Astronomical Data Analysis Software and Systems XIII
ASP Conference Series, Vol. 314, 2004
F. Ochsenbein, M. Allen, and D. Egret, eds.

VOTable JAVA Streaming Writer and Applications.

Pallavi Kulkarni, Ajit Kembhavi

Inter University Center For Astronomy And Astrophysics (IUCAA), Pune.

Sonali Kale

Persistent Systems Private Limited (PSPL), Pune.

Abstract. Virtual Observatory related tools use a new standard for data transfer called the VOTable format. This is a variant of the xml format that enables easy transfer of data over the web. We describe a streaming interface that can bridge the VOTable format, through a user friendly graphical interface, with the FITS and ASCII formats, which are commonly used by astronomers. A streaming interface is important for efficient use of memory because of the large size of catalogues. The tools are developed in JAVA to provide a platform independent interface. We have also developed a stand-alone version that can be used to convert data stored in ASCII or FITS format on a local machine. The Streaming writer is successfully being used in VOPlot (See Kale et al 2004 for a description of VOPlot).We present the test results of converting huge FITS and ASCII data into the VOTable format on machines that have only limited memory.

1. VOTable JAVA Streaming Writer

1.1. Introduction

VOTable java Streaming Writer is an API library for writing data in VOTable format. Streaming support is provided to deal with very large VOTable files without running out of memory. The writer does not create a tree structure in memory, so that the memory requirement is substantially reduced and very large VOTables can be written.

2. Streaming Writer Applications

2.1. ASCII and FITS to VOTable Converter

This is a tool for converting ASCII or FITS files to VOTable format. For ASCII files, it supports both ASCII files with column delimiters and ASCII files with fixed width columns. For FITS files, it supports FITS ASCII and binary tables. A number of useful data processing, browsing and visualization tools are available to deal with data in the VOTable format. But, most of the astronomy

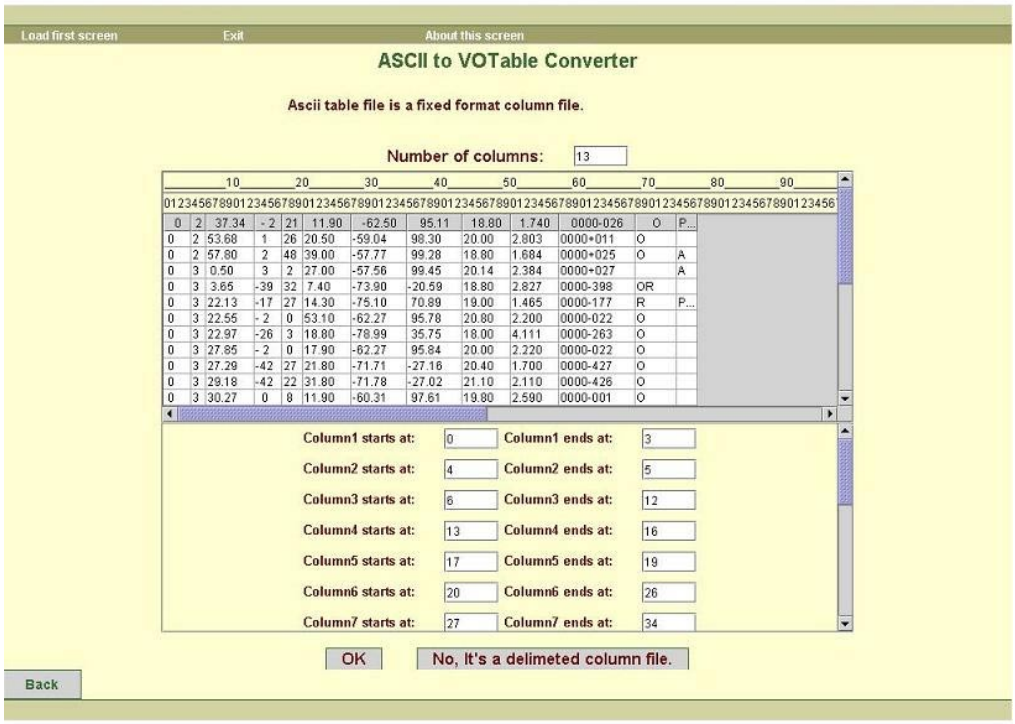

Figure 1. Screenshot 1

legacy data exists in the form of ASCII catalogues and FITS tables. To be able to use the new VOTable packages on legacy data, it is necessary to have a converter which transforms that data to the VOTable format. Tools are also needed to transform the outputs generated by various application packages to the VOTable format. We describe such a tool here.

