

Analysing ALMA data - the CASA software package Dirk Petry (ESO), June 2010

Outline

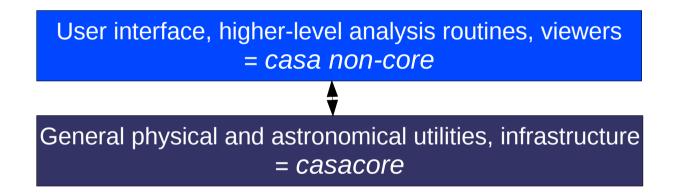
- \rightarrow What is CASA?
- \rightarrow Who develops CASA?
- → What are the main requirements and how does CASA meet them?
- \rightarrow How does CASA look and feel?
- $\rightarrow\,$ CASA status and release plans

- main features
- development team
- design and implementation
- the typical analysis session



CASA main features

- CASA = Common Astronomy Software Applications
- Development started in the 90s as the next generation of AIPS
- Refocussed in 2003 to be the ALMA/EVLA analysis package
- Has the intention to be a general software package to reduce both interferometer and single-dish data
- Internally consists of two parts:



- Implements the "Measurement Equation" (Hamaker, Bregman & Sault 1996)
- Internal data format is the "Measurement Set" (Kemball & Wieringa 2000)
- 1.5 Million lines of code (mostly C++)
- In public release under GNU Public License since December 2009



CASA – development team



D. Petry, MWA School, Amsterdam, June/July 2010



CASA – development team



Since mid 2008, two CASA developers at ESO, since Sept. 2009 three



CASA – development team

Originally only developed at NRAO (Socorro, NM), now

approx. 17 FTE developers are at work at

US (NRAO and others): 10.5 Japan (NAOJ): 3.0 Europe (ESO and others): 3.5

+ 1 CASA manager (NRAO Socorro) = Nick Elias

+ 1 Project Scientist (NRAO Socorro) = Jürgen Ott

+ a few 5% FTEs at ASTRON, ATNF, and other places Also involved: ALMA Computing Managers = B. Glendenning (NRAO), G. Raffi, P. Ballester (ESO)



Overall architecture:

1) A data structure

- 2) A set of data import/export facilities
- 3) A set of tools for data access, display, and editing
- 4) A set of tools for science analysis
- 5) A set of high-level analysis procedures ("tasks")
- 6) A programmable command line interface with scripting
- 7) Documentation



Overall architecture:

1) A data structure

Tables: Images, Caltables, and the Measurement Set (MS)

2) A set of data import/export facilities

the so-called *fillers*: ASDM \rightarrow MS, FITS \rightarrow Image, UVFITS \rightarrow MS, VLA \rightarrow MS, etc.

3) A set of tools for data access, display, and editing

tools to load/write data into/from casacore data types,

Qt-based table browser, viewer, and (beta) x/y plotter, matplotlib-based x/y plotter

4) A set of tools for science analysis

built around the *Measurement Equation* (developed in 1996) = a set of C++ classes for radio astronomical calibration and imaging

- 5) A set of high-level analysis procedures ("tasks") special procedures for each required task such as CLEAN etc.
- 6) A programmable command line interface with scripting

Python (augmented by IPython) gives a MATLAB-like interactive language

7) Documentation

an extensive cookbook (500 pages) + documentation through help commands (help, ?, pdoc) + online help pages for users and developers



Overall architecture:

1) A data structure

- a) Tables: Images, Caltables, and the Measurement Set (MS)
 - 2) A set of data import/export facilities

the so-called *fillers*: ASDM \rightarrow MS, FITS \rightarrow Image, UVFITS \rightarrow MS, VLA \rightarrow MS, etc.

3) A set of tools for data access, display, and editing

tools to load/write data into/from casacore data types,

Qt-based table browser, viewer, and (beta) x/y plotter, matplotlib-based x/y plotter

- 4) A set of tools for science analysis
- b) built around the Measurement Equation (developed in 1996) = a set of C++ classes for radio astronomical calibration and imaging

5) A set of high-level analysis procedures ("tasks") special procedures for each required task such as CLEAN etc.

6) A programmable command line interface with scripting

- c) Python (augmented by IPython) gives a MATLAB-like interactive language
 - 7) Documentation

an extensive cookbook (500 pages) + documentation through help commands (help, ?, pdoc) + online help pages for users and developers



CASA special features:

a) the Measurement Set (MS)

- developed by Cornwell, Kemball, & Wieringa between 1996 and 2000
- designed to store both interferometry (multi-dish) and single-dish data
- supports (in principle) any setup of radio telescopes
- supports description and processing of the data via the Measurement Equation
- fundamental storage mechanism: CASA Tables (inspired by MIRIAD)
- MS = table for radio telescope data (visibilities) + auxiliary sub-tables



