Categorisations of object types in SIMBAD

Anaïs Oberto[1],*, *Cécile* Loup[1], *Mark* Allen[1], *Caroline* Bot[1], *Laurent* Cambrésy[1], *Sébastien* Derrière[1], *Françoise* Genova[1], *Ada* Nebot[1], *Pierre* Ocvirk[1], *Arnaud* Siebert[1], and *Bernd* Vollmer[1]

[1] *Université de Strasbourg, CNRS, Observatoire astronomique de Strasbourg, UMR 7550, F-67000 Strasbourg, France*

Abstract. Astronomical objects may be classified into types in many ways, and the evolution of such categorisations changes with new discoveries and progress in astrophysical understanding. The SIMBAD database contains information on astronomical objects that have been studied in the published literature, including a field that specifies astronomical object types. As a record that is derived entirely from the literature, a given astronomical object in SIMBAD may have multiple object types, and the list of object types must be maintained and updated. The SIMBAD object type list currently contains some 200 types, that are organised into a hierarchy based on astrophysical concepts. The hierarchical structure also includes relations between object types, and this facilitates searches of SIMBAD to obtain lists of all of the astronomical objects in a given category independently of the publisher or the year of publication. We will explain the organisation of astronomical object types in SIMBAD and how they may be used in queries of the SIMBAD database, and visualised on all-sky maps.

1 Introduction

The SIMBAD database [1] contains a lot of information on astronomical objects. Among these the object type is of crucial importance. It can be just a wavelength (e.g. infrared or X-ray source), or give the basic nature of the object (e.g. star, galaxy, or molecular cloud), or give a much more detailed description on the nature, geometry, or evolutionary status of the object (e.g. Seyfert 1 or 2 galaxy, or Wolf-Rayet star). In practice, SIMBAD lists a main object type and a list of secondary types for each object. It is fundamental that the CDS team can give the most detailed object type as a main type to the users, which means that we have to rely on a robust algorithm to decide when the main type should be changed and when contradictory classifications should be checked.

Historically, the list of object types in SIMBAD was mostly phenomenological, and organized as a linear hierarchy. This list is still the one given to the users on the web site when clicking on 'Object types." It is also used by the cross-identification program COSIM (section 5). Objects with the highest priority for the main object type are at the bottom, while objects with the lowest priorities are at the top. Object types at the same level in the historical hierarchy are assumed to be incompatible, which lower the cross-identification score of COSIM.

This historical list will be replaced in the coming months. It has been reorganized as much as possible on astrophysical concepts, especially the evolutionary stage for the stars, taking advantage of

---

*e-mail: anais.oberto@astro.unistra.fr ORCID: 0000-0001-7141-1542

the wide expertise of the astronomers attached to the CDS team. For instance, the variable stars are not categorized separately anymore as in most cases they correspond to an evolutionary status or/and initial mass (e.g. a RR Lyrae is on the Horizontal Branch, and an Orion variable is an YSO). This new list is still hierarchical, though the hierarchy has changed, but the compatibilities between object types are much more complex than before. This should be of great help to better determine the "best" type and optimize cross-identifications with COSIM.

The hierarchical structure is made of relations, and can be used in SIMBAD in different sorts of use cases:

- Simple description of the astronomical object's classification

- Search in all types to obtain a list of objects that belongs to a category

- See the whole hierarchy and links between object types

- Visualise the localisation in the sky of a specific (or generic) type, and cross those information

- Compare two object types according to the levels and branches in the hierarchy tree

2 Web usage

Common usage of object types in SIMBAD is a text shown in the web page. For instance, the "Messier 1" object, also called "Crab Nebula", has the specific type "SuperNova Remnant" that is the expanding shell left by a stellar supernova explosion.

SIMBAD also gives more object types, that could be used to complete the main one. It helps to understand that this object was part of a group of observations within a general pipeline type (generally linked to catalogues).

3 SQL usage

The relations between object types can be structured in tables of the SIMBAD database. In order to do that, SQL is generated to create all the hierarchy. This can be used by the astronomers to query directly SIMBAD and filter objects on a category. For instance, the figure 1 shows the schema inside the database that could be used by anyone who wants to query SIMBAD.

Thanks to such a schema, other databases or services will be allowed to query SIMBAD programmatically and exploit at best the hierarchy of object types for solving complex queries.

4 Graphical visualisation

Documentation can be automatically generated from the stored information of this hierarchical organization, and can be displayed as a nice graph (thanks to the graphviz [1] open source tool) as you can see on figure 2.

In this kind of graph, the links between objects are more obvious, and on the left part you can see that we show different levels of granularity.

