



Object Types Validation in SIMBAD

Alexandre RICHARD (INAF)

Sébastien DERRIERE (CDS)

Andrea PREITE MARTINEZ (INAF)

Overview

- Scope and method
- SIMBAD object types
- Ontology of astronomical object types
- SIMBAD consistency checker

Scope and method

- General goal: exploring the possibilities offered by formal ontologies in the Astronomy field
 - Build ontology from scratch on a simple enough case
 - Develop application use-cases
- An ontology of astronomical object types
 - Object types are a well-known, well-sized field with potential ontology-based applications
 - SIMBAD's object types list as starting point
 - Defined concepts and use of automated reasoning

SIMBAD Object Types

SIMBAD object types (1/2)

- SIMBAD
 - Reference database of astronomical objects
 - 3.9 million objects
 - Identifiers, bibliography, measurements, otypes
- SIMBAD object types (*otypes*)
 - 198 different otypes
 - Covering most of the domain with a sensible covering factor (e.g. QSO exists, but not QSO1 nor QSO2)
 - Organized in a simple hierarchy
 - Used to annotate individual objects



CCDM J20137+1609AB -- Double or multiple star

Other object types:

*

(AG, BD, GC, GSC, HD, HIC, HIP, PPM, SAO, TYC, uvby98, YZ)

, ** (CCDM, **, ADS, IDS) , **IR** (2MASS) , **UV** (TD1)

SIMBAD object types (2 / 2)

12.00.00.00	multiple_object	mul	Composite object
12.01.00.00	· Region	reg	Region defined in the sky
12.01.05.00	· · Void	vid	Underdense region of the Universe
12.02.00.00	· SuperCIG	SCG	Supercluster of Galaxies
12.03.00.00	· CIG	CIG	Cluster of Galaxies
12.04.00.00	· GroupG	GrG	Group of Galaxies
12.04.05.00	· · Compact_Gr_G	CGG	Compact Group of Galaxies
12.05.00.00	· PairG	PaG	Pair of Galaxies
12.05.05.00	· · IG	IG	Interacting Galaxies
12.10.00.00	· GICl?	GI?	Possible Globular Cluster
12.11.00.00	· Cl*	Cl*	Cluster of Stars
12.11.01.00	· · GICl	GIC	Globular Cluster
12.11.02.00	· · OpCl	OpC	Open (galactic) Cluster
12.12.00.00	· Assoc*	As*	Association of Stars
12.13.00.00	· **	**	Double or multiple star
12.13.01.00	· · EB*	EB*	Eclipsing binary
12.13.01.01	· · · EB*Algol	Al*	Eclipsing binary of Algol type
12.13.01.02	· · · EB*betLyr	bL*	Eclipsing binary of beta Lyr type
12.13.01.03	· · · EB*WUMa	WU*	Eclipsing binary of W UMa type
12.13.01.08	· · · EB*Planet	EP*	Star showing eclipses by its planet
12.13.02.00	· · SB	SB*	Spectroscopic binary
12.13.11.00	· · CataclyV*	CV*	Cataclysmic Variable Star
12.13.11.02	· · · DQHer	DQ*	Cataclysmic Var. DQ Her type
12.13.11.03	· · · AMHer	AM*	Cataclysmic Var. AM Her type
12.13.11.05	· · · Nova-like	NL*	Nova-like Star
12.13.11.06	· · · Nova	No*	Nova
12.13.11.07	· · · DwarfNova	DN*	Dwarf Nova
12.13.12.00	· · XB	XB*	X-ray Binary
12.13.12.02	· · · LMXB	LXB	Low Mass X-ray Binary
12.13.12.03	· · · HMXB	HXB	High Mass X-ray Binary

