# reduce Users Manual

***Release X1***

**Kennneth Anderson**

October 27, 2014

# CONTENTS

CHAPTER

ONE

# INTRODUCTION

This document is version 1.0 of the `reduce` Users Manual. This manual will describe the usage of `reduce` as an application provided by the Gemini Observatory Astrodata package suite. `reduce` is an application that allows users to invoke the Gemini Recipe System to perform data processing and reduction on one or more astronomical datasets.

This document presents details on applying `reduce` to astronomical datasets, currently defined as multi-extension FITS (MEF) files, both through the application's command line interface and the application programming interface (API). Details and information about the `astrodata` package, the Recipe System, and/or the data processing involved in data reduction are beyond the scope of this document and will only be engaged when directly pertinent to the operations of `reduce`.

## 1.1 Reference Documents

- *The Gemini Recipe System: a dynamic workflow for automated data reduction*, K. Labrie *et al*, SPIE, 2010.
- *Developing for Gemini's extensible pipeline environment*, K. Labrie, C. Allen, P. Hirst, ADASS, 2011
- *Gemini's Recipe System; A publicly available instrument-agnostic pipeline infrastructure*, K. Labrie et al, ADASS 2013.

## 1.2 Overview

As an application, `reduce` provides interfaces to configure and launch the Gemini Recipe System, a framework for developing and running configurable data processing pipelines and which can accommodate processing pipelines for arbitrary dataset types. In conjunction with the development of `astrodata`, Gemini Observatory has also developed the compatible `astrodata_Gemini` package, the code base currently providing abstraction of, and processing for, Gemini Observatory astronomical observations.

In Gemini Observatory's operational environment "on summit," `reduce`, `astrodata`, and the `astrodata_Gemini` packages provide a currently defined, near-realtime, quality assurance pipeline, the so-called QAP. `reduce` is used to launch this pipeline on newly acquired data and provide image quality metrics to observers, who then assess the metrics and apply observational decisions on telescope operations.

Users unfamiliar with terms and concepts heretofore presented should consult documentation cited in the previous sections (working on the Recipe System User Manual).

## 1.3 Glossary

**adcc** – Automatated Data Communication Center. Provides XML-RPC and HTTP services for pipeline operations. Can be run externally to `reduce`. Users need not know about or invoke the `adcc` for `reduce` operations. `reduce` will launch an `adcc` instance if one is not available. See Sec. *The adcc* for further discussion on `adcc`.

**astrodata** (or Astrodata) – part of the **gemini_python** package suite that defines the dataset abstraction layer for the Recipe System.

**AstroData** – not to be confused with **astrodata**, this is the main class of the `astrodata` package, and the one most users and developers will interact with at a programmatic level.

**AstroDataType** – Represents a data classification. A dataset will be classified by a number of types that describe both the data and its processing state. The AstroDataTypes are hierarchical, from generic to specific. For example, a typical GMOS image might have a set of types like

'GMOS_S', 'GMOS_IMAGE', 'GEMINI', 'SIDEREAL', 'IMAGE', 'GMOS', 'GEMINI_SOUTH', 'GMOS_RAW', 'UNPREPARED', 'RAW' (see **types** below).

**astrodata_Gemini** – the **gemini_python** package that provides all observatory specific definitions of data types, **recipes**, and associated **primitives** for Gemini Observatory data.

**astrodata_X** – conceivably a data reduction package that could reduce other observatory and telescope data. Under the Astrodata system, it is entirely possible for the Recipe System to process HST or Keck data, given the development of an associated package, astrodata_HST or astrodata_Keck. Pipelines and processing functions are defined for the particulars of each telescope and its various instruments.

**Descriptor** – Represents a high-level metadata name. Descriptors allow access to essential information about the data through a uniform, instrument-agnostic interface to the FITS headers.

**gemini_python** – A suite of packages comprising **astrodata**, **astrodata_Gemini**, **astrodata_FITS**, and **gempy**, all of which provide the full functionality needed to run **Recipe System** pipelines on observational datasets.

**gempy** – a **gemini_python** package comprising functional utilities to the **astrodata_Gemini** package.

**MEF** – Multiple Extension FITS, the standard data format not only for Gemini Observatory but many observatories.

**primitive** – A function defined within an **astrodata_[X]** package that performs actual work on the passed dataset. Primitives observe tightly controlled interfaces in support of re-use of primitives and recipes for different types of data, when possible. For example, all primitives called `flatCorrect` must apply the flat field correction appropriate for the data's current AstroDataType, and must have the same set of input parameters.

**recipe** – Represents the sequence of transformations. A recipe is a simple text file that enumerates the set and order of **primitives** that will process the passed dataset. A **recipe** is the high-level pipeline definition. Users can pass recipe names directly to reduce. Essentially, a recipe is a pipeline.

**Recipe System** – The gemin_python framework that accommodates an arbitrary number of defined recipes and the primitives

**reduce** – The user/caller interface to the Recipe System and its associated recipes/pipelines.

**type** or **typeset** – Not to be confused with language primitive or programmatic data types, these are data types defined within an **astrodata_[X]** package used to describe the kind of observational data that has been passed to the Recipe System., Eg., GMOS_IMAGE, NIRI. In this document, these terms are synonymous with **AstroDataType** unless otherwise indicated.

CHAPTER

TWO

# INSTALLATION

The `astrodata` package has several dependencies like `numpy`, `astropy`, and others. All dependencies of `gemini_python` and `astrodata` are provide by the Ureka package, and users are highly encouraged to install and use this very useful package. It is an easy and, perhaps, best way to get everything you need and then some. Ureka is available at http://ssb.stsci.edu/ureka/.

Once a user has has retrieved the gemini_python package, available as a tarfile from the Gemini website (http://gemini.edu), and untarred only minor adjustments need to be made to the user environment in order to make astrodata importable and allow `reduce` to work properly.

## 2.1 Install

### 2.1.1 Recommended Installation

It is recommended to install the software in a location other than the standard python location for modules (the default `site-packages`). This is also the only solution if you do not have write permission to the default `site-packages`. Here is how you install the software somewhere other than the default location:

```
$ python setup.py install --prefix=/your/favorite/location
```

`/your/favorite/location` must already exist. This command will install executable scripts in a `bin` subdirectory, the documentation in a `share` subdirectory, and the modules in a `lib/python2.7/site-packages` subdirectory. The modules being installed are `astrodata`, `astrodata_FITS`, `astrodata_Gemini`, and `gempy`. In this manual, we will only use `astrodata`.

