rebuilt with rationalization of wavelength?
  - Answer -- No. Because it will just add more UCDs, not feasible as a replacement

UCD Working Group

- Will report precise UCD version 2
  - draft in 10/2003

Sebastien Derriere, CDS, F
Norman Gray, Starlink UK
Marco Leoni, ESO, D
Jonathan McDowell, CfA, US
Francois Ochsenbein, CDS, F
Pedro Osuna, ESA, E
Andrea Preite-Martinez, CDS, F
Guy Rixon, Cambridge, UK
Roy Williams, Caltech, US
…who else?

UCD WG Actions

- What is exact syntax?
  - in particular the class,modifier semantic
- What barriers to backward compatibility?
- Pruning UCDs
- Is the hierarchy solid?
  - If I have POS_EQ_RA, must there be POS_EQ and POS