The IRAF Spectroscopy Reduction Packages and Tasks

F. Valdes

IRAF Group, NOAO[1], PO Box 26732, Tucson, AZ 85726

Abstract. A variety of generic and instrument and data specific spectroscopy reduction packages are available in the NOAO suite of packages in IRAF. The instrument specific packages contain specialized routines written as IRAF command language scripts using complex and generic image processing and spectroscopy programs. The reduction routines connect all the reduction operations needed to fully reduce spectral data from particular instruments and customize their parameters. The package organization and the capabilities of the specialized reduction packages is presented.

1. Introduction

There are many types of spectroscopic instruments and data in astronomical use. The spectroscopic data, now usually in digital form, requires many reduction steps. In addition, there are many reduction recipes, scientific goals, and computer aptitudes among astronomers. This leads to the need for a diverse set of spectroscopic data reduction tools. This paper describes the approach taken in the NOAO[1] suite of IRAF[2] packages for handling this diversity.

The approach is to provide both a large set of generic tools which address a particular reduction operation in a very general way, a set of packages devoted to particular instruments or types of data, and specialized reduction tasks (programs) which integrate the reduction steps for particular instruments or types of data into a single command. This approach is applicable to other software systems although the software described in this paper is specific to IRAF.

Generic tools perform specific operations on spectra; for example determining dispersion functions from arc calibration spectra. High priority is given to generality in the type of spectra and algorithms. This is done by providing many parameters and, often times, multiple algorithms to allow tailoring the task to the requirements of the data. Emphasis on generality and modularity has a number of benefits. The most obvious is that a variety of spectral data is addressed once rather than writing new or modified versions for each new

[1]National Optical Astronomy Observatories, operated by the Association of Universities for Research in Astronomy, Inc. (AURA) under cooperative agreement with the National Science Foundation

[2]Image Reduction and Analysis Facility, distributed by the National Optical Astronomy Observatories

type of data encountered. Addressing individual reduction operations in separate programs allows new algorithms to be added without modifying other code and makes maintenance easier. General and flexible tools can also be exported and used at other institutions. I also believe that emphasis on generality and feedback from application to a variety of types of data also leads to better software. Finally, flexible, generic tools can be easily and quickly tied together in a command language scripting environment to produce specialized functions.

The most interesting, and recent, components of the IRAF spectroscopy software are the specialized data reduction packages and tasks. These are IRAF command language (CL) scripts selecting and combining the generic tools for the reduction of a particular type of spectroscopic data. The instrument specific packages collect only the generic tools needed for a particular instrument or type of data. The specialized reduction tasks provide a capability missing from the generic tool approach; namely, a complete and focused task for reducing spectroscopic data. The following sections of this paper describe the selection of spectroscopy packages and the capabilities of the specialized reduction tasks.

Specialized data reduction scripts offer several advantages over compiled tasks. The first is that scripts provide a very quick prototyping capability. They are also easier to maintain and modify for specific situations. To emphasize this last point, there have been instances where users were able to modify the scripts for their own purposes and, in one case, an instrument developer was able to build his own package and reduction task for his instrument. For compiled tasks users would need to learn SPP (the IRAF programming language) and how to compile, link, and install IRAF tasks. This leads to the point that scripts are architecture independent (although the commands in them will be architecture dependent), so there is no need to worry about multiple object libraries, binary directories, etc. Finally, command language scripts are easier to update as improvements are made to the generic tasks. For instance the recent addition of a nonlinear dispersion function capability required only the addition of another parameter and slight modification of the command invoking the modified generic task.

2. Spectroscopy Packages

IRAF commands are organized into packages. A package, when *loaded*, defines a set of commands, called *tasks*. Multiple packages may be loaded at the same time and commands from all packages are available. Typically, general image processing packages are always loaded while specific data reduction or analysis packages, such as the spectroscopic packages, are loaded as needed.

In this paper we discuss only the spectroscopy packages developed at NOAO and contained in the **noao** package. Other IRAF packages are are being developed at a number of institutions and these packages may also contain generally useful spectroscopy software.

The NOAO spectroscopy packages divide into two types, generic and instrument specific. The generic packages contain tasks for a variety of general spectral reduction operations. They typically have many parameters and algorithms to allow application to the widest possible types of data. Table 1 lists the current generic spectroscopy packages in the IRAF **noao** package.