Because you are not using the default location, you will need to add two paths to your environment. You might want to add the following to your .cshrc or .bash_profile, or equivalent shell configuration script.

C shell(csh, tcsh):

```
setenv PATH /your/favorite/location:${PATH}
setenv PYTHONPATH /your/favorite/location:${PYTHONPATH}
```

Bourne shells (sh, bash, ksh, ...)

```
export PATH=/your/favorite/location:${PATH}
export PYTHONPATH=/your/favorite/location:${PYTHONPATH}
```

If you added those lines to your shell configuration script, make sure your `source` the file to activate the new setting.

For csh/tcsh:

```
$ source ~/.cshrc
$ rehash
```

For bash:

```
$ source ~/.bash_profile
```

### 2.1.2 Installation under Ureka

Assuming that you have installed Ureka and that you have write access to the Ureka directory, this will install `astrodata` in the Ureka `site-packages` directory. WARNING: While easier to install and configure, this will modify your Ureka installation.

```
$ python setup.py install
```

This will also add executables to the Ureka `bin` directory and documentation to the Ureka `share` directory.

With this installation scheme, there is no need to add paths to your environment. However, it is a lot more complicated to remove the Gemini software in case of problems, or if you just want to clean it out after evaluation.

In tcsh, you will need to run `rehash` to pick the new executables written to `bin`.

## 2.2 Test the installation

Start up the python interpreter and import astrodata:

```
$ python
>>> import astrodata
```

Next, return to the command line and test that `reduce` is reachable and runs. There may be some delay as package modules are byte compiled:

```
$ reduce -h [--help]
```

This will print the reduce help to the screen.

If users have Gemini fits files available, they can test that the Recipe System is functioning as expected with a test recipe provided by the astrodata_Gemini package:

```
$ reduce --recipe test_one /path/to/gemini_data.fits
```

If all is well, users will see something like:

```
Resetting logger for application: reduce
Logging configured for application: reduce
                      --- reduce, v4890  ---
              Running under astrodata Version GP-X1
All submitted files appear valid
Starting Reduction on set #1 of 1

  Processing dataset(s):
        gemini_data.fits

================================================================================
RECIPE: test_one
================================================================================
 PRIMITIVE: showParameters
```

```
 -------------------------
 rtf = False
 suffix = '_scafaasled'
 otherTest = False
 logindent = 3
 logfile = 'reduce.log'
 reducecache = '.reducecache'
 storedcals = 'calibrations/storedcals'
 index = 1
 retrievedcals = 'calibrations/retrievedcals'
 cachedict = {'storedcals': 'calibrations/storedcals', 'retrievedcals':
              'calibrations/retrievedcals', 'calibrations': 'calibrations',
              'reducecache': '.reducecache'}
 loglevel = 'stdinfo'
 calurl_dict = {'CALMGR': 'http://fits/calmgr',
                'UPLOADPROCCAL': 'http://fits/upload_processed_cal',
                'QAMETRICURL': 'http://fits/qareport',
                'QAQUERYURL': 'http://fits/qaforgui',
                'LOCALCALMGR': 'http://localhost:%(httpport)d/calmgr/%(caltype)s'}
 logmode = 'standard'
 test = True
 writeInt = False
 calibrations = 'calibrations'
 .
Wrote gemini_data.fits in output directory


reduce completed successfully.
```

Users curious about the URLs in the example above, i.e. `http://fits/...`, see Sec. *Fits Storage* in Chapter 5, Discussion.

CHAPTER

THREE

# INTERFACES

## 3.1 Introduction

The `reduce` application provides a command line interface and an API, both of which can configure and launch a Recipe System processing pipeline (a 'recipe') on the input dataset. Control of `reduce` and the Recipe System is provided by a variety of options and switches. Of course, all options and switches can be accessed and controlled through the API.

## 3.2 Command line interface

We begin with the command line help provided by `reduce --help`, followed by further description and discussion of certain non-trivial options that require detailed explanation.

```
usage: reduce [options] fitsfile [fitsfile ...]
```

positional arguments:

```
fitsfile [fitsfile ...]
```

The [options] are described in the following sections.

### 3.2.1 Informational switches

**-h, –help** show the help message and exit

**-v, –version** show program's version number and exit

**-d, –displayflags** Display all parsed option flags and exit.

When specified, this switch will present the user with a table of all parsed arguments and then exit without running. This allows the user to check that the configuration is as intended. The table provides a convenient view of all passed and default values. Unless a user has specified a recipe (-r, –recipe), 'recipename' indicates 'None' because at this point, the Recipe System has not yet been engaged and a default recipe not yet determined.

Eg.,:

```
$ reduce -d --logmode console fitsfile.fits

--------------------   switches, vars, vals  --------------------

Literals                  var 'dest'              Value
```

```
------------------------------------------------------------------
['--invoked']            :: invoked              :: False
['--addprimset']         :: primsetname          :: None
['-d', '--displayflags'] :: displayflags         :: True
['-p', '--param']        :: userparam            :: None
['--logmode']            :: logmode              :: ['console']
['-r', '--recipe']       :: recipename           :: None
['--throw_descriptor_exceptions'] :: throwDescriptorExceptions :: False
['--logfile']            :: logfile              :: reduce.log
['-t', '--astrotype']    :: astrotype            :: None
['--override_cal']       :: user_cals            :: None
['--context']            :: running_contexts     :: None
['--calmgr']             :: cal_mgr              :: None
['--suffix']             :: suffix               :: None
['--loglevel']           :: loglevel             :: stdinfo
------------------------------------------------------------------

Input fits file(s):      fitsfile.fits
```

### 3.2.2 Configuration Switches, Options

**–addprimset <PRIMSETNAME>** Add this path to user-supplied primitives for reduction. eg., path to a primitives module.

**–calmgr <CAL_MGR>** This is a URL specifying a calibration manager service. A calibration manager overides Recipe System table. Not available outside Gemini operations.

