

# UCD1+ et ontologies

Stage de Sébastien Dérivaux



## Plan

- Les UCD1+
  - nouveautés par rapport aux UCD1
  - lien avec les ontologies
- Objectifs du stage
- Réalisations et discussion



## Les nouveaux UCD (UCD1+)

- Description sémantique **simple**, entre:
  - la complexité du langage naturel
  - la rigidité d'un modèle de données
- Mise au point d'un **vocabulaire**
  - définition associée à chaque mot
- Possibilité d'associer **plusieurs mots** pour former un **UCD1+**
  - phot.mag;em.opt.B
  - phys.size;instr.tel



## Les nouveaux UCD (UCD1+)

L'esprit reste le même qu'avec les UCD1:

- décrire des quantités (*qu'est-ce que c'est?*)
- avec un niveau de détail **raisonnable**
- pour pouvoir faire des comparaisons entre des ensembles hétérogènes (interopérabilité)

Mais nouvelle syntaxe...

<http://cdsweb.u-strasbg.fr/UCD/>



## Les nouveaux UCD (UCD1+)

On peut ré-utiliser un mot dans plusieurs UCD:

- phot.mag;em.opt.R
- phot.mag;em.opt.B

Plus de souplesse et de précision que les UCD1:

- ERROR
- stat.error;phot.mag;em.opt.R

Le **premier mot** porte l'essentiel du sens !

## UCD1+ vs Ontologies

Une ontologie :

- permet la description formelle d'un domaine, compréhensible par les spécialistes du domaine et les programmes;
- rend explicite ce qui est habituellement implicite;
- résulte d'une approche consensuelle;
- facilite l'interopérabilité entre les outils.

... des points communs avec les motivations pour avoir des UCD !

## Ontologie:


- **concepts**, comparables aux classes OO
- un concept possède des **propriétés** (slots)
- **facets**, qui sont des contraintes
  - › étoile n'est pas galaxie

Une propriété peut être commune à différents concepts. ex: **température** d'un **instrument**

## UCD1+:

- essentiellement description des **propriétés**
  - phys.temperature
  - phys.area
  - phot.flux
- mais aussi certains **concepts** dans le vocabulaire (instr, pos.ee)

## Le stage

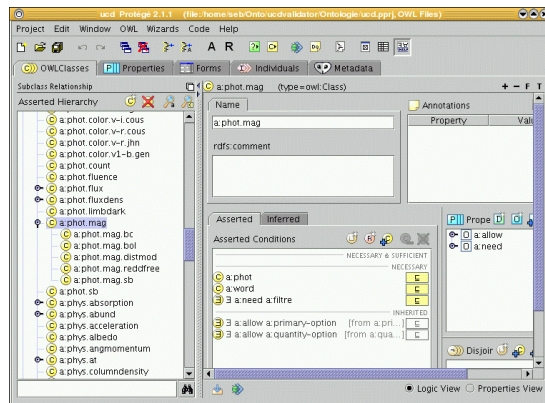
- 2 mois (juillet/août 2004)
- Utilisation de Protégé pour l'édition de l'ontologie 
- Encadrement A. Napoli + S. Derriere
- Documents sur le wiki:
  - <http://cds.u-strasbg.fr/twikiMDA/bin/view/Ontologies/StageOntoUCD2004>

## Objectifs initiaux

- Construction/validation d'un UCD
  - les UCD1+ sont composés de plusieurs mots: comment les choisir/assembler?
  - déterminer si un UCD est valide
- Trouver des concepts similaires
  - repose sur une connaissance du domaine
  - UCD → concept → concepts sim. → autres UCD
- Comparer/convertir des quantités
  - ex: convertir fréquence en longueur d'onde...
  - intégration de formules mathématiques?

## Validation d'UCD

- Stockage de règles syntaxiques dans l'ontologie
  - définition de catégories de mots (filtres, quantités, quantités\_photométriques, ...)
  - relations: phot.mag need Filtre
- Script de chargement de la liste des mots (+400 mots!) UCD dans Protégé
  - parseur avec javaCC
  - exportation de l'ontologie au format OWL



## Validation d'UCD

- Prototype:
  - Chargement avec OWLapi dans serveur java
  - Interface Web pour remplir le formulaire
  - TCP/IP
- Fonctionne pour les cas simples
- Problème de gestion des listes par le raisonneur
  - si [quantity allow a] et [quantity allow b], valide-t-on quantity;a;b ou quantity;b;a ?

## Classification par formulaire

- Aider un utilisateur à classer des concepts de façon interactive
- Prototype:
  - Tomcat, java/jsp
  - Librairie Jena (OWL et Racer)
- Formulaire construit dynamiquement:
  - à partir d'une ontologie (upload)
  - au fur et à mesure des choix de l'utilisateur

## Classification par formulaire

- Utilise le raisonneur pour trouver les concepts similaires, ou ceux subsumés
- Ontologie correspondante lourde à construire (revient à redéfinir tous les UCD)
- Problème récurrent: impossibilité de poser des contraintes **quantitatives** (numériques). Seules des contraintes qualitatives sont faisables.

## Besoin de maths...

- Prenons le découpage du spectre EM:
  - em.opt.B = [400-500nm] = [600-750THz]
  - em.opt.V = [500-600nm] = [500-600THz]
- Comment inférer ?
  - $\lambda = 445\text{nm} \rightarrow \text{em.opt.B}$
- Intégrer des formules mathématiques comme contraintes?

## Besoin de maths...

- Pas du tout possible avec OWL actuellement
- La question est à l'étude
- Pour l'instant: solution envisagée consisterait à stocker les formules de correspondance entre les UCD dans des commentaires + interprétation par un langage fonctionnel (hors ontologie/raisonneur)

## Conclusion

- Bon état de l'art... limitations actuelles de OWL pour faire de la science (perte par rapport aux ontologies à la LISP!)
- Intérêt de l'ontologie pour formaliser les règles de construction des UCD1+
- Besoin d'un outil de validation...
- Qq corrections dans le vocabulaire