Table 1. Generic Spectroscopy Packages

apextract	- Aperture spectra extraction package
longslit	- Long slit reduction package
onedspec	- One dimensional spectral reduction and analysis package
rv	- Radial velocity analysis package

The **apextract** package extracts one or more, one dimensional spectra from two dimensional images by summing data values within an aperture across image lines or columns. This assumes that the spatial and dispersion axes are close to the image axes; how close depends on the spatial extent of the spectral profile and the needs of the user. If significant distortions are present or very extended profiles are to be analyzed the **longslit** package provides tools for making geometric distortion corrections and wavelength and flux calibrations in two dimensions. The **onedspec** package contains all the tools for calibrating and analyzing one dimensional spectra. The radial velocity package provides tasks for measuring radial velocities from spectra.

The large number of tasks and task parameters in a generic package, plus the need to use tasks from several generic packages, can be confusing to novice users. The IRAF command environment allows creating new packages containing a set of the generic tasks from multiple packages, defining different default parameter values, and possibly adding additional tasks which are relevant to a particular instrument or type of data. The specialized spectroscopy reduction packages, listed in table 2, have been created in this way.

Table 2. Specialized Spectroscopy Packages

Fiber	argus	- CTIO ARGUS reduction package
	foe	- KPNO Fiber Optic Echelle reduction package
	hydra	- KPNO HYDRA (and NESSIE) reduction package
	fiber	- KPNO coude three fiber reduction package
	msred	- GENERIC multiple spectra reduction package
Slit	echelle	- GENERIC echelle reduction package
	goldcam	- KPNO GOLDCAM reduction package
	slit	- KPNO coude slit reduction package
	specred	- CTIO spectrophotometric reduction package
Scanner	iids	- KPNO IIDS reduction package
	irs	- KPNO IRS reduction package

The packages may be classified into three categories; optical fiber spectra, slit spectra of stellar objects, and one dimensional, beam-switched scanner spectra. The reduction operations required are similar in each category. Echelle spectra, both with fibers and with slits, are not categorized separately because,

apart from the form of the dispersion function, the reductions can be treated as multiple fiber or slit spectra. These are not the only possible classes of instruments and data types. Others will be added in the future.

Because the data for each category tend to be quite similar one should appreciate that data from similar, non-NOAO, instruments may also be reduced with the one of the reduction packages and the specialized reduction tasks after possibly minor parameter changes. Alternatively, one of the these packages may be taken as a starting point for developing IRAF-based, reduction packages for other instruments.

The fiber and slit categories of spectral data generally begin with two dimensional CCD images from which one dimensional spectra are extracted. This means that the generic tools collected in the package come from both the **apextract** and **onedspec** packages. In principle one could include the basic CCD reduction commands (from the IRAF **ccdred** package) and also include CCD reductions as part of the integrated reduction tasks; currently this is not done. Generally, observers separate the basic CCD operations, which may be done as the data is collected or as a separate data reduction session, from the spectral data reductions.

In the first category are spectrographs using optical fibers. These types of instruments have become increasingly common because they have many important features, such as obtaining many spectra per exposure, including a mix of object, sky, and arc spectra, and using very stable spectrographs located off the telescope. There are a number of such instruments within NOAO as well as at other observatories.

Argus is a multifiber instrument at Cerro Tololo Interamerican Observatory (CTIO) which positions fibers with a "fishing pole" style system in which each pole has an object and sky fiber with a fixed separation. *Hydra* is a multifiber instrument at Kitt Peak National Observatory (KPNO) which positions fibers with a robot positioner. *Nessie* is an older KPNO multifiber system using manually wired plugboards. The *Fiber Optic Echelle* is a two fiber (simultaneous object and arc spectra), echelle spectrograph on loan to KPNO. The image format consists of pairs of echelle orders. The **fiber** package refers to a three fiber configuration at the KPNO coude feed where two fibers can be used for arc spectra while one fiber is used for an object. The **msred** package provides a generic package for dealing with multiple spectra in a 2D format. While this is not limited to fibers it is most often used for that type of spectra.

The second category of slit reduction packages is specifically for stellar slit spectra which have been separately distortion corrected or which do not require distortion corrections. This class of spectra are very similar in their reduction requirements and the packages and tasks differ mostly in default parameters. The **echelle** package falls in this category because the reduction operations are basically the same as for single slit spectra but the operations are performed over multiple orders and, of course, an echelle dispersion function is used.

The scanner category consists of instruments which record one dimensional aperture spectra directly. These differ from the other categories mostly in requiring special beam-switch operations. At NOAO the two instruments are the *IIDS* and *IRS* scanners.

