Demonstration Project for the NVO: A Search for Brown Dwarfs

Bruce Berriman, John Good, Alex Szalay and Davy Kirkpatrick

Version 1.2

Version	Date	Modifications
1.2	8/5/2002	Added the scientific goals; add D. Kirkpatrick as author
1.1	6/26/2002	Response to technical questions from NVO participants – clean up of description
1.0	5/10/2002	Initial version

Scientific Goals:

Both the Two Micron All Sky Survey (2MASS) and the Sloan Digital Sky Survey (SDSS) have revolutionized our knowledge of substellar objects. These large-area surveys have revealed a continuum of sources from the lowest mass stars down to methane-rich brown dwarfs and have enabled the establishment of two new spectroscopic classes - the L and T dwarfs.

The low temperatures of these objects along with broad, deep absorptions by alkali metals make L and T dwarfs increasingly red in the i-z color of SDSS. In 2MASS, the onset of strong methane absorption particularly at H and Ks bands causes the J-Ks colors of T dwarfs to become very blue (J-Ks~0.0) after reaching a peak of J-Ks~2.1 for the latest L dwarfs. Both surveys are very powerful tools for finding nearby brown dwarfs. However, cooler and/or more distant brown dwarfs can be hard to pick out in either survey since such objects become single-band detections in both.

The coolest known brown dwarfs are Gl 570D (5.9 pc distant) and 2MASS 0415-0935 (~6 pc distant) with temperatures of 700-800 K. These objects have very deep methane bands at 1.65 and 2.2 microns and, were they more distant, they would be undetectable at H and Ks in the 2MASS survey. Cooler objects are also expected to be undetectable at H and K due to even stronger methane absorption. In the SDSS data, the cooler and more distant brown dwarfs become z-only detections. Burgasser (2001, PhD thesis, section 2.3.1) followed up a partial list of 41 J-only candidates over ~3500 sq. degrees of 2MASS data but found all sources to be spurious. Similarly Fan et al (2001, AJ, 122, 2833) found that the vast majority of the SDSS z-only detections are artifacts, and the bona fide brown dwarfs could only be culled from the chaff using follow-up data from other surveys or telescopes.

In this NVO pilot study, we will cross-correlate the 2MASS and SDSS data to search for z-only SDSS detections that have a matching J-only detection in 2MASS, thus providing a list of robust single-band detections from both surveys that can be investigated further with follow-up photometry and spectroscopy. The selection in principle is quite straightforward: T dwarfs have uniquely red z-J colors of 3.0 and higher, so we need only look for high-SNR single-band sources from both surveys which match this color constraint. Matching can be done within a small search radius (~2") since the surveys have almost identical epochs, eliminating to first order an allowance for proper motion. Finally, the study should return the 2MASS and Sloans image that contain the candidate matches. Visual examination of the images will aid in eliminating spurious matches due to noise spikes that escape the filters, binary pairs and Solar System objects.

Goals:

Cross-Matching and Publishing, two major long–term goals of the NVO:

Generalized cross-matching of sources from catalogs served from multiple sites and covering different wavelength regimes will be an important capability of the NVO. Ultimately, the NVO must deploy scaleable tool(s) that will cross-match small (MB-sized) and very large catalogs (TB-scale and eventually PB-scale). Currently, there are mature cross-matching tools that will cross-match GB-sized catalogs, but engines for cross-matching much larger scales are in the research phase. This demonstration project will begin the necessary evolution towards deploying NVO cross-matching tools. In this project, it is important that there is a clear separation between data provider and application, a separation mediated by evolving NVO middleware for publishing, service description, and directory services.

A search for brown dwarfs: We will deploy web-based cross-matching applications demonstrate their science value by identifying T-type brown dwarfs through cross-matching the 2MASS (infrared) and several optical catalogs that have been published in an NVO-compliant manner. A major target will be the Sloan (optical) source catalogs to identify objects with extreme colors. We will also investigate the use of the newly-released USNOB catalog in determining the completeness of such searches for brown dwarfs having less extreme colors: while the USNOB catalog is not as deep as the Sloan survey and will likely not contain any of the extreme T-dwarfs seen there, it does contain proper motion information and has much greater sky coverage. We will also consider a few other selection criteria, like high redshift QSO's, or high-redshift galaxies (dropouts) as part of our demo suite.

Technical Approach:

This project will take two parallel but related approaches. The first involves modest evolution of existing tools to deploy an operational service that will emphasize ease-of-use and accuracy over speed (but will run on the order of minutes not hours!). The second involves development of a fast cross-match engine that will eventually operate on very-large catalogs. These efforts will share a common set of requirements, and the results from existing code will be used to validate the results of the fast engine.