**–context <RUNNING_CONTEXTS>** Use <RUNNING_CONTEXTS> for primitives sensitive to context. Eg., –context QA When not specified, the context defaults to 'QA'.

**–invoked** Boolean indicating that reduce was invoked by the control center.

**–logmode <LOGMODE>** Set logging mode. One of 'standard', 'console', 'quiet', 'debug', or 'null', where 'console' writes only to screen and 'quiet' writes only to the log file. Default is 'standard'.

**–logfile <LOGFILE>** Set the log file name. Default is 'reduce.log' in the current directory.

**–loglevel <LOGLEVEL>** Set the verbose level for console logging. One of 'critical', 'error', 'warning', 'status', 'stdinfo', 'fullinfo', 'debug'. Default is 'stdinfo'.

**–override_cal <USER_CALS [USER_CALS ...]>** The option allows users to provide their own calibrations to `reduce`. Add a calibration to User Calibration Service. '–override_cal CALTYPE:CAL_PATH' Eg.,:

> –override_cal processed_arc:wcal/gsTest_arc.fits

**-p <USERPARAM [USERPARAM ...]>, –param <USERPARAM [USERPARAM ...]>** Set a primitive parameter from the command line. The form '-p par=val' sets the parameter in the reduction context such that all primitives will 'see' it. The form

```
-p ASTROTYPE:primitivename:par=val
```

sets the parameter such that it applies only when the current reduction type (type of current reference image) is 'ASTROTYPE' and the primitive is 'primitivename'. Separate parameter-value pairs by whitespace: (eg. '-p par1=val1 par2=val2')

See Sec. *Overriding Primitive Parameters*, for more information on these values.

**-r <RECIPENAME>, –recipe <RECIPENAME>** Specify an explicit recipe to be used rather than internally determined by a dataset's <ASTROTYPE>. Default is None and later determined by the Recipe System based on the AstroDataType.

**-t <ASTROTYPE>, –astrotype <ASTROTYPE>** Run a recipe based on this AstroDataType, which overrides default type or begins without initial input. Eg., recipes that begin with primitives that acquire data. `reduce` default is None and determined internally.

**–suffix <SUFFIX>** Add 'suffix' to output filenames at end of reduction.

**–throw_descriptor_exceptions** Boolean indicating descriptor exceptions are to be raised. This is a development switch.

### 3.2.3 Nominal Usage

The minimal call for reduce can be

```
$ reduce <dataset.fits>
```

While this minimal call is available at the Gemini Observatory (see Sec. *Fits Storage*), if a calibration service is unavailable to the user – likely true for most users – users should call `reduce` on a specified dataset by providing calibration files with the –overrride_cal option.

For example:

```
$ reduce --override_cal processed_arc:wcal/gsTest_arc.fits <dataset.fits>
```

Such a command for complex processing of data is possible because AstroData and the Recipe System do all the necessary work in determining how the data are to be processed, which is critcially based upon the determination of the *typeset* that applies to that data.

Without any user-specified recipe (-r –recipe), the default recipe is `qaReduce`, which is defined for various AstroDataTypes and currently used at the summit. For example, the `qaReduce` recipe for a GMOS_IMAGE specifies that the following primitives are called on the data:

```
prepare
addDQ
addVAR
detectSources
measureIQ
measureBG
measureCCAndAstrometry
overscanCorrect
biasCorrect
ADUToElectrons
addVAR
flatCorrect
mosaicDetectors
makeFringe
fringeCorrect
detectSources
measureIQ
measureBG
measureCCAndAstrometry
addToList
```

The point here is not to overwhelm readers with a stack of primitive names, but to present both the default pipeline processing that the above simple `reduce` command invokes and to demonstrate how much the `reduce` interface abstracts away the complexity of the processing that is engaged with the simplist of commands.

### 3.2.4 Overriding Primitive Parameters

In some cases, users may wish to change the functional behaviour of certain processing steps, i.e. change default behaviour of primitive functions.

Each primitive has a set of pre-defined parameters, which are used to control functional behaviour of the primitive. Each defined parameter has a "user override" token, which indicates that a particular parameter may be overridden by the user. Users can adjust parameter values from the reduce command line with the option,

**-p, –param**

If permitted by the "user override" token, parameters and values specified through the **-p, –param** option will *override* the defined parameter default value and may alter default behaviour of the primitive accessing this parameter. A user may pass several parameter-value pairs with this option.

Eg.:

```
$ reduce -p par1=val1 par2=val2 [par3=val3 ... ] <fitsfile1.fits>
```

For example, some photometry primitives perform source detection on an image. The 'detection threshold' has a defined default, but a user may alter this parameter default to change the source detection behaviour:

```
$ reduce -p threshold=4.5 <fitsfile.fits>
```

### 3.2.5 The @file facility

The reduce command line interface supports what might be called an 'at-file' facility (users and readers familiar with IRAF will recognize this facility). This facility allows users to provide any and all command line options and flags to `reduce` via in a single acsii text file.

By passing an @file to `reduce` on the command line, users can encapsulate all the options and positional arguments they might wish to specify in a single @file. It is possible to use multiple @files and even to embed one or more @files in another. The parser opens all files sequentially and parses all arguments in the same manner as if they were specified on the command line. Essentially, an @file is some or all of the command line and parsed identically.

To illustrate the convenience provided by an '@file', let us begin with an example *reduce* command line that has a number of arguments:

```
$ reduce -p GMOS_IMAGE:contextReport:tpar=100 GMOS_IMAGE:contextReport:report_inputs=True
  -r recipe.ArgsTest --context qa S20130616S0019.fits N20100311S0090.fits
```

Ungainly, to be sure. Here, two (2) *user parameters* are being specified with **-p**, a *recipe* with **-r**, and a *context* argument is specified to be **qa** . This can be wrapped in a plain text @file called *reduce_args.par*:

```
S20130616S0019.fits
N20100311S0090.fits
--param
GMOS_IMAGE:contextReport:tpar=100
GMOS_IMAGE:contextReport:report_inputs=True
-r recipe.ArgsTests
--context qa
```

This then turns the previous reduce command line into something a little more *keyboard friendly*:

```
$ reduce @reduce_args.par
```

The order of these arguments is irrelevant. The parser will figure out what is what. The above file could be thus written like:

```
-r recipe.ArgsTests
--param
GMOS_IMAGE:contextReport:tpar=100
GMOS_IMAGE:contextReport:report_inputs=True
--context qa
S20130616S0019.fits
N20100311S0090.fits
```

---

**Note:** Comments are accommodated, both line and in-line. '=' signs *may* be used but this has meaning only for arguments that expect unitary values. The '=' is entirely unnecessary.