3. The Specialized Spectroscopy Reduction Tasks

For some astronomers having the appropriate tools in a package is sufficient. They prefer to reduce their data one step at a time with control over all the options of the general tools. However, other astronomers prefer to have a single program which handles all the data reductions in a standard way with a minimum of user interaction. Some reasons for this are for preliminary (often called quick-look) reductions, for processing large amounts of data uniformly and quickly, for repetitive operations which require many steps using the general tools, and for beginning or infrequent users.

Most of the instrument and data specific packages now have an integrated reduction task combining individual reduction steps and providing some degree of automation. These tasks are IRAF command language (CL) scripts for the reasons described previously. The reduction tasks have the following goals:

1. Perform all the reduction operations in an integrated task
2. Guide the user through the reduction steps
3. Address instrument and data specific requirements
4. Collect and limit the number of parameters to those specific to the problem
5. Perform interactive reference calibrations first
6. Perform bulk of processing automatically and noninteractively
7. Perform record keeping and processing checks

Though the tasks provide a complete reduction path (excluding basic CCD calibrations), these integrated tasks are still quite flexible. Task parameters allow selecting just one operation, a set of operations, or all operations. Thus, one still has the option of doing the reductions a step at a time or all at once.

An important point about instrument and data specific reduction tasks is they can contain functions for dealing with specific features or peculiarities of a particular type of data or instrument. Examples of this are computing fiber throughputs or merging pairs of arc observations because a single arc may not have sufficient lines or fully illuminate all the fibers in a multifiber spectrograph. Though these specialized functions may not be addressed by any single generic task they can be performed using a combination of general image processing tasks and parts of the generic spectroscopy tasks. The reduction scripts naturally allow including these specialized, compound commands saving the user many separate steps.

A necessary aspect and virtue of the very powerful and general generic tasks is that the user is given a choice of many parameters to specifying data dependent properties and algorithm options. For some this makes the tasks overly complex; even with the IRAF parameter mechanism which gives most parameters recommended default values. Furthermore, when there are many such tasks involved in the data reduction process the number of parameters, distributed among different parameter sets, to be kept track of becomes quite large. The integrated, instrument and data specific reduction tasks can reduce and simplify the parameters by collecting all the parameters in one place, combining parameters which are common to several tasks into a single parameter, and hiding parameters which have fixed values or which are not applicable to the particular instrument or data.

One of the features of data reductions when there is a lot of data is that once calibration information is set up much of the work can be done noninteractively or even as a background or batch process. Also, for a consistent set of data some of the calibration data can be defined noninteractively by reference to other similar calibration data. The prime example of this is wavelength calibration using arc lamp spectra. Once the arc lines have been identified in one arc spectrum other arc spectra will be sufficiently similar that the computer can find the lines and determine new wavelength calibrations. In the data reduction scripts the user is guided to do the required interactive operations first and only once; if a step has been done previously then it is not repeated. When all the required interactive calibration steps have been accomplished, the bulk of the data can be processed noninteractively and the script may even submit a background or batch job.

When doing many reduction operations on many observations, keeping track of what has been done and what is left to be done can become a significant problem. This problem is compounded by instruments which multiplex many observations simultaneously. The best example of this is multifiber spectrographs which can produce a hundred or more spectra per exposure. The integrated reduction tasks can be designed to do this record keeping consistently and to include processing checks to prevent reprocessing or out of order operations.

4. Spectral Reduction Outline

This section discusses the types of reduction operations performed by the specialized reduction tasks excluding the one dimensional scanner instruments. The description roughly follows the actual flow of the programs. Space does not allow much usage and algorithmic detail. Table 3 shows the basic reduction operations provided by the tasks. These basic operations generally consist of a number of steps. In addition, there are setup and display steps.

Table 3. Basic Spectroscopy Reduction Operations

Fibers	Slits	Scanners
Extract	Extract and sky subtract	Flat field
Flat field	Wavelength calibrate	Wavelength calibrate
Wavelength Calibrate	Flux calibrate	Beam switch
Sky Subtract		Flux calibrate

4.1. Setup Steps

The setup steps consist of setting parameters and preparing input lists. The parameters are easily reviewed and set with the IRAF parameter editor. The parameters likely to be adjusted are those which vary due to the CCD used, such as pixel scales and orientations, and selecting the operations desired. The types of input lists which might need to be prepared are long lists of input

images, aperture identifications for multifiber data, and arc assignment lists if the default assignment by observation time is not desired.

The aperture identification list for multifiber data is a very valuable feature. This list defines the aperture numbers and object titles for each fiber. It also defines which fibers are object, sky, arc, or other type of spectra. For data with many fibers, record keeping becomes a major concern and the initial preparation of this list allows the identifications of the spectra to be maintained through any sorting and analysis of the spectra. In the future there may be instruments which produce this file automatically for the user.