Conversion of ASCII files The tool can handle two types of ASCII files:

1. ASCII files with column delimiters.
2. ASCII files with fixed width columns.

The tool first determines which of the two types the ASCII file is. It then determines the delimiter, or the width of individual column depending upon the file type. It reports to the user the number of columns detected, and this number can be changed if required in special cases. Finally, it determines the data type of individual columns by scanning the first few lines in the file and displays the file metadata which can be edited. When the file is converted to VOTable format, it is written to output stream using the VOTable Java Streaming Writer. The tool works in streaming mode by reading the ASCII file one line at a time and writing the VOTable one row at a time. It can therefore be used to convert big ASCII files to VOTable files without running out of memory.

Conversion of FITS files The tool makes use of nom.tam.fits library to read FITS files. Metadata is read from the headers of individual Header Data Units.

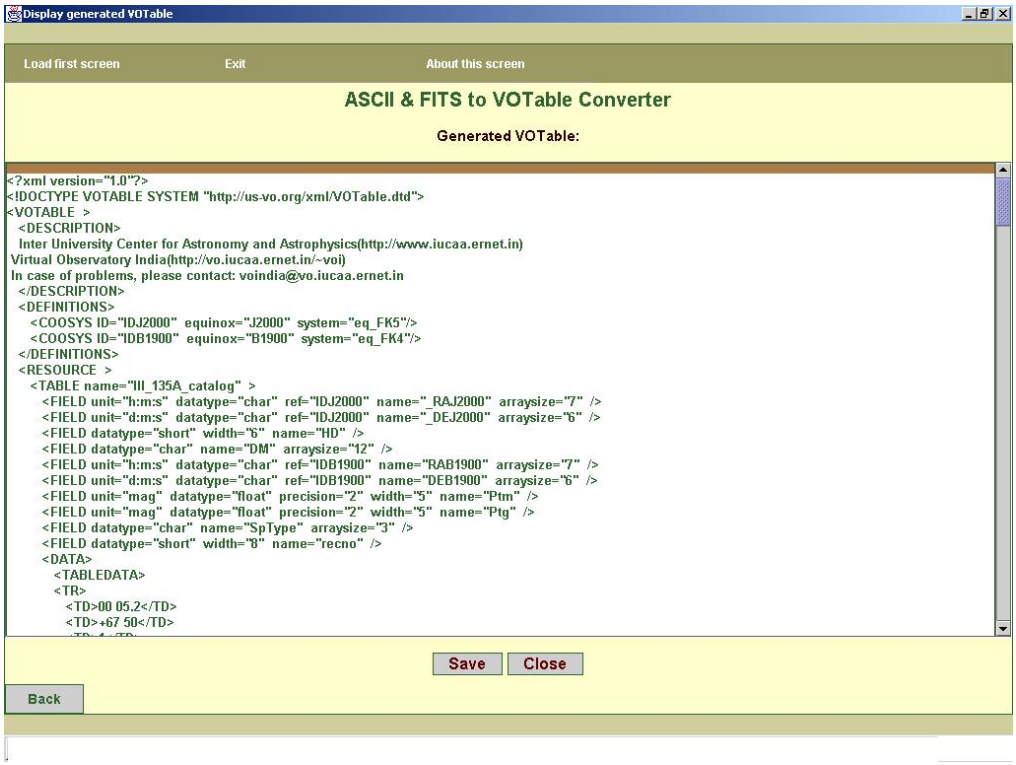

Figure 2. Screenshot 2

Metadata is displayed to the user for confirmation and it can be edited as well. After the FITS tables are converted to VOTable, the tool makes use of the VOTable Java Streaming Writer to write the data to the output stream.

2.2. VOPlot

VOPlot is a tool for visualizing astronomical data. It is an application for plotting different astronomical graphs using data stored in VOTable format. VOPlot makes use of VOTable Java Streaming Writer to export the data in VOTable format. See Kale et al (2004) for a description of VOPlot.

3. Some Test Results

3.1. For ASCII files:

1. The tool was found to require about 8.03 minutes to convert the Tycho-II catalogue with 1058332 rows having size 354MB on a P-IV machine with 384MB RAM.
2. The tool was found to require about 0.7 minutes to convert the Burbidge catalogue with 22484 rows having size 2.29MB on a P-IV machine with 384MB RAM.

3.2. For FITS files:

1. The tool was found to require about 0.016 minutes to convert the UGC catalogue with 12939 rows having size 1.22MB on a P-IV machine with 384MB RAM.
2. The tool was found to require about 0.1276 minutes to convert the hd catalogue with 100000 rows having size 7.16 MB on a P-IV machine with 384MB RAM.

4. Relevant Websites

- Inter University Center for Astronomy and Astrophysics - `http://www.iucaa.ernet.in`
- Persistent Systems Private Limited - `http://www.persistent.co.in`
- Centre de Donnés astronomiques de Strasbourg - `http://cdsweb.u-strasbg.fr`
- VOTable - `http://vizier.u-strasbg.fr/doc/VOTable/`

References

Kale, S. 2004, this volume, 350