White space is the only significant separator of arguments: spaces, tabs, newlines are all equivalent when argument parsing. This means the user can 'arrange' their @file for clarity.

Eg., a more readable version of the above file might be written as:

```
# reduce parameter file
# yyyy-mm-dd
# GDPSG

# Spec the recipe
-r
    recipe.ArgsTests  # test recipe

# primitive parameters here
# These are 'untyped', i.e. global
--param
    tpar=100
    report_inputs=True

--context
    qa                # QA context

S20130616S0019.fits
N20100311S0090.fits
```

---

All the above examples of `reduce_args.par` are equivalently parsed. Which, of course, users may check by adding the **-d** flag:

```
$ reduce -d @redpars.par

--------------------   switches, vars, vals  --------------------

Literals                      var 'dest'              Value
------------------------------------------------------------------
['--invoked']                 :: invoked              :: False
['--addprimset']              :: primsetname          :: None
['-d', '--displayflags']      :: displayflags         :: True
['-p', '--param']             :: userparam            :: ['tpar=100', 'report_inputs=True']
['--logmode']                 :: logmode              :: standard
['-r', '--recipe']            :: recipename           :: ['recipe.ArgTests']
['--throw_descriptor_exceptions'] :: throwDescriptorExceptions        :: False
['--logfile']                 :: logfile              :: reduce.log
['-t', '--astrotype']         :: astrotype            :: None
['--override_cal']            :: user_cals            :: None
['--context']                 :: running_contexts     :: ['QA']
['--calmgr']                  :: cal_mgr              :: None
['--suffix']                  :: suffix               :: None
```

```
['--loglevel']                :: loglevel             :: stdinfo
------------------------------------------------------------------

Input fits file(s):   S20130616S0019.fits
Input fits file(s):   N20100311S0090.fits
```

### 3.2.6 Recursive @file processing

As implemented, the @file facility will recursively handle, and process correctly, other @file specifications that appear in a passed @file or on the command line. For example, we may have another file containing a list of fits files, separating the command line flags from the positional arguments.

We have a plain text 'fitsfiles' containing the line:

```
test_data/S20130616S0019.fits
```

We can indicate that this file is to be consumed with the prefix character "@" as well. In this case, the 'reduce_args.par' file could thus appear:

```
# reduce test parameter file

@fitsfiles       # file with fits files

# AstroDataType
-t GMOS_IMAGE

# primitive parameters.
--param
    report_inputs=True
    tpar=99
    FOO=BAR

# Spec the recipe
-r recipe.ArgTests
```

The parser will open and read the @fitsfiles, consuming those lines in the same way as any other command line arguments. Indeed, such a file need not only contain fits files (positional arguments), but other arguments as well. This is recursive. That is, the @fitsfiles can contain other at-files", which can contain other "at-files", which can contain ..., *ad infinitum*. These will be processed serially.

As stipulated earlier, because the @file facility provides arguments equivalent to those that appear on the command line, employment of this facility means that a reduce command line could assume the form:

```
$ reduce @parfile @fitsfiles
```

or equally:

```
$ reduce @fitsfiles @parfile
```

where 'parfile' could contain the flags and user parameters, and 'fitsfiles' could contain a list of datasets.

Eg., fitsfiles comprises the one line:

```
test_data/N20100311S0090.fits
```

while parfile holds all other specifications:

```
# reduce test parameter file
# GDPSG

# AstroDataType
-t GMOS_IMAGE

# primitive parameters.
--param
    report_inputs=True
    tpar=99            # This is a test parameter
    FOO=BAR            # This is a test parameter

# Spec the recipe
-r recipe.ArgTests
```

### 3.2.7 Overriding @file values

The `reduce` application employs a customized command line parser such that the command line option

**-p** or **–param**

will accumulate a set of parameters *or* override a particular parameter. This may be seen when a parameter is specified in a user @file and then specified on the command line. For unitary value arguments, the command line value will *override* the @file value.

It is further specified that if one or more datasets (i.e. positional arguments) are passed on the command line, *all fits files appearing as positional arguments in the parameter file will be replaced by the command line arguments.*

Using the parfile above,

Eg. 1) Accumulate a new parameter:

```
$ reduce @parfile --param FOO=BARSOOM

parsed options:
--------------------
AstroDataType: GMOS_IMAGE
FITS files:    ['S20130616S0019.fits', 'N20100311S0090.fits']
Parameters:    tpar=100, report_inputs=True, FOO=BARSOOM
RECIPE:        recipe.ArgsTest
```

Eg. 2) Override a parameter in the @file:

```
$ reduce @parfile --param tpar=99

parsed options:
--------------------
AstroDataType: GMOS_IMAGE
FITS files:    ['S20130616S0019.fits', 'N20100311S0090.fits']
Parameters:    tpar=99, report_inputs=True
RECIPE:        recipe.ArgsTest
```

Eg. 3) Override the recipe:

```
$ reduce @parfile -r=recipe.FOO

parsed options:
--------------------
AstroDataType:    GMOS_IMAGE
```

```
FITS files:       ['S20130616S0019.fits', 'N20100311S0090.fits']
Parameters:       tpar=100, report_inputs=True
RECIPE:           recipe.FOO
```

Eg. 4) Override a recipe and specify another fits file

```
$ reduce @parfile -r=recipe.FOO test_data/N20100311S0090_1.fits

parsed options:
--------------------
AstroDataType:    GMOS_IMAGE
FITS files:       ['test_data/N20100311S0090_1.fits']
Parameters:       tpar=100, report_inputs=True
RECIPE:           recipe.FOO
```

## 3.3 Application Programming Interface (API)

**Note:** The following sections discuss and describe programming interfaces available on `reduce` and the underlying class Reduce.