4.2. Aperture Definitions

The first step in all the reduction tasks is to define a reference set of apertures. Once done the reference apertures are automatically used to set the apertures in all other spectra. The reference apertures may be used unchanged or as a starting point for automatic recentering, resizing, or even retracing (determining the path of the spectrum along the dispersion through the 2D image). While the intent is that only the initial reference apertures have an interactive component it is also possible to review the apertures for each spectrum if desired. For multifiber data with hundreds of spectra it is clearly important not to have to do this for each observation.

An initial set of apertures is determined automatically based on task parameters such as the number of spectra expected. The user may then review the definitions graphically. Since this step is usually only performed once it would be wise to check such things as the aperture assignments, the background regions, and the apertures widths. For multifiber instruments, which may suffer from missing (broken) fibers, this is important to avoid misidentifications of spectra.

The positions of the spectra along the dispersion are fit by a smooth function. The user may review the the function fits to the measured positions and adjust the fitting parameters. While this may be done for all apertures it is usually only necessary to do this for one aperture and then apply the same parameters to all other apertures. Again, for large numbers of fibers, it is important to not have to look at each fiber except if truly desired.

4.3. Flat Fielding

A flat fielding step is part of the fiber data reduction tasks. This step is optional if one wants to flat field the CCD images separately. However, flat fielding of fiber data by two dimensional division is very sensitive to even a small fraction of a pixel shift in the fiber positions. When there are no shifts it is entirely equivalent to divide the extracted spectra by the extracted flat field spectra. While it is not equivalent when there are small shifts, the approximation is close and preferable to dealing with artifacts caused by dividing misaligned flat field and object fiber spectra.

A flat field is also used for determining fiber throughput corrections. Alternatively, a sky flat field may be used for the throughput corrections if the illumination of the fibers is different than the flat field illumination. Only the relative total flux through each fiber is used so a high S/N sky exposure is not required for this step.

When the flat field spectra are extracted the user has the option of fitting a smooth function to the average of the flat field fiber spectra and then dividing the individual spectra by this smooth mean. This removes the large scale shape of the flat field spectrum and, thereby, approximately preserves the pixel statistics of the object images.

The operations performed in this step are to extract the flat field fiber spectra, possibly remove the average spectrum shape, extract the sky flat field spectra if given, compute throughput corrections to be incorporated in the final flat field spectra, and finally normalize so the mean flat field through all fibers is unity. Only the fitting of the average flat field shape is possibly interactive. Once the flat field has been derived this step is not done again when the task is repeatedly executed.

The fiber flat fielding algorithm, which requires many operations using the generic image processing tasks, is an example of special operations that can be implemented for data and instrument specific packages and reduction tasks.

4.4. Wavelength Calibration

Virtually all spectroscopic data needs to be wavelength calibrated. In the specialized reduction tasks this step is divided into a initial setup step, which is partly interactive, and then new calibrations are derived and applied as needed when the object spectra are processed.

The initial calibration setup consists of extracting one arc calibration reference spectrum and having the user identify the lines (only a subset need be identified and further lines are added automatically) and define the dispersion function to be fit to the wavelengths as a function of pixel position. For multifiber spectra the other fiber arc spectra are extracted and automatically used to find lines and determine dispersion functions. This step has the option of interactive review if it appears arc lines are misidentified. At this point a common linear dispersion system is defined. All spectra will then have the same dispersion sampling. Note that in V2.10 it is not necessary to linearize the spectra though it might still be desirable.

In the instrument and data categories, echelle data was not identified as a separate class. This is because the only significant step which is different is computing and applying an echelle dispersion function (a function of pixel position and order number) instead of a one dimensional function. In the reduction scripts it was found that one could simply substitute the echelle dispersion task and otherwise the logical flow of the processing remained unchanged.

For fiber spectra there is the possibility of using a different type of calibration to monitor zero point shifts in the wavelengths. This is an another example of a special purpose algorithm. For the *Nessie* instrument, because it is time consuming to take many arc lamp exposures due to manually moving the fibers to a small region in a special plugboard for illumination, users can take quick exposures of scattered light from the dome lights (mercury lines) or possibly sky lines, and then select a set of such lines to be monitored and used for making wavelength zero point corrections.

Another special purpose algorithm for *Nessie* data is to combine two or more arc exposures in which only some fibers are illuminated in each exposure. This is because the arc illumination covers only a small region of the field. By

creating a simple file giving which fibers in each exposure are to be used this instrumental peculiarity is dealt with.