Ontology of Astronomical Object Types

Feb 18th, 2008

Ontology overview

- Contents:
 - Concept-centered: 367 concepts, 108 defined
 - 303 restrictions on 27 properties
 - 13 annotations
- Implementation:
 - Formalized using description logic SHOIQ(D)
 - OWL 1.1 (extension of OWL-DL)
 - Could currently be reduced to SHIQ / OWL-DL
- Tools:
 - Protégé editor and Protégé-OWL API
 - Automated reasoning (with Pellet)

Ontology construction

- Study to determine the technologies to use
- First ontology version directly derived from SIMBAD
- IVOA Technical Note (Feb.19, 2007)
- Second version enhanced with domain experts
 - AGN: Paolo Padovani (ESO)
 - YSO, diffuse matter: Laurent Cambrésy (CDS)
- Real world-linkable information added
 - Measurements
 - Keywords from databases and registries
- Use-case: SIMBAD consistency checker

Ontology highlights

- **Annotations** help keep additional information centralized in the ontology without a reasoning load increase
 - e.g. `simbad:shortCode`, `vizier:kwd`, `interface:naturalName`
- **Restrictions** neither necessary nor sufficient on objects usually expressed backwards to include the knowledge without over-constraining the concepts
 - `VariableObject hasMeasurement some VariabilityPeriod` : too strong
 - `VariabilityPeriod isMeasuredFor only VariableObject` : better

Protégé concept view

The screenshot shows the Protégé 3.4 beta interface with the following components:

- Window Title:** ObjectTypes_0.99b Protégé 3.4 beta (file:/Applications/Onto/OWLfiles/ObjectTypes_0.99b.pprj, OWL / RDF Files)
- Menu Bar:** File, Edit, Project, OWL, Reasoning, Code, Tools, Window, Help
- Toolbar:** Standard Protégé navigation and editing tools.
- Subclass Explorer:**
 - Project: ObjectTypes_0.99b
 - Root: owl:Thing
 - Class: AstrObject
 - CompoundObject
 - AGN
 - AssociationOfStars
 - ClusterOfGalaxies
 - ClusterOfStars
 - GalaxiesGroup
 - Galaxy
 - GalaxyPair
 - LINER
 - MultipleStar
 - DoubleStar
 - AccretingWhiteDwarf
 - CataclysmicVariable
 - EclipsingBinary** (selected)
 - EclipsingBinaryAlgol
 - EclipsingBinaryBetaLyr
 - EclipsingBinaryWUMa
 - EllipsoidalVariableStar
 - SpectroscopicBinary
 - SymbioticStar
 - XRayBinary
 - PlanetaryNebula
 - SuperClusterOfGalaxies
 - EMSource
 - InterStellarMedium
 - StellarObject
 - SubStellarObject
 - SupermassiveBlackHole

- Class Editor:**
- Class: EclipsingBinary (instance of owl:Class) [Inferred View]
- Table of Properties and Values:

Property	Value	Type	Lang
gcvs:code	E	string	
interface:naturalName	Eclipsing Binary	string	
interface:searchKeywords	eclipsing#binary#star	string	
misc:description	Binary systems with orbital planes so close to the observer's L...	string	
owl:equivalentClass	DoubleStar and (hasProcess some Eclipse)	owl:Class	
rdfs:comment	Eclipsing Binary	string	
rdfs:subClassOf	DoubleStar	owl:Class	
rdfs:subClassOf	VariableStellarObject	owl:Class	
simbad:name	EB*	string	
simbad:shortCode	EB*	string	
vizier:kwd	Binaries:eclipsing	string	
- Asserted Conditions:**
- DoubleStar hasProcess some Eclipse (NECESSARY & SUFFICIENT)
- VariableStellarObject (NECESSARY)
- StellarObject or MultipleStar (INHERITED)
- hasComponent only StellarObject (from VariableStellarObject)
- hasComponent exactly 2 owl:Thing (from MultipleStar)
- hasConstituent only AstrObject (from DoubleStar)
- hasConstituent min 2 (from CompoundObject)
- hasProcess some (Explosion or Eclipse or SupernovaExplosion or Pulsation or Rotation) (from CompoundObject)
- Disjoints:** (Empty section)
- View Modes:** Logic View (selected), Properties View

Protégé property view

The screenshot shows the Protégé 3.4 beta interface. The title bar indicates the file path: `file:/Applications/Onto/OWLfiles/ObjectTypes_0.99b.pprj, OWL / RDF Files`. The menu bar includes File, Edit, Project, OWL, Reasoning, Code, Tools, Window, and Help. The toolbar contains various icons for file operations and navigation.

