OVERVIEW OF VIRTUAL OBSERVATORY TOOLS

Mark G. Allen
Observatoire astronomique de Strasbourg (UMR 7550)

ABSTRACT

I provide a brief introduction and tour of selected Virtual Observatory tools to highlight some of the core functions provided by the VO, and the way that astronomers may use the tools and services for doing science. VO tools provide advanced functions for searching and using images, catalogues and spectra that have been made available in the VO. The tools may work together by providing efficient and innovative browsing and analysis of data, and I also describe how many VO services may be accessed by a scripting or command line environment. Early science usage of the VO provides important feedback on the development of the system, and I show how VO portals try to address early user comments about the navigation and use of the VO.

Key words: Virtual Observatory; Software.

1. INTRODUCTION

Multi-wavelength science requires that

- data from different telescopes
- the analysis tools used to handle the data
- the on-line astronomy services
- archived data and information

to be readily compatible. The Virtual Observatory is the framework for such an interoperable system. The vision is to have all astronomy resources available to a user with the same ease of use as if they were all installed on your desktop computer.

The tools available in the VO now includes a wide variety of different types of software applications and services with different levels of use of VO standards. Many of the current tools are VO-enabled versions of existing tools, where additional VO capabilities have been added such as being able to search VO registries for data, or the ability to pass information interoperably to other applications. Some VO tools have been designed and built with VO in mind from the beginning, and have VO data searching, or registry browsing built into the core capabilities of the tool. As the VO is still at a relatively early stage in terms of scientific functionality, there are many prototype or ‘proof of concept’ VO tools that are implementing new and innovative capabilities, exploring the different ways in which the VO infrastructure may be used.

The currently available VO services also include many existing services that have implemented VO protocols on top of existing services as a kind of translation layer. In this way the existing services, many of which were already providing VO-like services, now expose those capabilities to the VO without the need to create, or recreate their service from scratch. As with VO tools, there are also many services that have been developed specifically for VO, with the various components such as registry access, VO formats, and meta-data incorporated into the system design from the outset.

The existing VO tools and services may also be combined in various ways to offer a wider range of capabilities than can be offered by a single tool or service. There are messaging mechanisms that may be implemented by the tool developers to make one tool talk to another, and there are also ways of scripting VO services at different levels using for example the python scripting language. Scripting VO services and tools together with languages already in wide use by astronomers puts the ability to custom build tools into the hands of the astronomer. In this way, many of the VO capabilities, exposed by graphical user interfaces (GUI), may also be used in customisable scripting mode to handle more complex processes than can be made available in GUIs. Also, scripting allows combining VO capabilities with conventional astronomy software.

The levels of robustness and maturity of the currently available tools varies, and as with any new instrument scientific users must be alert to the possibility of problems. Prototypes, while innovative and offering unique capabilities, are sometimes hard to use or unreliable. The VO standards are relatively new, and some of the core
Figure 1. The diversity of astronomy data. The VO aims to provide a system in which many different types of Astronomy data may be used. It is this diversity of data that distinguishes Astronomy in the use of e-Infrastructures, as many other fields have more uniform data sets. This collage of data includes all sky maps of the microwave background and the infrared sky from WMAP and DIRBE satellites, as well as multi-wavelength images of individual sources (M81, as compiled by AstroVirtel). Also the use of theoretical simulations is well advanced in the VO with access to the Millenium simulation data shown. Time series data are envisaged in the VO such as the MACHO microlensing event shown in the figure.
standards have not converged to a stable version which causes compatibility issues with tools that have implemented different versions. Other issues also arise from the availability and level of description provided for the content of the VO, that is the data provided by data centres and archives. Often the data desired has not yet been published to the VO, and sometimes the level of description is not sufficient to select the required data. The national and international projects are well aware of these issues, and are working toward stabilising the standards and accelerating the up-take of VO by major data providers. Moreover, the VO is ultimately a science driven endeavour, and early science usage of the systems is essential for them to evolve into the robust systems we require for astronomy research.

2. TOUR OF SOME VO TOOLS

Here we present a brief tour of some of the current VO tools. This is in no way a complete list, rather I aim to simply highlight various aspects and capabilities of VO tools. I consider the broad areas of visual browsing of the available data, the interoperability of tools, the different ways current VO services use the VO, and scripting capabilities.

2.1. Visual Browsing of Data

The three major types of astronomy data relevant to the VO are images, catalogues and spectra. Various tools allow searching, manipulation and analysis of these common data types, with a significant level of overlap in the capabilities which is to be expected for such commonly used data types.

