CASA

B.E. Glendenning

What is CASA?

- CASA is the baseline post-processing package for EVLA and ALMA data
- It is a suite of applications for the reduction and analysis of radio-astronomical data (derived from the former AIPS++ package)
- The algorithms are written in C++; interface in python/ipython
- Plotting is done with the matplotlib library and Qt
- The Viewer and tablebrowser are Qt-based
- It is fully scriptable, with in-line help and scientist-written documentation (notably the cookbook)
- Telescope data (visibility and single-dish) are stored in a MeasurementSet (MS); filler converts EVLA SDM+BDF data to the MS
- It contains functionality for manipulating/plotting/… core infrastructure data types (e.g., Tables, Measures, …)
- Interferometric calibration and imaging are done via the Hamaker, Bregman, Sault formalism (Measurement Equation)
- It contains image analysis and other mathematical functionality

General Status

- Have had Beta (patch) releases every ~3 months since October 2007
  - Initially restricted, now available (after registration) to anyone
  - Tutorial at synthesis imaging summer school, ~50 students (positive feedback)
- Used every day for EVLA correlator data translation & at the ALMA Test Facility (ATF)
- Generally very capable, although too much expertise is sometimes required
- The task paradigm (a'la AIPS) is good, but the user interface needs attention in some areas

Current Capabilities

- Data Import
  - VLA (EVLA) archive
  - External EVLA and ALMA fillers complete
  - UVFITS (also for export)
- Flagging
  - UV-plot based including time, channel averaging
  - Viewer flagging
  - Manual flagging
- Calibration
  - Polarization
  - VLA flux density calibrator images
  - Spline fitting and polynomial bandpass determination
  - Flexible combination of multiple spectral windows
- Imaging
  - Mosaicing (various types)
  - Widefield imaging (W-projection much faster than faceting)
  - Multi-scale clean (experimental)
  - Analysis includes image math, statistics, image plane fitting

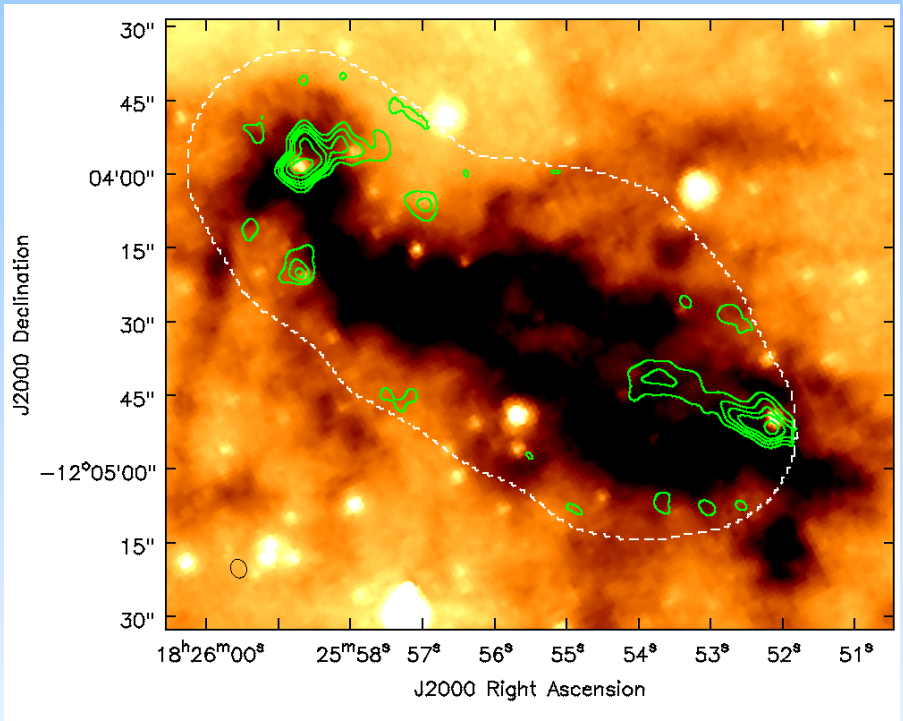
Needed Developments

- User Interface (GUIs)
- Demonstrate satisfactory speed with large datasets
- Post-observation corrections (antenna location errors, for example)
- Plotting speed and flexibility
- Viewer flagging and autoflagging improvements
- More sophisticated image analysis tasks
- Continuum subtraction with many spectral windows
- Improvements to calibration solution visualization
- Logger information clean-up and streamlining
- Algorithms developed by the algorithm development group need to get implemented within CASA (see B. Butler talk)