The `reduce` application is essentially a skeleton script providing the described command line interface. After parsing the command line, the script then passes the parsed arguments to its main() function, which in turn calls the Reduce() class constructor with "args". Class Reduce() is defined in the module `coreReduce.py`. `reduce` and class Reduce are both scriptable, as the following discussion will illustrate.

### 3.3.1 reduce.main()

The main() function of reduce receives one (1) parameter that is a Namespace object as returned by a call on ArgumentParser.parse_args(). Specific to reduce, the caller can supply this object by a call on the parseUtils.buildParser() function, which returns a fully defined reduce parser. As usual, the parser object should then be called with the parse_args() method to return a valid reduce parser Namespace. Since there is no interaction with sys.argv, as in a command line call, all Namespace attributes have only their defined default values. It is for the caller to set these values as needed.

As the example below demonstrates, once the "args" Namespace object is instantiated, a caller can set any arguments as needed. Bu they must be set to the correct type. The caller should examine the various "args" types to determine how to set values. For example, args.files is type list, whereas args.recipename is type string.

Eg.,

```
>>> from astrodata.adutils.reduceutils import reduce
>>> from astrodata.adutils.reduceutils import parseUtils
>>> args = parseUtils.buildParser("Reduce,v2.0").parse_args()
>>> args.files
[]
>>> args.files.append('S20130616S0019.fits')
>>> args.recipename = "recipe.FOO"
>>> reduce.main(args)
--- reduce, v2.0 ---
Starting Reduction on set #1 of 1
Processing dataset(s):
S20130616S0019.fits
...
```

Processing will proceed as usual.

### 3.3.2 Class Reduce and the runr() method

Class Reduce is defined in `astrodata.adutils.reduceutils` module, `coreReduce.py`.

The reduce.main() function serves mainly as a callable for the command line interface. While main() is callable by users supplying the correct "args" parameter (See *reduce.main()*), the Reduce() class is also callable and can be used directly, and more appropriately. Callers need not supply an "args" parameter to the class constructor. The instance of Reduce will have all the same arguments as in a command line scenario, available as attributes on the instance. Once an instance of Reduce() is instantiated and instance attributes set as needed, there is one (1) method to call, **runr**(). This is the only public method on the class.

---

**Note:** When using Reduce() directly, callers must configure their own logger. Reduce() does not configure logutils prior to using a logger as returned by logutils.get_logger(). The following example will illustrate how this is easily done. It is *highly recommended* that callers configure the logger.

---

Eg.,

```
>>> from astrodata.adutils.reduceutils.coreReduce import Reduce
>>> reduce = Reduce()
>>> reduce.files
[]
>>> reduce.files.append('S20130616S0019.fits')
>>> reduce.files
['S20130616S0019.fits']
```

Once an instance of Reduce has been made, callers can then configure logutils with the appropriate settings supplied on the instance. This is precisely what `reduce` does when it configures logutils.

```
>>> from astrodata.adutils import logutils
>>> logutils.config(file_name=reduce.logfile, mode=reduce.logmode,
                    console_lvl=reduce.loglevel)
```

At this point, the caller is able to call the runr() method on the "reduce" instance.

```
>>> reduce.runr()
All submitted files appear valid
Starting Reduction on set #1 of 1
Processing dataset(s):
S20130616S0019.fits
...
```

Processing will then proceed in the usual manner.

CHAPTER

FOUR

# SUPPLEMENTAL TOOLS

The astrodata package provides a number of command line driven tools, which users may find helpful in executing reduce on their data.

With the installation and configuration of `astrodata` and `reduce` comes some supplemental tools to help users discover information, not only about their own data, but about the Recipe System, such as available recipes, primitives, and defined AstroDataTypes.

If the user environment has been configured correctly these applications will work directly.

## 4.1 listprimitives

The application `listprimitives` is available as a command line executable. This tool displays available primitives for all AstroDataTypes, their parameters, and defaults. These are the parameters discussed in Sec. *Overriding Primitive Parameters* that can be changed by the user with the **-p, –param** option on reduce. under the AstroDataTypes. The help describes more options:

```
$ listprimitives -h

Usage: listprimitives [options]

Gemini Observatory Primitive Inspection Tool, v1.0 2011

Options:
-h, --help            show this help message and exit
-c, --use-color       apply color output scheme
-i, --info            show more information
-p, --parameters      show parameters
-r, --recipes         list top recipes
-s, --primitive-set   show primitive sets (Astrodata types)
-v, --verbose         set verbose mode
--view-recipe=VIEW_RECIPE
                      display the recipe
```

### 4.1.1 listprimitives information

The following section presents examples of the kind of information that `listprimitives` may provide.

Show available recipes:

```
$ listprimitives -r

================================================================================

RECIPES_Gemini
--------------------------------------------------------------------------------
   1. basicQA
   2. checkQA
   3. makeProcessedArc.GMOS_SPECT
   4. makeProcessedBias
   5. makeProcessedDark
   6. makeProcessedFlat
   7. makeProcessedFlat.GMOS_IMAGE
   8. makeProcessedFlat.GMOS_SPECT
   9. makeProcessedFlat.NIRI_IMAGE
   10. makeProcessedFringe
   11. qaReduce.GMOS_IMAGE
   12. qaReduce.GMOS_SPECT
   13. qaReduce.NIRI_IMAGE
   14. qaReduceAndStack.GMOS_IMAGE
   15. qaStack.GMOS_IMAGE
   16. reduce.F2_IMAGE
   17. reduce.GMOS_IMAGE

Subrecipes
--------------------------------------------------------------------------------
   1. biasCorrect
   2. correctWCSToReferenceCatalog
   3. darkCorrect
   4. flatCorrect
   5. lampOnLampOff
   6. makeSky
   7. overscanCorrect
   8. prepare
   9. skyCorrect
   10. standardizeHeaders
   11. thermalEmissionCorrect

================================================================================
```

Users can request display the contents of a particular recipe, as listed above. This will present the 'stack' of primitives that will be called by the Recipe System when the particular recipe is either specified through the `reduce` command line by the user, or selected internally by the Recipe System itself.

For example, a user may like to see the primitive stack called by the default 'QA' recipe for GMOS_IMAGE data. As seen in the above example, these 'qa' recipes are defined for several AstroDataTypes.