The Topcat tool (http://www.star.bristol.ac.uk/~mbt/topcat/) is designed for analysis and manipulation of tabular data, in particular astronomical catalogues. It has an interactive graphical interface, and aims to do the things that astronomers want to do with tables. This includes viewing and editing the data and the metadata that describes the table (Fig 2). It includes powerful cross-matching features, statistical functions and several visualisation windows. In terms of the VO it makes a major contribution to the interoperability of tabular data from different sources because of the large number of formats it supports including FITS, VOTable, SQL, CSV etc. Topcat also has direct methods for searching VO registries (via the load dialogue), and most of its functions can also be accessed by using the underlying STIL (http://www.star.bristol.ac.uk/~mbt/stil/) library directly.

There are many astronomy software tools for displaying and analysing images (SAOImage DS9, gaia, IDL image programs, and more...) and most of these have implemented ways to access external images and catalogues. Here I highlight the VO features of the Aladin sky atlas (http://aladin.u-strasbg.fr/aladin.gml, Bonnarel et al. (2000)) which is designed to allow easy access to various data providing services. Alongside the direct access to services like VizieR (http://vizier.u-strasbg.fr) and SIMBAD (http://simbad.u-strasbg.fr), Aladin can send coordinate based queries to VO registries to find all the VO image, catalogue and spectra services that provide data covering a specified region of the sky. Figure 3 shows the result of such a query.

Aladin uses the metadata supplied in registry entries to enhance the selection of data, for example to use the image metadata to show the field of view outlines of the available images overlaid on images of the relevant region of the sky. The use of metadata for this purpose is much faster than other methods that require downloading the image. This is particularly useful when the result of a query contains many images, and combined with the image overlay, colour composite and controlled transparency of images in Aladin makes for a powerful way of determining the detailed coverage and overlap of a set of images and catalogues (see Fig 3). Aladin can also display the result of a query for spectra, showing the sky locations of the spectra on the image, and provides an easy connection to other VO tools for spectroscopy to load the spectra.

VOSpec (http://esavo.esa.int) is one such tool specialised for combining and analysing spectra made available in the VO. VOSpec may directly load spectra from Aladin, but also includes its own detailed query interface to find spectra of a given object or sky coordinate position. This is particularly powerful for building up a multi-archive spectrum of an object as shown in Fig 5 which shows data from the extreme ultraviolet through to the mid-infrared wavelengths. VOSpec enables a high level of interoperability between different units including...
Figure 2. Table data and metadata in topcat. Topcat is an excellent tool for handling all the different types of tables used in Astronomy. Large tables, tables with many columns, different coordinate systems, complex selections, math functions, and a wide range of convertible formats are all handled well in this well documented tool.
flux densities and has also made some important steps toward handling photometric points alongside of continuous spectra in order to build spectral energy distributions. Much of the flexibility of combining multi-archive spectra comes from the use of unit specification of the dimensions of the unit and the scaling factor to the SI system (Osuna & Salgado, 2005) as shown in Fig 6. Theoretical spectra published to the VO may also be manipulated in VOSpec which has explored the possibilities of how one may search for and apply fitting techniques for matching models to observations. Other capabilities of VOSpec are to be found in Baines, D. et al. (2009).

In addition to general purpose spectral tools like VOSpec, other services provide spectral and model fitting services to the VO. VOSA (VO Sed Analyser) is a web tool (http://svo.laeff.inta.es/theory/vosa) designed to automatically determine physical parameters of stars from comparisons with collections of theoretical models, that may work with several objects at the same time. The main benefit here is that the service does the technical homogenising of the models so that they may be used together. This kind of approach is paving the way for a more general description for photometric points in the VO.
3. VO TOOLS WORKING TOGETHER

There are various ways in which VO software applications may work together. Via VO standards different tools may access data and services published to the VO, but there is also a way in which VO tools may ‘talk’ to each other interactively. For example once a table is loaded into Aladin, it may ‘broadcast’ that table to Topcat avoiding the need to save and load the table into Topcat. Beyond simple loading of data into other applications, the ‘VO messaging’ standard behind these interactions, also allows for interactive highlighting of selected sources across multiple applications. For example sources that are selected in a colour-colour diagram in Topcat can be shown overlaid on various images in Aladin. The previously mentioned interaction of Aladin catalogues and spectra in VOSpec also makes use of this, and many applications are starting to use messaging for new and innovative connections between VO tools, and also to connect to conventional astronomy software like IDL and IRAF.

