Tutorial T2:
Data Inspection and Editing

Andy Biggs
ALMA Regional Centre (ESO)

Preliminaries

- Move to a directory where you intend to work
  - Should have plenty of space
  - Dataset we'll be working on is 1.6 GB in size
- Download data, script and final list of flags
  - You've already done this, right...?
- Tutorial script contains skeletal CASA commands
  - Tips and hints as well
  - Goal is to complete script

Start CASA

- On the command line type:
  - **`casapy`**
- After a short while, this brings up:
  - An (i)Python prompt
    - The command-line interpreter is Python
    - All Python functionality is available in CASA
    - e.g. **`range`** command, numpy, loops, etc.
  - A logger window
    - Useful information is reported here by each task

Locate the data

- List directory from within CASA
  - Can use usual unix commands i.e. '**ls**'
  - Often requires a '**!**' e.g. '**!display plotants.png**'
- The dataset we will use is **`all_avg.ms/`**
  - This is a directory!
    - **`.ms`** = Measurement Set (MS)
    - Contains tables and visibility data
    - Rarely have to worry about contents of MS
- CASA works directly on this MS
  - One doesn't load the data into CASA

Running CASA tasks (1)

- Set inputs individually
  - `inp taskname`
  - `parameter1=x`
  - `parameter2=y`
  - `inp` (to review)
  - `go`
- Parameter names can be tab-completed
  - Avoids typos and creation of useless variables
- `tget taskname`
  - Returns previous inputs (stored in `.last` file)

Running CASA tasks (2)

- Use a “one-liner”
  - `taskname(parameter1=x,parameter2=y)` or…
  - `taskname(x,y)` (if x and y are the first two inputs)
- Assemble one-liners in a python script
  - `execfile('andys_script.py')`
  - “Best practice”

View summary of the data

- Relevant task is **`listobs`**
- Summary information includes:
  - Which sources are included in the file
  - Spectral window properties
    - Bandwidth, central frequency, number of channels, correlations
  - Which antennas were in the array
  - Observation sequence
    - Which source with which spectral windows when?
- Spectral windows
  - Spectral windows are selected via a (0-based) index
  - Non-science spws often included e.g. ALMA

View antenna locations

- Relevant task is **`plotants`**
- Shows 2-D map of antenna positions
- Interesting information includes:
  - Size of the array
  - Antenna distribution
    - Are there any outliers?
- Often used to select reference antenna
  - Choose one with a large range of baseline lengths
  - For e-MERLIN, ‘Mk2’ is usually used

Viewing (*u,v*) coverage

- Relevant task is **`plotuv`**
- Shows projected baseline for each visibility
  - *w* coordinate is not shown (obviously)
- Complete coverage (no gaps) would be nice...
  - Incomplete coverage produces image sidelobes
  - The bigger the gaps, the bigger the sidelobes
- Wide-bandwidths help a lot
  - Separate track for each frequency point
  - $(u, v) = B / \lambda$

Viewing data

- Relevant task is **`plotms`**
  - 2-D data plotter e.g. amplitude against frequency
- **`plotms`** is a very powerful task
  - Data can be selected on almost every property
    - Antenna, baseline, timerange, spw, correlation, ...
  - Data can be displayed in many different ways
    - >50 axis possibilities (time, frequency, phase, amplitude, ...)
- Very useful for:
  - Visualising data
  - Finding bad data

Flag versions

- Relevant task is **`flagmanager`**
- Two main modes of operation:
  - Save current list of flags to a named version
    - Do this regularly!
  - Replace the current flags with those from a named version
    - In the unlikely event that you screw up!
- Versions are stored in a **`.flagversion`** directory

Flagging data

- Some data normally needs to be deleted ("flagged")
- Typical reasons include:
  - An antenna is not working very well or at all
    - Problem may be time-variable
    - May only affect some spectral windows
  - Data taken whilst telescopes are off-source (slewing)
    - Many telescopes now flag this automatically
  - Edge channels of the spectral windows
    - Amplitude of these is usually lower
- Problems are usually antenna-based!
- Usually affect all correlations