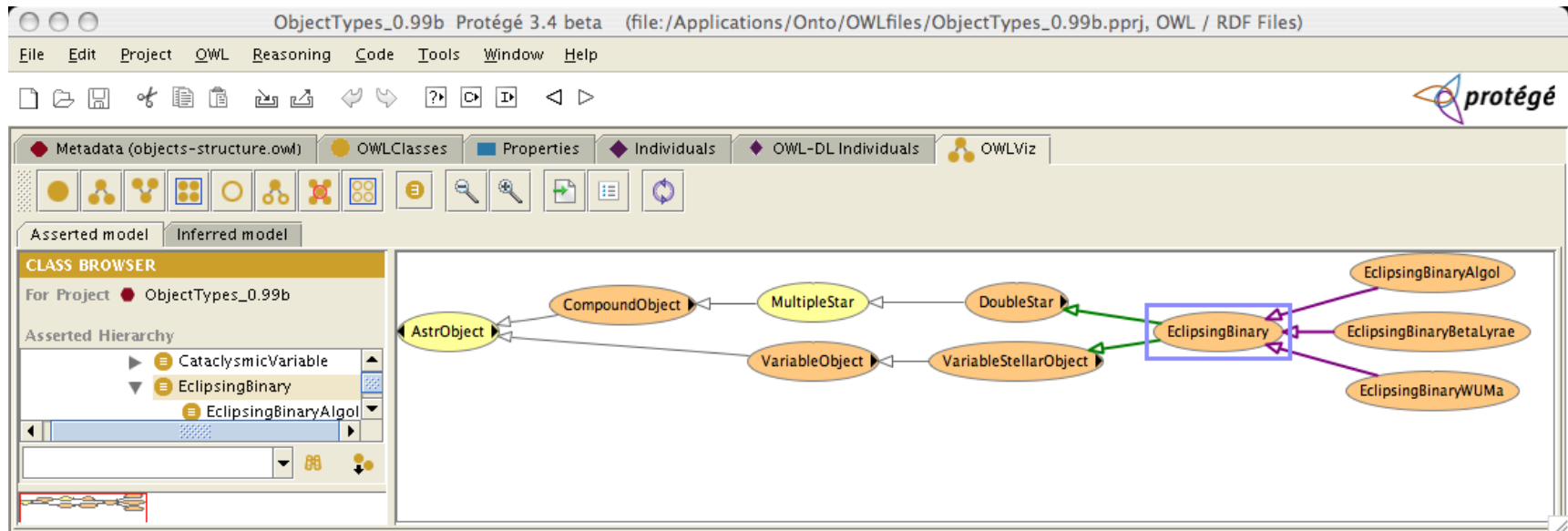
The main workspace is divided into several panes:

- PROPERTY BROWSER:** Shows a tree view of properties for the project 'ObjectTypes_0.99b'. The 'hasComponent' property is selected under the 'hasConstituent' hierarchy.
- PROPERTY EDITOR:** Displays the configuration for the 'hasComponent' property. It includes a table of associated properties and their values.