4. VO VIA THE COMMAND LINE

VO services may be accessed in various programatic ways. The tools that present these services in GUI windows access the VO essentially via the basic standards and protocols. Astronomers may also choose to access registries or VO services in a programatic way, in order to include calls to the VO from their own programs. To assist this mode of using the VO there are various libraries available for calling VO ‘tasks’ from a python scripting environment, and for accessing the VO from the UNIX command line. Python scripting resources may be found at the AstroGrid web pages (http://www.astrogrid.org). The UNIX VO-CLI command line tools for the VO can be found at http://iraf-nvo.noao.edu/vocli, with tasks to query and access VO image, and table data, resolving object names to coordinates and searching for data collections and services. The VO-CLI services also offer a simple form based interface that allows one to interactively construct queries, and see the command line syntax for making the same query programatically (see Fig 7).

5. THE FRONT DOOR TO THE VO

New users to the VO may find the array of tools and interfaces complicated and confusing at first. The VO is perceived as the new technology for astronomy software and as such it raises a high level of expectation for the capabilities and user friendliness of the system. There is indeed a diversity of different kinds of interfaces, from simple sky coordinate search input forms, to more detailed query interfaces with constraints on things like VO service types. Some VO teams have made significant efforts to construct intuitive interfaces, and others as described earlier are sometimes at the prototype stage of their development. While confusing at first, it is well worth the effort to discover these new tools for doing science as they offer capabilities that are not available elsewhere - searching global archives for data on a given source for example. Early user feedback (at meetings such as this one) has however indicated a strong desire for a single entry point into the VO, so that new users can more easily find the tools and services relevant to their scientific needs. These comments are sometimes misinterpreted as a call for the VO to be designed as a monolithic software system, but I suspect the main issue is simply the need for some order in the chaos, so that a scientist can at least find a starting point for using the VO. Some groups have made efforts to provide a front door to the VO for their users, for example the US NVO front web pages (http://www.us-vo.org/) seeks to put ‘the Universe at your fingertips’ with a clean set of icons with links to basic VO services with concise descriptions. The AstroGrid VODesktop tools (Tedds, J. et al. (2009)) provide an entry point to the VO with a search and query interface that works interoperably with other VO tools, with the complexity (and power) more up front than in the front web page approach.

Other information on the available VO tools and services are available on VO project web pages such as the software and tutorials pages of the EuroVO (http://www.euro-vo.org). The IVOA has only limited information addressing scientific users of the VO, but does provide an open list of VO tools available (and editable) on the IVOA Applications Working Group pages (http://www.ivoa.net/cgi-bin/twiki/bin/view/IVOA/IvoaApplications). News items about new capabilities of VO tools are available via the IVOA newsletter which is given a more prominent web presence at http://www.ivoa.net/newsletter.

6. CHALLENGES AND OUTLOOK

In addition to creating appropriate entry points to the VO, there remain challenges for the development and uptake of VO tools. It really only makes sense to use
Figure 7. NVO command line access to VO

VO tools if they provide capabilities not available elsewhere, so it is essential that VO tools concentrate on the data access and interoperability capabilities that are enabled by the VO infrastructure. These tools must also mature beyond prototypes and become robust enough for scientists to rely on them for research tasks. Current VO tools show that the VO has been successful in implementing efficient data access mechanisms, searching for data and loading it into appropriate tools is relatively easy via the various tools. The scientific description of the data is however not sufficient for many of the use cases put forward for the VO, in particular the description of photometry and the associated filter systems that would be required for building SEDs. Catalogues will list what the filter name is, but linking this with actual filter transmission curves or the calibration method is difficult. Work on such data description models and metadata is in progress, but a more significant effort is required in this area. Initial up-take of VO publishing of scientific data products by data centres is encouraging, and the expected high levels of VO compliance of future projects, such as indicated in the European AstroNet infrastructure roadmap (http://www.astronet-eu.org) indicates that the astronomy community sees the VO as the right approach for distributing data and services. There are however key datasets and archives that are not currently available in the VO, so there is also more effort required to engage with data providers and to make it easy for them to publish to the VO so that users see a more complete set of data when searching with VO tools.

Overall, VO tools allow a new way of doing astronomy by combining data that comes from different telescopes and using on-line services to access information and computational resources. It is well worth the effort to try VO tools because they offer capabilities that are not available easily or efficiently via conventional astronomy software. The VO is presently in transition to an operational phase with tools and services at various levels of maturity. The growth and development of the VO relies critically on feedback from early science use and communication between the VO projects and the astronomy community at meetings such as this is essential for building the VO as an effective tool for doing science.

REFERENCES

Baines, D. et al. in this volume.
Tedds, J. et al. in this volume.
This page is left intentionally blank.