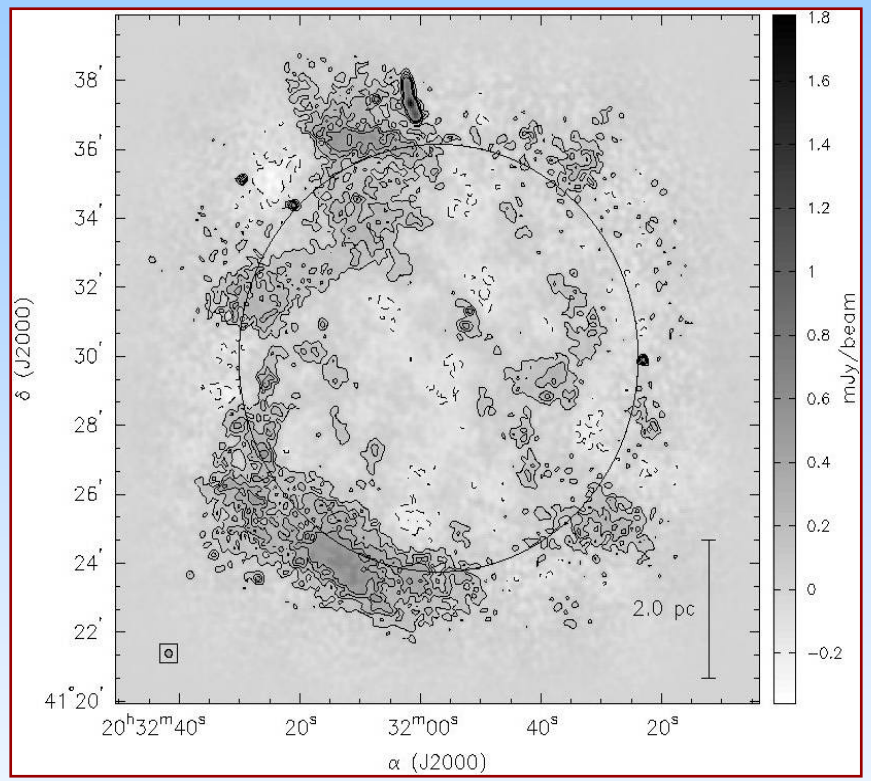
CASA Strengths for SRO

- Full data import (e.g., complex correlator setups)
- Able to handle large datasets
- Wide-band imaging using Multiscale MFS
- W-projection imaging
- Non-linearized polarization calibration (for high dynamic range), frequency-dependent D terms
- Spline G (gain) solutions
- Low-level data inspection/modification tools; scriptability in general

Production of Scientific Images in CASA Becoming Routine

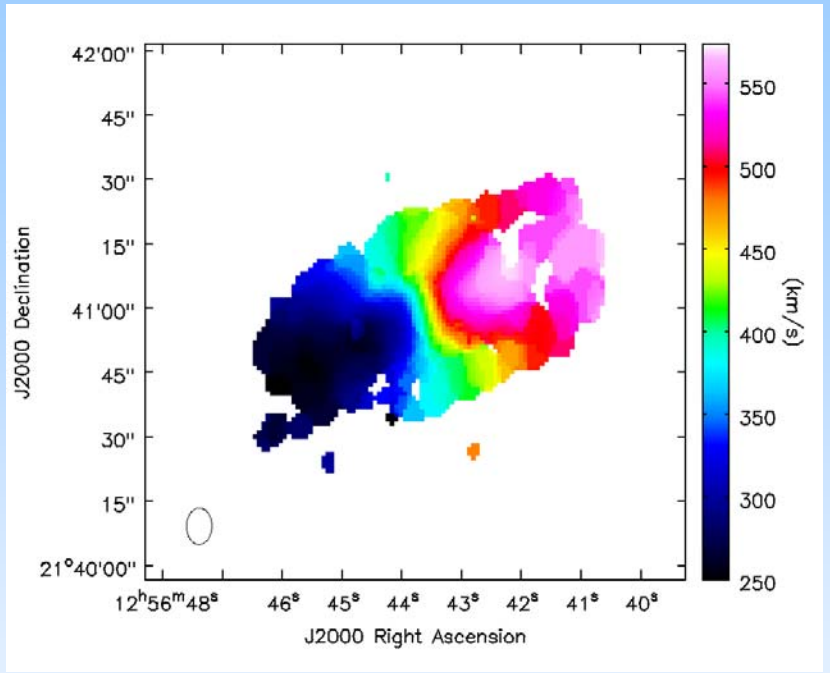


Green contours show SMA ^{12}CO (2-1) integrated intensity superposed on a GLIMPSE 8 μm image of the Infrared Dark Cloud (IRDC) G19.3+0.07. Six-pointing SMA mosaic imaged in CASA – calibration of SMA data coming soon. Brogan et al. (in prep).