4.5. Flux Calibration

The slit instrument reduction tasks include flux calibration as one of the processing options. This option is not included in the fiber tasks because there has been no demand for it; it could be added if desired.

Flux calibration consists of extracting *all* the standard star observations first and computing a sensitivity function from the ratio of observed verse reference calibration fluxes. The standard star extractions follow the same steps as the object processing except they are done before any objects and, of course, they are not initially flux calibrated. After each standard star is extracted a query is made for the name of the standard star. It would be nice to have this done automatically. Currently it is up to the observer to identify which objects were observed since image header information is subject to error. There is an optional interactive step to allow reviewing and modifying the bandpasses used for the calibrations. Typically, the same bandpasses as used in the reference calibration observations are used so that no interpolation is required. However, if the dispersion is a poor match to the calibration data it may be necessary to define additional bandpasses.

The ratios of the observed fluxes through a set of bandpasses to the fluxes found in the calibration data is recorded in a file. After all the standard stars have been processed the entire set of ratios is fit by a smooth function of wavelength. This is a very interactive step in order to evaluate the photometric conditions and define an acceptable sensitivity function. The sensitivity function is recorded as a spectrum which is then applied to object spectra as needed.

4.6. Object Processing

After the various calibration steps have been completed, the object spectra are processed. One goal of the reduction tasks is to keep track of the processing state of each observation. Thus, if a specified input spectrum has been previously processed it will be skipped or reprocessed depending on a user set processing option. The ability to skip processed spectra is particularly useful in allowing wildcard input lists so that the user doesn't have to keep track of which spectra have been processed.

Another goal of the reduction tasks is to allow the object processing to be done completely noninteractively; though with options to allow interactive review and intervention if desired. The tasks keep track of whether interactive steps have been selected and allow these steps to be turned off during processing. When all interactive steps have been turned off the task may automatically submit a background or batch process if the user has selected this option.

In the following the processing steps are described in the order in which they are performed. Most of the steps have parameters which allow selecting or skipping each step. The first step is assigning the apertures. This is done by inheriting the reference apertures and providing task parameters to select recentering of the apertures based on a part of the spectrum, resizing the aperture widths to accommodate seeing changes, interactively reviewing the apertures, and retracing based on the object being extracted. Whether these steps are

done depends on the stability of the instrument and characteristics such as the strength of the spectra; for example, it would be worse to retrace very weak spectra than to assume it is the same as the reference spectrum.

The spectra are then extracted, with sky subtraction in the case of slit spectra. This step may also include options such as optimal weighting and cosmic ray detection. If a 1D fiber flat field has been computed it is divided into the extracted object spectra.

If wavelength calibration has been selected, calibration arc spectra are assigned. This is usually the two nearest in time but any selection may be made with an input table. The assigned arc spectra are extracted if needed. Note that extraction of arc spectra is only done as needed and later objects using the same arc spectra may not need to extract them again. However, for slit spectra one generally wants to use the same apertures as the object spectra. Therefore, if the object apertures are recentered or retraced the user may choose to have the arc spectra be reextracted with the new apertures. For fiber spectra the reference aperture is generally used.

The arc lines are automatically reidentified and new dispersion functions are fit. For the special algorithm using night sky or dome vapor lamp lines for tracking zero point shifts it is these type of spectra that are extracted and the new dispersion function is the original arc dispersion function with a zero point shift applied. Alternatively, some fiber instruments provide one or more fibers to take simultaneous arc spectra. In this case the spectra from these fibers are compared against the same fibers in the master arc calibration to determine zero point changes while keeping the same dispersion shape.

The one or more dispersion calibration spectra are then interpolated either in time or spatially in the case of simultaneous arc fibers. The final interpolated dispersion functions are then applied to the object spectra. This application consists of either resampling the spectra to a common linear dispersion sampling or, in IRAF V2.10, the actual function may be stored with the spectrum and no interpolation done.

Sky subtraction with fiber spectra requires selecting the sky spectra, combining them, and subtracting the final sky from the object fibers. The sky spectra selection is done based on aperture or beam numbers. It has an interactive option to plot all the sky spectra and allow deletion of those which might be contaminated by objects. The combining options include various choices of algorithms such as medians or sigma clipping.

Flux calibration is a straightforward, noninteractive, application of extinction corrections and the sensitivity functions.

An interactive option, which might be selected if the reduction tasks are used at the telescope for quick reductions is to automatically plot the final spectra using an interactive spectra plotting task.

References

There are a variety of documents covering the IRAF NOAO spectroscopy packages. There are overviews of the packages, detailed descriptions for each task, and user's guides for each of the specialized instrument reduction tasks.