Property	Value	Type	Lang
interface:naturalName	has among its (direct) components	string	
owl:inverseOf	isComponentOf ↔ hasComponent	owl:ObjectProperty	
rdfs:comment	Indicates that the DOMAIN CompoundObject has the f...	string	
rdfs:domain	CompoundObject	owl:Class	
rdfs:range	AstrObject	owl:Class	
rdfs:subPropertyOf	hasConstituent ↔ isConstituentOf	owl:ObjectProperty	

Below the table, the 'Domain' is set to 'CompoundObject' and the 'Range' is set to 'AstrObject'. On the right, there are checkboxes for property characteristics: Functional, InverseFunctional, Symmetric, and Transitive. An 'Inverse' section shows 'isComponentOf' as the inverse property.

Protégé graph visualization



SIMBAD Consistency Checker

Feb 18th, 2008

SIMBAD objects, multiple otypes (1/2)

- Each object has associated otypes
- Some identifiers correspond to catalogs having an otype
- Each object detected in a catalog inherits the catalog's dominant otype

SIMBAD objects, multiple otypes (2/2)



SIMBAD query result



[CDS](#) · [Simbad](#) · [VizieR](#) · [Aladin](#) · [Catalogues](#) · [Nomenclature](#) · [Biblio](#) · [Tutorial](#) · [Developer's corner](#)

other query modes : [Identifier query](#) [Coordinate query](#) [Criteria query](#) [Bibliography query](#) [Script submission](#) [Output options](#) [Help](#)

Object query : M31

C.D.S. - SIMBAD4 rel 1.070 - 2008.02.08CET16:15:37

[Available data](#) [Basic data](#) [Identifiers](#) [Plot & images](#) [Bibliography](#) [Measurements](#) [External archives](#) [Notes](#)

Basic data :

M 31 -- LINER-type Active Galaxy Nucleus

query around with radius arcmin

Other object types:

LIN () , G (LEDA, 2MASX, MCG, UGC, Z, [M98c]) , AGN ([VV2000c], [VV2003c], [VV98c]) , Rad (2C, DA, [DGW65]) , IR (IRAS, IRC, RAFGL) , GiC (GIN) , GiG (K79) , QSO ([VV2006]) , X (XSS)

ICRS coord. (ep=2000 eq=2000) : 00 42 44.31 +41 16 09.4 (~Unknown) [10800.00 10800.00 90] D [1999ApJS..125..409C](#)

FK5 coord. (ep=2000 eq=2000) : 00 42 44.31 +41 16 09.4 (~Unknown) [10800.00 10800.00 88] D [1999ApJS..125..409C](#)

FK4 coord. (ep=1950 eq=1950) : 00 40 00.07 +40 59 43.6 (~Unknown) [10800.00 10800.00 0] D [1999ApJS..125..409C](#)

Gal coord. (ep=2000 eq=2000) : 121.1743 -21.5728 (~Unknown) [10800.00 10800.00 72] D [1999ApJS..125..409C](#)

Radial velocity / Redshift / cz : km/s -301 [7] / z -0.001004 [0.000023] / cz -300.99 [6.90] D [2002LEDA.....0P](#)

Morphological type:

Sb D ~

Angular size (arcmin):

186.2 61.65 35 (~) D ~

Fluxes (4) :

B 4.3 [~] E ~

J 2.094 [0.016] C [2006AJ....131.1163S](#)

H 1.283 [0.017] C [2006AJ....131.1163S](#)

K 0.984 [0.017] C [2006AJ....131.1163S](#)

Consistency checker goals (1/2)

- Consistency check of an object's otypes WRT ontology
- Consistency check of measurements WRT otypes
- If inconsistent, indicate the cause(s) of the problem
- If consistent, suggest possibly missing otypes

Consistency checker goals (2/2)

Basic data :

M 31 -- LINER-type Active Galaxy Nucleus

query around with radius arcmin

Other object types: **LIN** () , **G** (LEDA,2MASX,MCG,UGC,Z,[M98c]) , **AGN** ([VV2000c],[VV2003c],[VV98c]) , **Rad** (2C,DA,[DGW65]) , **IR** (IRAS,IRC,RAFGL) , **GiC** (GIN) , **GiG** (K79) , **QSO** ([VV2006]) , **X** (XSS)