Show the primitive stack for the 'qa' GMOS_IMAGE type:

```
$ listprimitives --view-recipe qaReduce.GMOS_IMAGE

================================================================================
RECIPE: qaReduce.GMOS_IMAGE
================================================================================
# This recipe performs the standardization and corrections needed to convert
# the raw input science images into a single stacked science image

prepare
addDQ
```

```
addVAR(read_noise=True)
detectSources
measureIQ(display=True)
measureBG
measureCCAndAstrometry
overscanCorrect
biasCorrect
ADUToElectrons
addVAR(poisson_noise=True)
flatCorrect
mosaicDetectors
makeFringe
fringeCorrect
detectSources
measureIQ(display=True)
measureBG
measureCCAndAstrometry
addToList(purpose=forStack)


================================================================================
```

`listprimitives` is in need of refinement and work continues on building a tool that will present primitives and parameters in a more focused way, i.e., report only those primitives and parameters relevant to a given dataset. As it currently stands, users can request that `listprimitives` display primitive parameters (as may be passed to `reduce` through the **-p** or **–param** option, Sec. *Overriding Primitive Parameters*), but this results in a list of all AstroDataTypes, their primitives and associated parameters. Admittedly, this list is rather ungainly, but users may see, for example, that the primitive `detectSources` has several user-tunable parameters:

```
detectSources
    suffix: '_sourcesDetected'
    centroid_function: 'moffat'
    threshold: 3.0
    sigma: None
    fwhm: None
    method: 'sextractor'
    max_sources: 50
```

See the discussion in Sec. *Overriding Primitive Parameters* on command line override of primitive parameters, and where overriding the 'threshold' parameter is dicussed specifically.

## 4.2 typewalk

As with `listprimitives` and `reduce`, the command line application `typewalk` becomes available once the user has configured astrodata correctly. `typewalk` examines files in a directory or directory tree and reports the types and status values through the AstroDataType classification scheme. Running `typewalk` on a directory containing some Gemini datasets will demonstrate what users can expect to see. If a user has downloaded gemini_python X1 package with the 'test_data', the user can move to this directory and run `typewalk` on that extensive set of Gemini datasets.

By default, `typewalk` will recurse all subdirectories under the current directory. Users may specify an explicit directory with the **-d** or **–dir** option.

`typewalk` provides the following options [**-h, –help**]:

```
-h, --help            show this help message and exit
-b BATCHNUM, --batch BATCHNUM
                      In shallow walk mode, number of files to process at a
                      time in the current directory. Controls behavior in
                      large data directories. Default = 100.
--calibrations        Show local calibrations (NOT IMPLEMENTED).
-c, --color           Colorize display
-d TWDIR, --dir TWDIR
                      Walk this directory and report types. default is cwd.
-f FILEMASK, --filemask FILEMASK
                      Show files matching regex <FILEMASK>. Default is all
                      .fits and .FITS files.
-i, --info            Show file meta information.
--keys KEY [KEY ...]  Print keyword values for reported files.Eg., --keys
                      TELESCOP OBJECT
-n, --norecurse       Do not recurse subdirectories.
--or                  Use OR logic on 'types' criteria. If not specified,
                      matching logic is AND (See --types). Eg., --or --types
                      GEMINI_SOUTH GMOS_IMAGE will report datasets that are
                      either GEMINI_SOUTH *OR* GMOS_IMAGE.
-o OUTFILE, --out OUTFILE
                      Write reported files to this file. Effective only with
                      --types option.
--raise               Raise descriptor exceptions.
--types TYPES [TYPES ...]
                      Find datasets that match only these type criteria.
                      Eg., --types GEMINI_SOUTH GMOS_IMAGE will report
                      datasets that are both GEMINI_SOUTH *and* GMOS_IMAGE.
--status              Report data processing status only.
--typology            Report data typologies only.
--xtypes XTYPES [[XTYPES ...]
                      Exclude <xtypes> from reporting.
```

Files are selected and reported through a regular expression mask which, by default, finds all ".fits" and ".FITS" files. Users can change this mask with the **-f, –filemask** option.

As the **–types** option indicates, `typewalk` can find and report data that match specific type criteria. For example, a user might want to find all GMOS image flats under a certain directory. `typewalk` will locate and report all datasets that would match the AstroDataType, GMOS_IMAGE_FLAT.

A user may request that a file be written containing all datasets matching AstroDataType qualifiers passed by the **–types** option. An output file is specified through the **-o, –out** option. Output files are formatted so they may be passed *directly to the reduce command line* via that applications 'at-file' (@file) facility. See *The @file facility* or the reduce help for more on 'at-files'.

Users may select type matching logic with the **–or** switch. By default, qualifying logic is AND, i.e. the logic specifies that *all* types must be present (x AND y); **–or** specifies that ANY types, enumerated with **–types**, may be present (x OR y). **–or** is only effective when the **–types** option is specified with more than one type.

For example, find all GMOS images from Cerro Pachon in the top level directory and write out the matching files, then run reduce on them (**-n** is 'norecurse'):

```
$ typewalk -n --types GEMINI_SOUTH GMOS_IMAGE --out gmos_images_south
$ reduce @gmos_images_south
```

Find all F2_SPECT and GMOS_SPECT datasets in a directory tree:

```
$ typewalk --or --types GMOS_SPECT F2_SPECT
```

This will also report match results to stdout, colourized if requested (**-c**).

Users may find the **–xtypes** flag useful, as it provides a facility for filtering results further by allowing certain types to be excluded from the report.

