

# Arecibo Software Development for Spectral Line Data Reduction

Steve Torchinsky

2004 October 5

## Introduction

This memo gives a very brief overview of software at Arecibo, with emphasis on future development of spectral-line data reduction software, in preparation for the meeting to be held on Monday, October 11th, 2004 at 9am in the library.

Software requirements at the Observatory can be divided into five main categories.

1. telescope control
2. datataking
3. user interface to datataking and telescope control
4. data monitoring
5. data processing

In fact, there are more categories such as web development, scheduling, and proposal management. These will not be discussed here.

## 1 Telescope Control

Telescope Control is handled by the Vertex control-loop system. The direct interface to the Vertex system is a set of higher level routines which run on a VME crate called PNT. The PNT computer can issue low-level requests, such as “go to AZ-ZA and move at a constant rate”. The PNT computer uses the low level commands to implement RA-DEC tracking, and scanning patterns. User access to PNT is via the datataking program that runs on COR (see below). CIMA (see

below) implements scanning patterns by sending commands to the datataking program running on COR, which in turn communicates with PNT, which in turn communicates with the Vertex system through a serial port.

Phil Perillat is responsible for PNT and COR.

## 2 Datataking

Datataking for spectral-line observations is done either by the Interim Correlator (COR), or the WAPP. This memo addresses software effort for the WAPP and future spectral line machines.

The WAPP is a cluster of four dual-processor Pentium based systems running the GNU/Linux operating system. The WAPP writes spectral line data in FITS format. This is a standard file format which is recognized by a large number of applications developed worldwide. A utility developed by NASA called fitsview can read and display data in any FITS file, including those written by the WAPP.

Phil Perillat is responsible for COR.

Jeff Hagen is responsible for WAPP.

## 3 User Interface to Datataking and Telescope Control

Observers run the telescope and WAPP datataking using the Control Interface Module for Arecibo (CIMA). CIMA is a graphical user interface based on the Tcl/Tk widget set.

Datataking with the Interim Correlator can also be controlled from CIMA. Observers may also prefer to run a Tcl script on COR to run their experiment, instead of using CIMA.

Jeff Hagen is responsible for CIMA. Future versions of CIMA will be developed by a collaboration of Jeff Hagen and Mikael Lerner.

## 4 Data Monitoring

Observers can run an online data display on the Dataview computer. The ALFA data display shows raw bandpass plots, waterfall plots, and updated text of the FITS header. Data from the WAPP is multicast on the WAPP subnet using Jeff Hagen's socket communication library. This method

allows data display without increasing disk i/o, and without relying on sometimes sporadic Network File Share (NFS). Future developments for the online data display include data accumulation for display of integrated data and on/off subtraction. The ALFA data display will be adapted for use with single-pixel observations.

Mikael Lerner is responsible for Dataview.

Data display for the Interim Correlator is possible by running the IDL routine CORMON ALL. This routine works by polling the saved data file, and plotting results after data has been written to disk.

Phil Perillat is responsible for CORMON.

## 5 Data Processing

### 5.1 Current Data Processing Software

There is a large library of routines for data processing at Arecibo written in the Interactive Data Language (IDL). Many of these were written by Phil Perillat in support of the Interim Correlator. There are also large contributions from past and present staff members, notably Tapasi Ghosh, Chris Salter, Snezana Stanimirovic, Karen O'Neil, Gomathi Thai. Visiting observers have also contributed, Carl Heiles being the main contributor.

The Arecibo IDL library can be used to process spectral line data from the WAPP using routines which read the FITS files and translate the data to the COR structure. Phil Perillat has developed the translation routines.

A common complaint regarding the Arecibo IDL routines is its complicated user interface. To address this issue, Erik Muller developed a text-based, menu driven user interface to some of the COR routines. This user interface is called ICRed (Interim Correlator Reduction) and is a prototype which demonstrates an easy-to-use user interface.

An independent development, also in IDL, is WapRed (WAPP Reduction) which is written by Erik Muller. WapRed uses the same menu philosophy as ICRed, but it reads the FITS files and processes the data directly without going through a translation stage. It can do basic data manipulations including on/off and co-adding. It can also split FITS files according to keywords which has been used to create FITS files which are compatible with other packages, such as IRAF.

## 5.2 Future Directions

Data reduction software at Arecibo must be developed within the context of the worldwide effort on spectral line data reduction software. There are many packages for reduction of Astronomical spectral line data. A short list of some of these is: CLASS, AIPS, IRAF, SPECX, XS, DRP. In addition, the astronomical community has made available a large number of routines written for generalized data processing packages such as IDL, MatLab, and PerlIDL to name a few.

The development of the WAPP as a spectral line machine also introduced the Arecibo philosophy on data reduction software. The motivation behind choosing a standard file format was to allow the users the choice of data reduction packages. In practice, this has worked imperfectly. FITS files can be read by nearly all the packages listed above, however their implementation is often incomplete. Also, packages such as CLASS, can not deal with telescope specific information, and require calibrated data as input.

The effort available at Arecibo for data reduction software development is limited. The full-time programmers have responsibilities for telescope control, datataking, user interface, and data display, all mentioned above. In addition to research, Astronomy staff members are expected to spend 50% of their time on observatory service tasks. Postdoctoral Research Associates spend 25% of their time on service tasks. Astronomy staff members can devote some of their service time to software development, but they also have other service responsibilities such as observer friending, receiver friending, observatory documentation, conference organizing, student supervising, spectrum management.

It is not sensible to use our limited resources to develop data reduction software which recreates the functionality of other packages. The hope that third-party software would be capable of processing WAPP data turned out to be only partly realized. For example, fitsview can display WAPP data, but it is not a data reduction package. However, with some programming effort, IRAF was successfully used by Ellen Howell for reduction of ALFA data.

There are two main areas where we can concentrate our effort. We can continue the development of the Arecibo IDL library. The second possibility is the development of an Arecibo software pipeline which produces calibrated spectra, coadded integrations, and maps.

A common complaint we have heard about the Arecibo IDL routines is the fact that they have a complicated and inconsistent user interface. Erik Muller addressed this issue by developing the ICRed wrapper for the Interim Correlator routines. It is also possible to develop a graphical user interface using the IDL widget toolkit.

We have had the request many times that a visiting observer to Arecibo should be able to simply walk away with calibrated spectra and maps, printed out and ready for publication. In other words, Arecibo should be a point-and-shoot telescope, like an auto-focus, auto-exposure digital camera. It is possible to approach this functionality in the following way. Calibration, co-adding, on/off

subtraction, can all be done by a pipeline program. There would be no options to the program. It processes data files according to what is contained in the header, and outputs the data product in a FITS file. The resulting FITS file can be displayed with fitsview, or it can be imported to another data reduction package for more sophisticated data processing such as baseline removal, line fitting, RFI excision.

### 5.3 Recommendations

I propose that the pipeline software as described above should be developed, and that Mikael Lerner should be responsible for leading the effort. This follows naturally from the work he is doing on the data display.

I recommend that the IDL library should continue to be developed by those people using IDL, but that it is not necessary to develop a user friendly interface. Arecibo users who run IDL are already familiar with IDL and the Arecibo routines. A friendly user interface will not have such a great impact on those people to warrant the enormous effort required to make the interface. New users will have the option to use their favorite package once the pipeline is in place, or they can simply take the data “as-is” and produce output using fitsview.

Ideally, the continued IDL development would have a coordinator. That person would work towards having a uniform interface, either by writing wrappers for contributed routines, or by requiring developers to pass data and arguments in a standard way. The coordinator would be responsible for quality assurance by ensuring that the routines produce the correct results.