ICRS coord. (ep=2000 eq=2000) : 00 42 44.31 +41 16 09.4 (~Unknown) [10800.00 10800.00 90] D [1999ApJS..125..409C](#)

FK5 coord. (ep=2000 eq=2000) : 00 42 44.31 +41 16 09.4 (~Unknown) [10800.00 10800.00 88] D [1999ApJS..125..409C](#)

FK4 coord. (ep=1950 eq=1950) : 00 40 00.07 +40 59 43.6 (~Unknown) [10800.00 10800.00 0] D [1999ApJS..125..409C](#)

Gal coord. (ep=2000 eq=2000) : 121.1743 -21.5728 (~Unknown) [10800.00 10800.00 72] D [1999ApJS..125..409C](#)

Radial velocity / Redshift / cz : km/s -301 [7] / z -0.001004 [0.000023] / cz -300.99 [6.90] D [2002LEDA.....0P](#)

Morphological type: Sb D -

Angular size (arcmin): 186.2 61.65 35 (~) D -

Fluxes (4) : B 4.3 [-] E ~

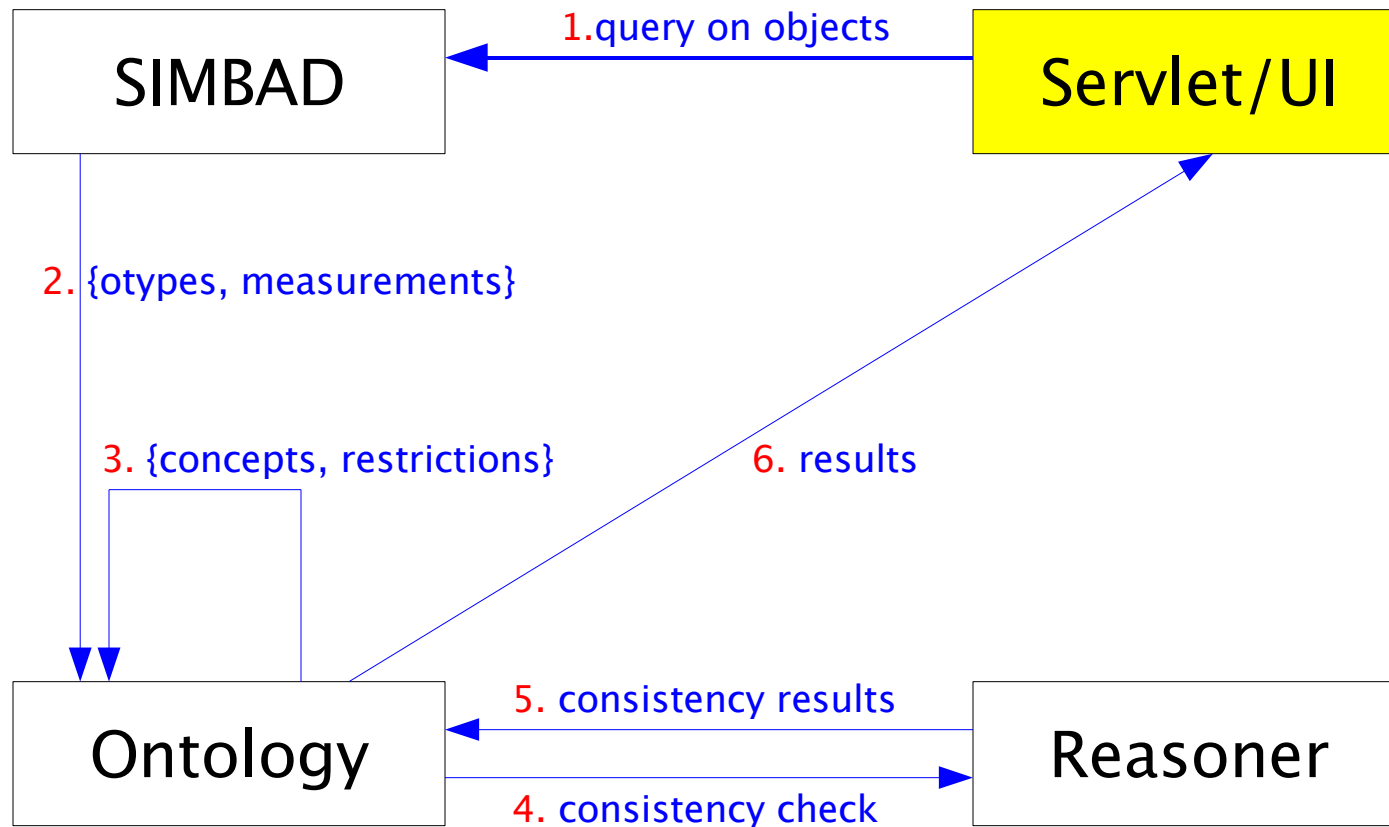
J 2.094 [0.016] C [2006AJ....131.1163S](#)