For example, find GMOS_IMAGE types, but exclude ACQUISITION images from reporting:

```
$ typewalk --types GMOS_IMAGE --xtypes ACQUISITION

directory: ../test_data/output
   S20131010S0105.fits ............... (GEMINI) (GEMINI_SOUTH) (GMOS) (GMOS_IMAGE)
   (GMOS_RAW) (GMOS_S) (IMAGE) (RAW) (SIDEREAL) (UNPREPARED)

   S20131010S0105_forFringe.fits ..... (GEMINI) (GEMINI_SOUTH) (GMOS) (GMOS_IMAGE)
   (GMOS_S) (IMAGE) (NEEDSFLUXCAL) (OVERSCAN_SUBTRACTED) (OVERSCAN_TRIMMED)
   (PREPARED) (SIDEREAL)

   S20131010S0105_forStack.fits ...... (GEMINI) (GEMINI_SOUTH) (GMOS) (GMOS_IMAGE)
   (GMOS_S) (IMAGE) (NEEDSFLUXCAL) (OVERSCAN_SUBTRACTED) (OVERSCAN_TRIMMED)
   (PREPARED) (SIDEREAL)
```

Exclude ACQUISITION images that have already had some processing done:

```
$ typewalk --types GMOS_IMAGE --xtypes ACQUISITION PREPARED

directory: ../test_data/output
   S20131010S0105.fits ............... (GEMINI) (GEMINI_SOUTH) (GMOS) (GMOS_IMAGE)
   (GMOS_RAW) (GMOS_S) (IMAGE) (RAW) (SIDEREAL) (UNPREPARED)
```

With **–types** and **–xtypes**, users may really tune their searches for very specific datasets.

CHAPTER

FIVE

# DISCUSSION

## 5.1 Fits Storage

The URLs that appear in `test_one` recipe example (Sec. *Test the installation*), reference web services available within the Gemini Observatory's operational environment. They will *not* be available directly to users running `reduce` outside of the Gemini Observatory environment.

In the context of `reduce` and the Astrodata Recipe System, FitsStorage provides a calibration management and association feature. Essentially, given a science frame (or any frame that requires calibration) and a calibration type requested, FitsStorage is able to automatically choose the best available calibration of the required type to apply to the science frame. The Recipe System uses a machine-oriented calibration manager interface in order to select calibration frames to apply as part of pipeline processing.

Though this service is not currently available to general gemini_python users, plans to provide this as a local calibration service are in place and expected for *Future Enhancements*.

## 5.2 The adcc

As a matter of operations, `reduce` and the Recipe System depend upon the services of what is called the `adcc`, the Automated Data Communication Center. The `adcc` provides services to pipeline operations through two proxy servers, an XML-RPC server and an HTTP server. The XML_RPC server serves calibration requests made on it, and retrieves calibrations that satisfiy those requests from the Gemini FITS Store, a service that provides automated calibration lookup and retrieval.

The `adcc` can be run externally and will run continuously until it is shutdown. Any instances of `reduce` (and the Recipe System) will employ this external instance of the `adcc` to service a pipeline's calibration requests. However, a user of `reduce` need not start an instance of the `adcc` nor, indeed, know anytihng about the `adcc` *per se*. If one is not available, an instance of the `adcc` will be started by `reduce` itself, and will serve that particular `reduce` process and then terminate.

This note is provided should users notice an `adcc` process and wonder what it is.

## 5.3 Future Enhancements

### 5.3.1 Intelligence

One enhancement long imagined is what has been generally termed 'intelligence'. That is, an ability for either `reduce` or some utility to automatically do AstroDataType classification of a set of data, group them appropriately, and then pass these grouped data to the Recipe System.

As things stand now, it is up to the user to pass commonly typed data to `reduce`. As shown in the previous section, *typewalk*, `typewalk` can help a user perform this task and create a 'ready-to-run' @file that can be passed directly to `reduce`. Properly implemented 'intelligence' will *not* require the user to determine the AstroDataTypes of datasets.

### 5.3.2 Local Calibration Service

The Fits Storage service will be delivered as part of a future release and will provide the calibration management and association features of *Fits Storage*: for use with the public release of the *gemini_python* data reduction package. This feature will provide automatic calibration selection for both pipeline (recipe) operations and in an interactive processing environment.

CHAPTER

SIX

# 6. ACKNOWLEDGMENTS

The Gemini Observatory is operated by the Association of Universities for Research in Astronomy (AURA), Inc., under a cooperative agreement with the NSF on behalf of the Gemini partnership: the National Science Foundation (United States), the Science and Technology Facilities Council (United Kingdom), the National Research Council (Canada), CONICYT (Chile), the Australian Research Council (Australia), Ministerio da Ciencia e Tecnologia (Brazil), and Ministerio de Ciencia, Tecnologia e Innovacion Productiva (Argentina).

APPENDIX

A

# *REDUCE* DEMO

Original demo author: Kathleen Labrie, October 2014

## A.1 Setting up

First install Ureka, which can be obtained at http://ssb.stsci.edu/ureka/.

The second step is to install `gemini_python` as described in *Section 2 - Installation*. Please do make sure that the command *reduce* is in your `PATH` and that `PYTHONPATH` includes the location where the modules `astrodata`, `astrodata_FITS`, `astrodata_Gemini`, and `gempy` are installed.

The demo data is distributed separately. You can find the demo data package `gemini_python_datapkg-X1.tar.gz` on the Gemini website where you found the gemini_python package. Unpack the data package somewhere convenient:

```
tar xvzf gemini_python_datapkg-X1.tar.gz
```

In there, you will find a subdirectory named `data_for_reduce_demo`. Those are the data we will use here. You will also find an empty directory called `playground`. This is your playground. The instructions in this demo assume that you are running the `reduce` command from that directory. There is no requirements to run `reduce` from that directory, but if you want to follow the demo to the letter, this is where you should be for all the paths to work.

## A.2 Introduction to the Demo

In this demo, we will reduce a simple dither on source GMOS imaging sequence. We will first process the raw biases, and then the raw twilight flats. We will then use those processed files to process and stack the science observation.

Instead of the default Quality Assessment (QA) recipe that is used at the Gemini summits, we will use another recipe that will focus on the reduction rather than on the multiple measurements of the QA metrics used at night. QA metrics, here the image quality (IQ), will only be measured at the end of the reduction rather than throughout the reduction. Another difference between the standard QA recipe and the demo recipe, is that the demo recipe does stack the data, while the stacking is turned off in the QA context.

The demo recipe is essentially a Quick Look recipe. It is NOT valid for Science Quality. Remember that we you are using is a QA pipeline, not a Science pipeline.

## A.3 The Recipes

To process the biases and the flats we will be using the standard recipes. The system will be able to pick those automatically when it recognizes the input data as GMOS biases and GMOS twilight flats.

For the science data, we will override the recipe selection to use the Demo recipe. If we were not to override the recipe selection, the system would automatically select the QA recipe. The Demo recipe is more representative of a standard Quick-Look reduction with stacking, hence probably more interesting to the reader.