The Measurement Set

MAIN none		ANTENNA_ID FEED_ID DATA_DESC_ID	ANTENNA ANTENNA_ID		FI		ΞT	(ORBIT_ID) (PHASED_ARRAY_ID)	POLARIZATION POLARIZATION_ID row number DATA_DESCRIPTION none			
		PROCESSOR_ID (PHASE_ID) FIELD_ID (PULSAR_GATE_ID) ARRAY_ID				POINTING SYSCAL WEATHER			(SOURCE)	SOURCE_ID explicit	(DOPPLER) FIELD	SPW_ID (PULSAR_ID)
		OBSERVATION_ID STATE_ID	FEED FE	ED_ID	explicit	MAIN FREQ_OFFS SYSCAL	ET	ANTENNA_ID SPW_ID BEAM_ID				
			SYSCAL			SYSCAL	BEAM_ID (PHASED_FEED_ID)	(DOPPLER)	DOPPLER_ID explicit	SPW_ID	SOURCE_ID TRANSITION_ID	
(FREQ_OFFSET) none	ANTENNA_ID FEED_ID SPW_ID	DATA_DESC	RIPTION DA	TA_DESC_ID	row number	MAIN	SPW_ID POLARIZATION_ID (LAG_ID)				
(SYSCAL) none		ANTENNA_ID FEED_ID SPW_ID	PROCESSOF	R PROCESS	DR_ID	row number	MAIN	TYPE_ID MODE_ID (PASS_ID)				
POINTING none		ANTENNA_ID POINTING_MODEL_ID	FIELD	FIELD_ID		row number	MAIN	SOURCE_ID (EPHEMERIS_ID)				
			OBSERVATIO	ON OBSERVATI	ON_ID	row number	MAIN HISTORY	none				
(WEATHER) none		ANTENNA_ID	STATE	STATE_ID		row number	MAIN	none				
HISTORY none		OBSERVATION_ID OBJECT_ID	SPW	SPW_ID	row number	DATA_DESCI FEED FREQ_OFFSI SOURCE		(RECEIVER_ID) (DOPPLER_ID) (ASSOC_SPW_ID)				
FLAG_CMD none		none				SYSCAL						
	Legend:							1. Tables not r	referenced by ot	her tables		
	[Table Name] [Key defined in this table] [key definition method] [referenced by] [referenced keys] (optional) reference to table outside the MS definition						Leve		elerenced by ou	ler tables		
							Leve	2: Tables refer	enced by level 1			
							Level 3: Tables referenced by level 2					

V1, D.Petry, 13.2.09



CASA special features:

b) the *Measurement Equation* (Hamaker, Bregman, & Sault 1996 + Sault, Hamaker, & Bregman 1996) implemented as a set of C++ classes for radio astronomical calibration and imaging

$$\vec{V}_{ij} = \vec{M}_{ij}\vec{B}_{ij}\vec{G}_{ij}\vec{D}_{ij}\int \vec{E}_{ij}\vec{P}_{ij}\vec{T}_{ij}\vec{F}_{ij} S\vec{I}_{v}(l,m) e^{-i2\pi(u_{ij}l+v_{ij}m)} dl \, dm + \vec{A}_{ij}$$

where

the vectors are: V = visibility = f(u, v), I = Image to be calculated,

A = additive baseline-based error component

the matrices are: M = multiplicative, baseline-based error component

- *B* = bandpass response
- G = generalised electronic gain
- D = polarisation leakage
- E = antenna voltage pattern
- *P* = paralactic angle
- T = tropospheric effects
- F = ionospheric Faraday rotation
- S = mapping of I to the polarization basis of the observation

other variables and indices are:

l, *m* = image plane coordinates, *i*, *j* = telescope ID pairs = baseline, u, v = Fourier plane coordinates



CASA special features:

 b) the Measurement Equation (Hamaker, Bregman & Sault 1996) implemented as a set of C++ classes for radio astronomical calibration and imaging (continued)

Assuming, e.g., independence of the matrices from (l,m), the ME can be solved for individual calibration components.

$$\vec{V}_{ij}^{obs} = \vec{B}_{ij}\vec{G}_{ij}\vec{D}_{ij}\vec{P}_{ij}\vec{T}_{ij}\vec{F}_{ij} \vec{V}_{ij}^{ideal}$$

ideal visibility known from calibrator source

 \Rightarrow have set of linear equations.

The actual calculation of the component is then a χ^2 minimization.

The calibrater (cb) tool contains a set of *solvers* for the different calibration components.