The generation is done automatically with a program reading information of links in the database and generating a file in "dot" format, that is used to generate this graph. So that, any modification in the SIMBAD hierarchy will automatically be reported in the graphical documentation.
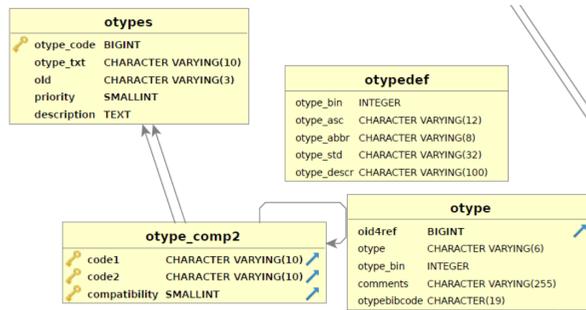
Figure 1. A table (here "otype_comp2") will manage the links of compatibility between 2 object types. Then, for instance, anyone would be able to search easily young stars using the hierarchy and thanks to a new specific function like "otype_childs" with this type of query : `SELECT * FROM basic WHERE otype IN otype_childs('YSO')`
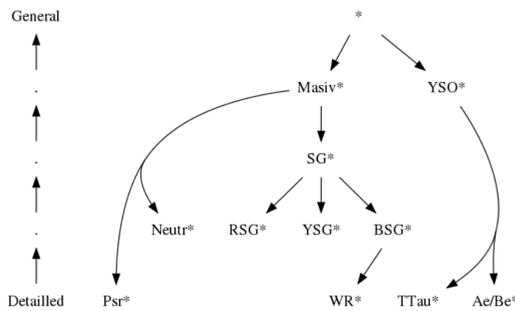

Figure 2. Graph automatically generated (partial hierarchy : a sub- tree under "Stars")

5 Cross-matching object types

Each SIMBAD object is a composition of all observations on many telescopes and satellites. That is why we could have secondary object types, depending on the points of view. In any case, before the insertion of astronomical objects in SIMBAD, we should compare these object types to check if there is no discrepancy according to the types hierarchy. The "COSIM" software (Comparisons of Objects for SIMbad [2] : internal cross-identification tool) is used to compare astronomical objects coming from a table, and compute a score of compatibility on each data, and specially on object type:

```
... Search around ICRS 20:16:31.91+37:39:09.1
... 2 object(s) found from coo
    1/1: PoC/*iC ( -10)  60.0"D ( 2.9)
+++ 2/2: **/*iC (   1)   8.8"B ( 0.0)
```

COSIM is going up and down in this hierarchy to quantify the relation between object types. We also have links between object types to know if they are compatible or not, outside the hierarchy. Both of these relations are used to compute a score for cross-identification.

In the previous example, the compatibility of a star in a cluster is much higher with a double star than the compatibility with a part of a cluster.

6 Full sky usage: SIMBAD object types in a map (MOC[3])

Once the information is well organised in the database, astronomers can use it to query directly SIMBAD and visualise the coverage of objects on one or more categories.
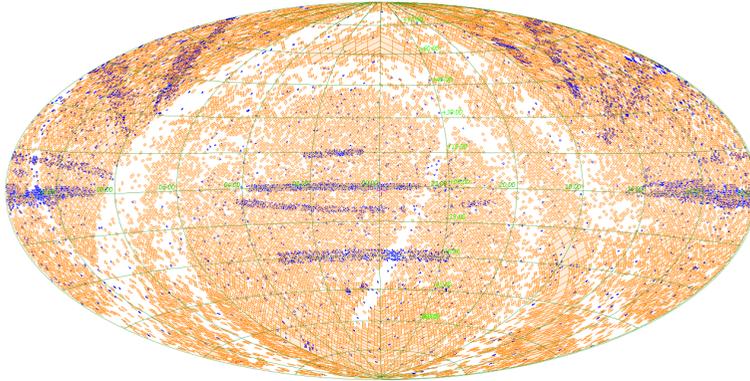

Figure 3. Visualisation of localisation of Infra-Red (orange) and Seyfert (blue) objects in SIMBAD

7 Conclusion

This work is still in progress, and should not be taken "as is." It opens a new perspective to improve in the long term the classification of objects and work has begun to share information with the astronomical ontology [4].

References

[1] M. Wenger, F. Ochsenbein, D. Egret, P. Dubois, F. Bonnarel, S. Borde, F. Genova, G. Jasniewicz, S. Laloë, S. Lesteven et al., Astronomy and Astrophysics Supplement **143**, 9 (2000), `astro-ph/0002110`

[2] C. Brunet, A. Oberto, M. Brouty, M. Buga, C. Loup, E. Perret, B. Vollmer, F. Woelfel, *COSIM: the necessary evolution of a cross-identification tool along with data evolution*, in *These Proceedings* (2018)

[3] P. Fernique, T. Boch, T. Donaldson, D. Durand, W. O'Mullane, M. Reinecke, M. Taylor, *MOC - HEALPix Multi-Order Coverage map Version 1.0*, IVOA Recommendation 02 June 2014 (2014), `1505.02937`

[4] S. Derriere, A. Preite-Martinez, A. Richard, L. Cambrésy, P. Padovani, *Ontology of Astronomical Object Types Version 1.3*, IVOA Note 03 March 2010 (2010)