The standard recipe to process GMOS biases is named `recipe.makeProcessedBias` and contains these instructions:

```
# This recipe performs the standardization and corrections needed to convert
# the raw input bias images into a single stacked bias image. This output
# processed bias is stored on disk using storeProcessedBias and has a name
# equal to the name of the first input bias image with "_bias.fits" appended.

prepare
addDQ
addVAR(read_noise=True)
overscanCorrect
addToList(purpose="forStack")
getList(purpose="forStack")
stackFrames
storeProcessedBias
```

The standard recipe to process GMOS twilight flats is named `recipe.makeProcessedFlat.GMOS_IMAGE` and contains these instructions:

```
# This recipe performs the standardization and corrections needed to convert
# the raw input flat images into a single stacked and normalized flat image.
# This output processed flat is stored on disk using storeProcessedFlat and
# has a name equal to the name of the first input flat image with "_flat.fits"
# appended.

prepare
addDQ
addVAR(read_noise=True)
display
overscanCorrect
biasCorrect
ADUToElectrons
addVAR(poisson_noise=True)
addToList(purpose="forStack")
getList(purpose="forStack")
stackFlats
normalizeFlat
storeProcessedFlat
```

The Demo recipe is named `recipe.reduceDemo` and contains these instructions:

```
# recipe.reduceDemo

prepare
addDQ
addVAR(read_noise=True)
overscanCorrect
biasCorrect
ADUToElectrons
```

```
addVAR(poisson_noise=True)
flatCorrect
makeFringe
fringeCorrect
mosaicDetectors
detectSources
addToList(purpose=forStack)
getList(purpose=forStack)
alignAndStack
detectSources
measureIQ
```

For the curious, the standard bias and flat recipes are found in `astrodata_Gemini/RECIPES_Gemini/` and the demo recipe is in `astrodata_Gemini/RECIPES_Gemini/demos/`. You do not really need that information as the system will find them on its own.

## A.4 The Demo

The images will be displayed at times. Therefore, start ds9:

```
ds9 &
```

### A.4.1 The Processed Bias

The first step is to create the processed bias. We are using the standard recipe. The system will recognize the inputs as GMOS biases and call the appropriate recipe automatically.

The biases were taken on different dates around the time of the science observations. For convenience, we will use a file with the list of datasets as input instead of listing all the input datasets individually. We will use a tool named `typewalk` to painlessly create the list.

```
cd <your_path>/gemini_python_datapkg-X1/playground

typewalk --types GMOS_BIAS --dir ../data_for_reduce_demo -o bias.list

reduce @bias.list
```

This creates the processed bias, `N20120202S0955_bias.fits`. The output suffix `_bias` is the indicator that this is a processed bias. All processed calibrations are also stored in `./calibrations/storedcals/` for safe keeping.

If you wish to see what the processed bias looks like:

```
reduce N20120202S0955_bias.fits -r display
```

*Note: This will issue an error about the file already existing. Ignore it. The explanation of what is going on is beyond the scope of this demo. We will fix this, eventually. Remember that this is a release of software meant for internal use; there are still plenty of issues to be resolved.*

### A.4.2 The Processed Flat

Next we create a processed flat. We will use the processed bias we have just created. The system will recognize the inputs as GMOS twilight flats and call the appropriate recipe automatically.

The "public" RecipeSystem does not yet have a Local Calibration Server. Therefore, we will need to specify the processed bias we want to use on the *reduce* command line. For information only, internally the QA pipeline at the summit uses a central calibration server and the most appropriate processed calibrations available are selected and retrieve automatically. We hope to be able to offer a "local", end-user version of this system in the future. For now, calibrations must be specified on the command line.

For the flats, we do not really need a list, we can use wild cards:

```
reduce ../data_for_reduce_demo/N20120123*.fits \
   --override_cal processed_bias:N20120202S0955_bias.fits \
   -p clobber=True
```

This creates the processed flat, `N20120123S0123_flat.fits`. The output suffix `_flat` is the indictor that this is a processed flat. The processed flat is also stored in `./calibrations/storedcals/` for safe keeping.

The `clobber` parameter is set to True to allow the system to overwrite the final output. By default, the system refuses to overwrite an output file.

If you wish to see what the processed flat looks like:

```
reduce N20120123S0123_flat.fits -r display
```

### A.4.3 The Science Frames

We now have all the pieces required to reduce the science frames. This time, instead of using the standard QA recipe, we will use the Demo recipe. Again, we will specify the processed calibrations, bias and flat, we wish to use.

```
reduce ../data_for_reduce_demo/N20120203S028?.fits \
   --override_cal processed_bias:N20120202S0955_bias.fits \
                  processed_flat:N20120123S0123_flat.fits \
   -r reduceDemo \
   -p clobber=True
```

The demo data was obtained with the z' filter, therefore the images contain fringing. The `makeFringe` and `fringeCorrect` primitives are filter-aware, they will do something only when the data is from a filter that produces fringing, like the z' filter. The processed fringe that is created is stored with the other processed calibrations in `./calibrations/storedcals/` and it is named `N20120203S0281_fringe.fits`. The `_fringe` suffix indicates a processed fringe.

The last primitive in the recipe is `measureIQ` which is one of the QA metrics primitives used at night by the QA pipeline. The primitive selects stars in the field and measures the average seeing and ellipticity. The image it runs on is displayed and the selected stars are circled for visual inspections.

The fully processed stacked science image is `N20120203S0281_iqMeasured.fits`. By default, the suffix of the final image is set by the last primitive run on the data, in this case `measureIQ`.

This default naming can be confusing. If you wish to set the suffix of the final image yourself, use `--suffix _myfinalsuffix`.

### A.4.4 Clean up

It is good practice to reset the RecipeSystem state when you are done:

```
superclean --safe
```

Your files will stay there, only some hidden RecipeSystem directories and files will be deleted.

## A.5 Limitations

The X1 version of the RecipeSystem has not been vetted for Science Quality. Use ONLY for quick look purposes.

The RecipeSystem currently does not handle memory usage in a very smart way. The number of files one can pass on to `reduce` is directly limited by the memory of the user's computer. This demo ran successfully on a Mac laptop with 4 GB of memory.