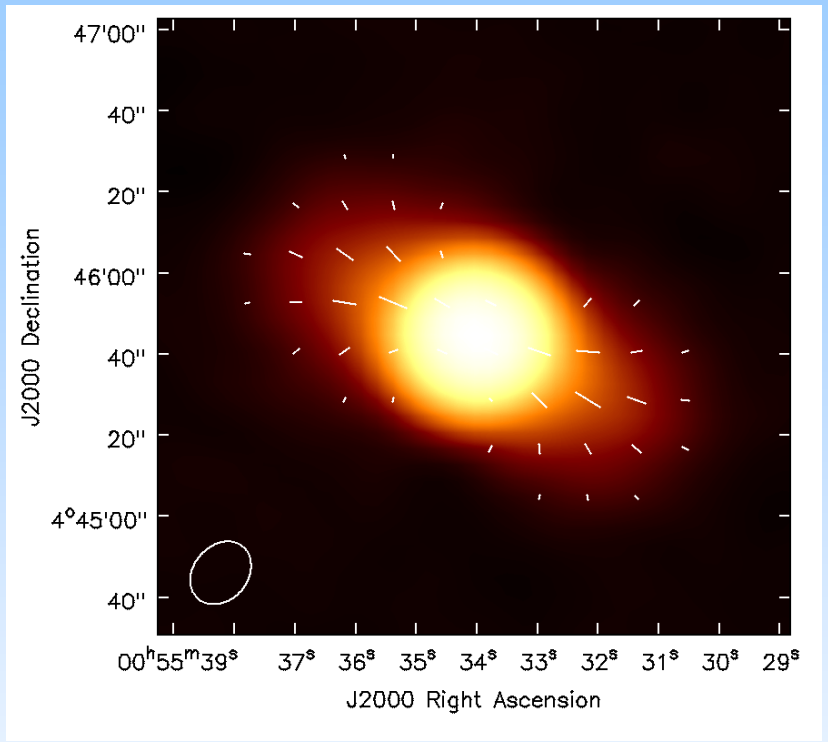


An extended radio counterpart of TeV J2032+4130 in the Cygnus OB association. VLA 3.6 cm continuum 5 point mosaic, D configuration, multi-scale clean. Butt et al. (2008)