H 1.283 [0.017] C [2006AJ....131.1163S](#)

K 0.984 [0.017] C [2006AJ....131.1163S](#)

LIN && G && AGN && Rad && IR && GiC && GiG && QSO && X
=> Consistent?

SIMBAD consistency checker



Consistency results (1/2)

M_31
 CSO_619
 IRAS_20266+3856

Draw

INCONSISTENT ITEMS

nb of items: 1:

[CSO_619](#)

CONSISTENT ITEMS

nb of items: 2:

[IRAS_20266+3856](#)

[M_31](#)

M_31

OTYPES: [G, GiC, GiG, AGN, LIN, QSO, Rad, IR, X]

SUBSUMERS FROM OTYPES: [Galaxy, GalaxyInCluster, GalaxyInGroup, AGN, LINER, HighPowerRadioQuietAGN, RadioSource, InfraredSource, XRaySource]

Suggested Additional otypes: rG mul

--- M_31 is **CONSISTENT**

CSO_619

OTYPES: [* QSO, blu]

SUBSUMERS FROM OTYPES: [StellarObject, HighPowerRadioQuietAGN, BlueObject]

inconsistent otypes: StellarObject [] and HighPowerRadioQuietAGN [QSO]*

----- HighPowerRadioQuietAGN CANNOT be component of StellarObject

----- StellarObject can be component of HighPowerRadioQuietAGN

--- CSO_619 is **INCONSISTENT**: inconsistency from otypes

IRAS_20266+3856

OTYPES: [PN, Mas, IR]

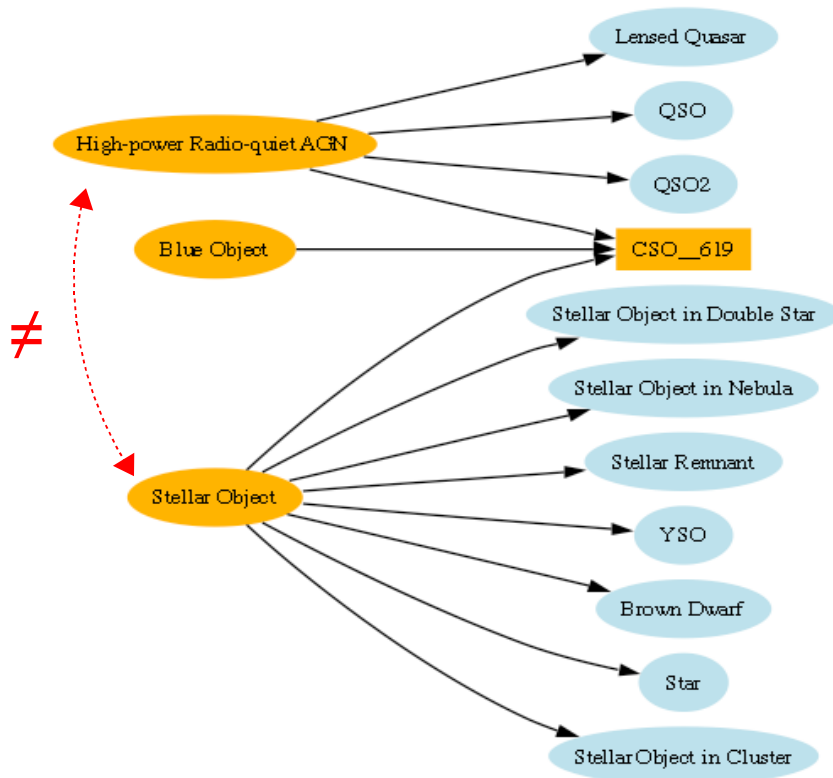
SUBSUMERS FROM OTYPES: [PlanetaryNebula, PlanetaryNebulaShell, Maser, InfraredSource]