Requirements:

1. Support full sky searches
2. Interchange information using a standard document format (probably a specific VOTable).
3. Run from a web interface that allows users write their own custom queries, and select catalogs, input quite general search criteria, and specify table columns that will be served on output.
4. Interact with the component archives through web services, probably using SOAP as the envelope that contains NVO-compliant data formats.

5. Display results in a format that is easy to comprehend.
6. Deliver results from the “fast” engine within one minute for real-time demonstration purposes (even if over a small area of the sky), to make sure that attention is kept
7. The architecture will use simple and portable modules.

Development Steps

Search Criteria

Develop search criteria – color, signal-to-noise and search radius - for identifying brown - dwarf matches between 2MASS and Sloan catalogs. Proper motion can most likely be ignored in the first instance because the two surveys are contemporaneous. Understand how USNOB can be brought to bear on this pilot study. Engage IPAC brown dwarf specialists on this issue. Discuss other possible searches within selected people in the broader community (high-z quasars, galaxies).

Cross match engines.

See Appendix A for detailed notes on the development steps and the cross-match engines.

1. **Design of algorithms:** Develop and document the algorithms and software design for the two engines. These should follow a common format, in so far as possible. Algorithms will address the sorting and indexing ability of remote catalog services.

2. **Use of existing tools:** IPAC is developing a cross-identification engine through funding provided NPACI Digital Sky. It will be delivered in August 2002; it will deploy an operational version of the “compare” program already in use at IPAC. It will derive an initial list of candidate matches within the search area, and then apply a probabilistic algorithm to identify matches (the Psearch algorithm in use at NED that makes use of uncertainties in position).
3. **Development of protocols**: Work with the NVO Metadata Group to standardize the representation of positional error in catalogs. This will entail selecting or creating part of the UCD namespace to indicate concepts such as “simple error” and “error covariance matrix”. And, work with the NVO Metadata Group to standardize the publication of full

sky search services.

4. Fast engine: Alex and colleagues have already developed a prototype server, the Sky Query server.

Portals and front-ends

Develop a common front end for the web-based requests to the cross-match services**.** Need to consider what is needed to allow users to run cross-matches. We will demonstrate ways to browse cross-match responses, using available NVO-compliant browsers for displaying and studying results, and for displaying meta-information (e.g. area covered . . .).

Schedule (2002 is implied in all dates!)

1. Define search criteria; investigate whether USNOB is valuable – July 15
2. Work with NVO Metadata Group to identify necessary standards work to enable publishing, discovery, and utilization of the component services.
3. Deliver documented algorithms and designs – August 15
4. Agree on and prototype front end – August 15
5. Deploy cross-match engines - September 10
6. Do cross-matches between 2MASS, Sloan, USNOB, and other published optical catalogs as needed – September 30
7. Validation of results – cross-compare the cross-comparisons, compare results with known brown dwarfs – December 1.
8. Presentation of results to the community – January 6.

Appendix: More detailed notes on development steps

1. Develop and describe algorithms that do the cross-identifications. Perform Monte-Carlo simulations of a few searches, to estimate response time. Compare algorithms that stream all data to one place and compare them to recursive algorithms.
2. Consider interchange protocols between computational and data resources, resisting the temptation to go around the protocol for a quick fix, but rather making refinements through interaction with the NVO community.
3. Define what the front end should do. As the minimum, we need an interactive

browsing capability of the tables and columns in the databases. We need to be able to get the covered sky areas, positional accuracies, the names of the primary table, and the primary attributes (UCD). The covered sky area should be displayed; possibly with an outline of the intersecting area (maybe this is too hard for now). We should show an image of the area, and possible mark the objects finally found. We need to deliver a VOTable document that contains the output data, and their likelihoods.

4. The visualization of the above should be done using tools that use NVO-compliant protocols
5. We will need to decide on the query language to express the complex match and search criteria, if the remote data service supports such criteria. The extensions have to have a part that describes the fuzzy spatial join, and a part that describes the common area. We also need to work with the NVO Metadata Group on the emerging protocol to describe the remotely published archives. The full grammar may include SQL if the remote service has that capability. The extensions developed by SkyQuery represent a possible start. It is able to handle non-detections as well as detections.
6. Design the low-level web services that serve the cross-match portal. The minimal set should include a description of the resource (area, name of primary table, wavelengths, etc); a method to get the names and descriptions of tables, the method to get the names and descriptions of attributes in a given table, the ability to return results from a simple SQL query, and a service that can perform recursive cross-identifications.
7. Decide, based on the outcome of (1), on how much of the cross matching is done inside the databases, and how much is done at the portal. Design the appropriate stored procedures, in particular their interfaces.