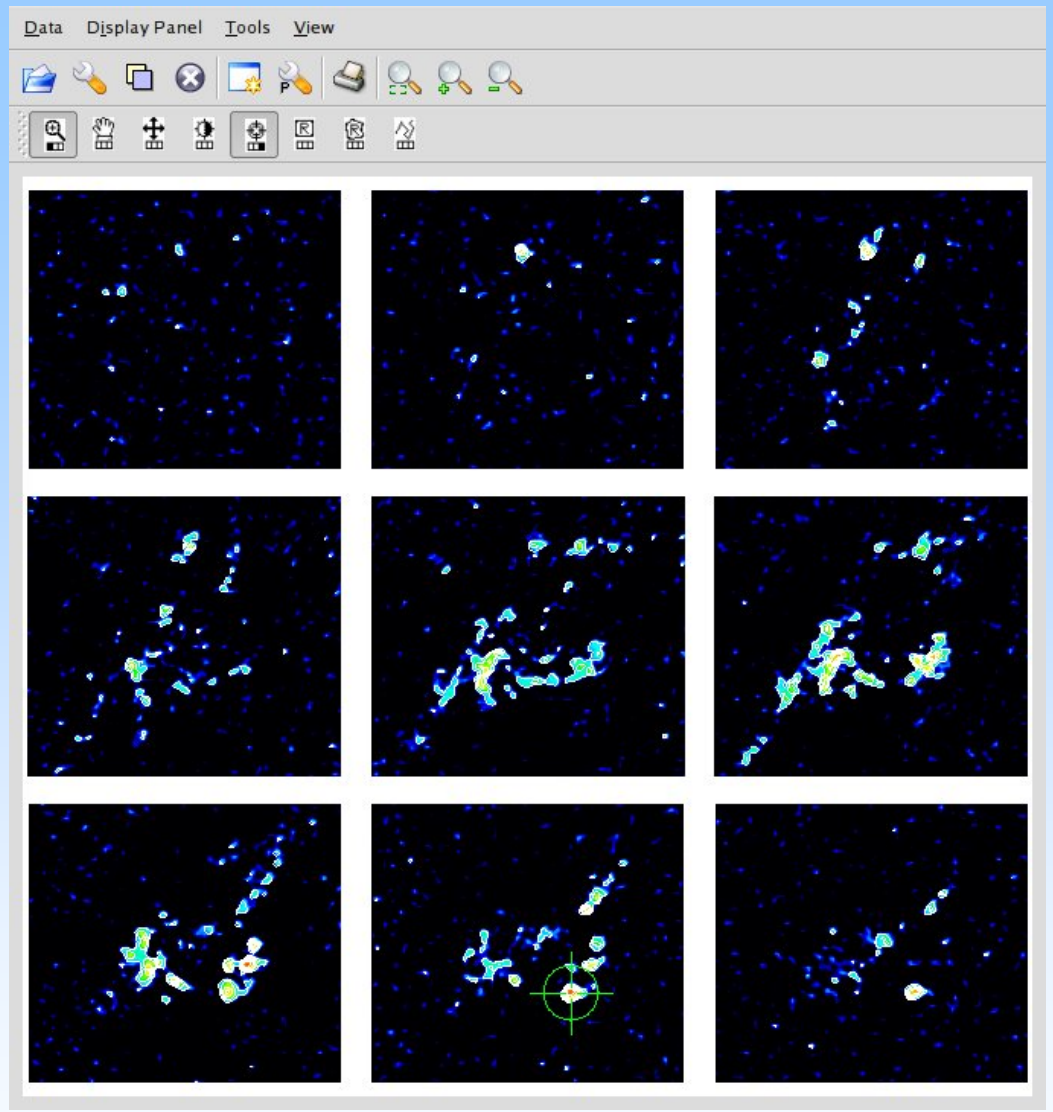
Data Calibrated and Imaged in CASA Tutorials at NRAO Synthesis Imaging Workshop June 2008

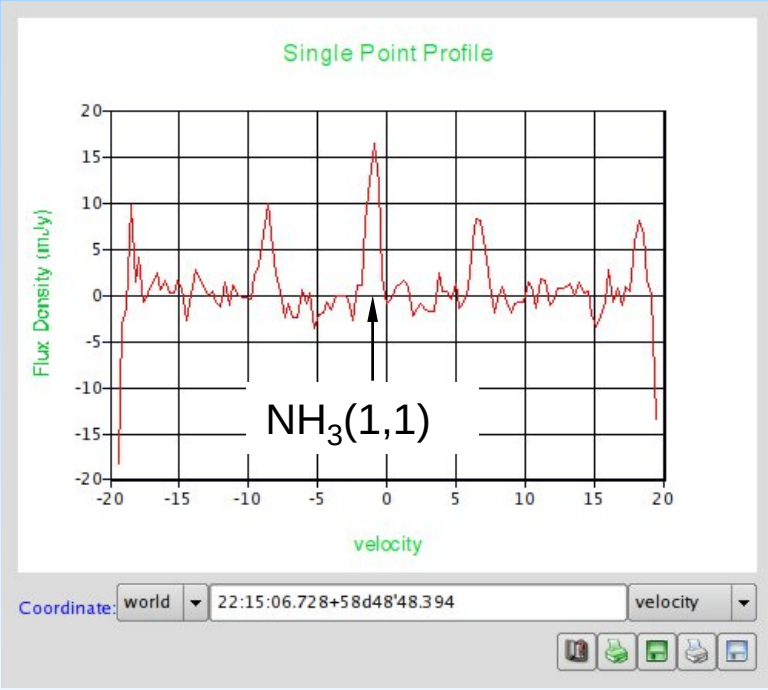


CO(1—0) kinematics (moment 1) of the galaxy NGC4826 from the BIMA SONG survey (data originally published in Helfer, Thornley, Regan et al. 2003)



E-field vectors in Jupiter magnetosphere. Archival VLA 6 cm D-configuration full Stokes polarization data.