Suggested Additional otypes: Rad mul Neb, ClD

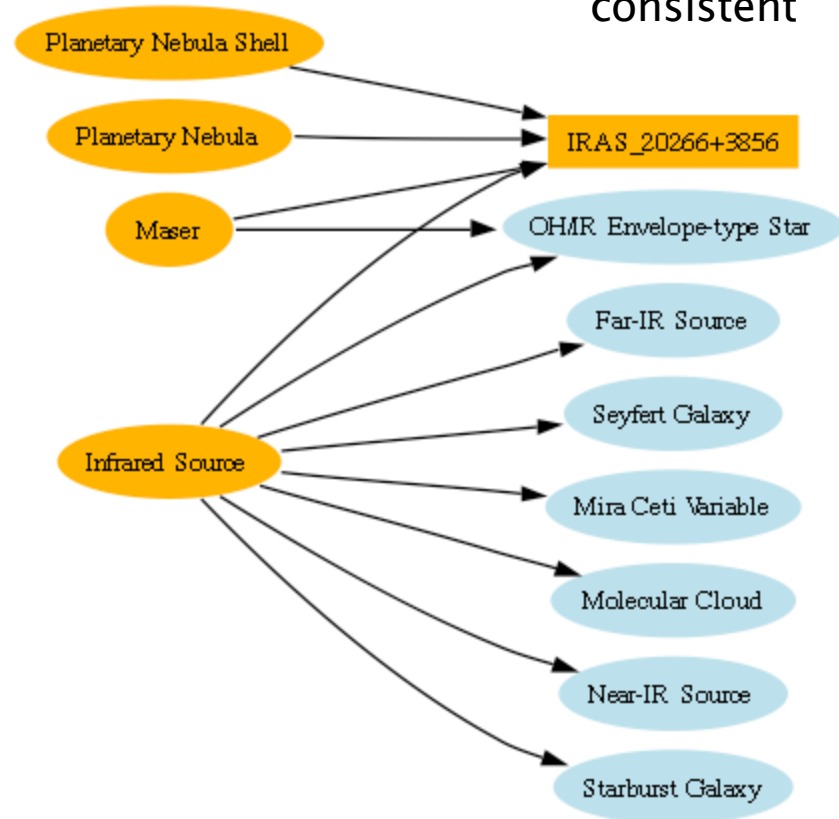
--- IRAS_20266+3856 is **CONSISTENT**

Consistency results (2/2)

CSO 619: inconsistent



IRAS 20266+3856:
consistent



SIMBAD consistency checker highlights

- Ontology independent from SIMBAD
- Local graph visualization of otypes within the ontology using Graphviz
- Check time: ~2.4s/object (25 objects/min)
- The SIMBAD data \leftrightarrow concepts link can work both ways
- With little tagging beforehand, ontology could be used with another database