IRAS 22134 - a young ring cluster. VLA mosaic of $NH_3(1,1)$ and associated spectrum at one point in the ring. Main and hyperfine components are visible in the spectrum. Shepherd & Kumar 2008, ApJ, in prep.

Backup Slides

Requirements Areas

- General
  - General and Relation to Pipeline
  - Operational Issues
  - Performance
- User Interface
  - General User Interface
  - Graphical User Interface (GUI)
  - Command Line Interface (CLI)
  - Interface programming, parameter passing & feedback
  - Documentation & Help Facility
- Data Handling
  - Data Handling - General Data Requirements
  - Data Import and Export
  - Images and Other Data Products
  - Foreign Data
  - Interaction with the Archive
- Calibration and Editing
  - General Calibration and Editing
  - Atmospheric Calibration
  - Interferometer Data Calibration
  - Single Dish Data Calibration (ALMA)
  - Mosaicing Considerations for Calibration
  - Ancillary & Diagnostic Data considerations for calibration
- Imaging
  - General Imaging
  - Interferometric Imaging
  - Mosaicing and Single Dish (ALMA) Imaging

Requirements Areas (2)

- Analysis
  - General Analysis
  - Spectral Line Analysis
  - Image Cube Analysis & Manipulation
  - Single Dish Specific Analysis (ALMA)
- Visualization and Plotting
  - General Visualization and Plotting
  - Display Appearance and Interactivity
  - Visibility Data Visualization
  - Display Other Data
  - Image Cube Manipulation
  - Single Dish Plotting
- Special Features
  - Simulation
  - VLBI *(not being actively developed)*
  - Solar System Objects
  - Pulsars *(not being actively developed)*

Imaging & Deconvolution

- Mosaic imaging
  - Joint deconvolution (Miriad style) and by *gridding convolution*
  - Mosaicing with heterogenous arrays (ALMA, CARMA)
- Widefield imaging: *W-projection* and faceting
  - W-projection more than 1 order of magnitude faster than faceting
- Multiple algorithms for single dish and interferometry combination
  - Feathering
  - Single Dish as a model for deconvolution
  - *True joint deconvolution using both visibility data and raster single dish software*
    - Requires data with well-calibrated weights between the single dish and interferometry data (ALMA), and testing
- *Full beam polarimetric imaging*
  - Targeted at friendly VLA users on a “shared risk” basis
- *Multiscale clean*
- MEM & NNLS (toolkit level only so-far)

Calibration

- Standard gain & bandpass calibration
  - Sampled and Polynominal/Spline solutions available
  - Flux density reference scaling
  - Sampled baseline-based solution available
  - Solution normalization
  - Phase-only, *Amp-only options*
  - Auto-interpolation of flagged channels in bandpass
- Polarization calibration
  - Linearized instrumental polarization (D-terms) solutions available
  - *Channelized option for frequency-dependent instrumental polarization*
  - Optional solution for source polarization
  - Polarization position-angle solution support (for circular basis)

Calibration (2)

- Additional features
  - *Flexible combinations of data (over scan,field,spw) for solving ("fan in")*
  - *Flexible distribution of solutions to data ("fan out")*
  - Smoothing
  - Interpolation and accumulation (incremental)
  - Solution plotting, including interactive flagging
- TBD
  - Cal table alignment with Science Data Model
  - gain spline improvements (fan-in/out, better phase-tracking, etc.)
  - additional flexibility in data flagging by calibration application
  - additional smoothing and interpolation modes
  - extend full polarization calibration support to linear basis (gain/source polarization disentanglement; position angle solve)
  - parameterized (e.g., poly and/or spline) instrumental polarization option

Performance

- For “small” data CASA is comparable but slower than other packages. A previous benchmarking campaign showed:
  - Test Case 1 - Polarized continuum: AIPS++/AIPS=1.1, AIPS++/Miriad=2.4
  - Test Case 2 – 1 & 3 mm spectral line: AIPS++/Gildas = 1.1
  - Test case 3 – 7mm, fast switching spectral line: AIPS++/AIPS = 1.5
- Recent “intermediate” tests show that CASA is much faster than AIPS (~6x)
  - No magic: CASA tiling prevents reading through the data many times
  - AIPS is faster when memory is larger than the raw data size
- Started: Terabyte initiative
  - Flag, calibrate, image 1 TB (raw data size) data = 10h of peak data
  - Cluster (16 nodes, 128 cores) purchased, working on simulating the data and initial timing tests
  - Should have results Q2 2009.
- *Testing EVLA sized data sets is the important exercise!*