CASA special features:

c) A programmable command line interface with scripting

- Framework Architecture of 17 tools can be bound to any scripting language, presently selected is **Python (augmented by IPython)**
 - at atmosphere library
 - ms Measurement Set utilities
 - mp Measurement Set Plotting, e.g. data (amp/phase) versus other quantities
 - cb Calibration utilities
 - cp Calibration solution plotting utilities
 - im Imaging utilities
 - ia Image analysis utilities
 - fg flagging utilities
 - tb Table utilities (selection, extraction, etc.)
 - me Measures utilities
 - tp table plot
 - vp voltage patterns
 - qa Quanta utilities
 - cs Coordinate system utilities
 - pl matplotlib functionality
 - sd ASAP = ATNF Spectral Analysis Package (single-dish analysis imported from ATNF)
 - sm simulation



CASA special features:

c) A programmable command line interface with scripting

(continued)

Python (augmented by IPython)

Gives features such as

- tab completion
- autoparenthesis
- command line numbering
- access to OS, e.g.

```
Lines starting with '!' go to the OS.
```

```
a = !ls *.py to capture the output of 'ls *.py'.
```

```
!cmd $myvar expands Python var myvar for the shell.
```

- history
- execfile()
- comfortable help



CASA special features:

c) A programmable command line interface with scripting (continued)

In addition to toolkit: high-level tasks for the standard user

toolkit (implemented in C++) tasks (implemented in Python)

e.g. the task *importfits* is based on the tool *ia* (image analysis):

```
#Python script
casalog.origin('importfits')
ia.fromfits(imagename,fitsimage,whichrep,whichhdu,zeroblanks,overwrite)
ia.close()
```

CASA 3.0.1 comes with 91 implemented tasks.



CASA status

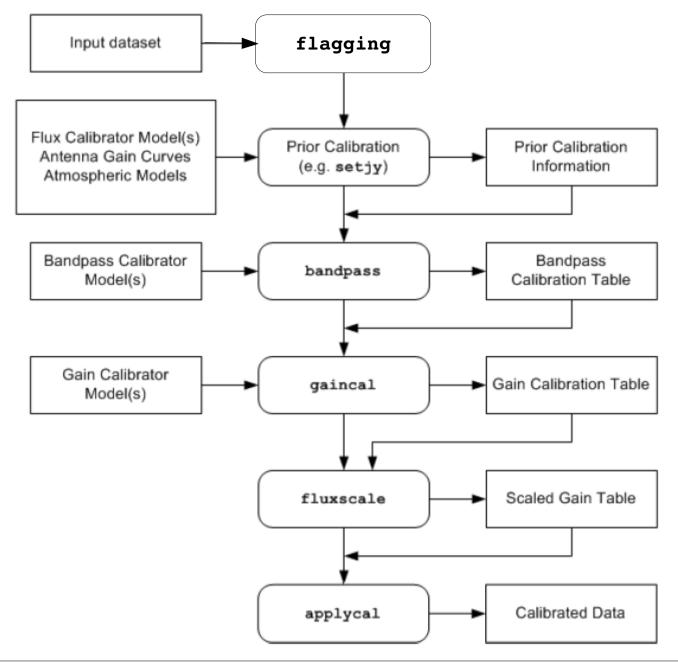
- Since Dec 2009 in public release under GPL = anybody can download, no warranty (see http://casa.nrao.edu), limited support (help desk, needs registration)
- Tutorials for the user community regularly given
- The first public release was CASA 3.0.0 (Dec 2009), release 3.0.2 published this month
- Development platforms: Linux (RHEL) + Mac OS X
- Supported platforms (binary distribution): RHEL, Fedora, openSuSE, Ubuntu, Max OS X
- Code kept in *svn* repository at NRAO, Socorro
- Presently have approx. 4300 modules, 1.5E6 lines of code, 1E6 lines of comments
- The core functionality (*casacore*, also available at http://code.google.com/p/casacore/) is also used by other projects
- Hot topics:
 - Support for High Performace Computing and Parallelisation
 - Advanced Imaging: wide fields, continuum imaging over wide spectral ranges
 - Interoperability: using CASA for other observatories and VLBI



How does CASA look and feel?

A typical analysis session

Part 1: flagging and calibration





How does CASA look and feel?

A typical analysis session Part 2: imaging and

image analysis

Calibrated Data applycal imaging image cube numerical viewing, analysis plotting

publication-ready plots and numerical results



Pictures from a typical analysis session

1) Startup:

open terminal and start casapy

Available tasks and tools are listed and the logger window is opened.

>	🖌 dpetry@M83:~/temp/ca	asa-bologna2010	
	CASA Version 3.0.1	ologna2010]\$ casapy (r11099) 2010/04/15 04:08:39 UTC	
	tasklist taskhelp help taskname toolhelp help par.parame	he following commands: - Task list organized by category - One line summary of available tasks - Full help for task - One line summary of available tools Hetername - Full help for parameter name * tasks are available after asap_init() is run	
	Activating auto-log Filename : ip Mode : ba Output logging : Fa Raw input log : Fa Timestamping : Fa State : ac CASA <2>: []	ackup alse alse alse	ved.



The logger provides functionality for monitoring and debugging command execution.

🗕 🖶 🖶 🗃 🗼 🗶	Search Message:	💏 Filter: Time 🔻
ne	Priority Origin	Message
2010-04-23 12:04:03	INFO plotms:::	###### Begin Task: plotms #####
2010-04-23 12:04:03	INFO	plotms::::casa
2010-04-23 12:04:04	INFO	plotms::::casa
2010-04-23 12:04:04	INFO plotms:::	###### End Task: plotms #####
2010-04-23 12:04:04	INFO plotms:::	###################################
2010-04-23 12:08:11	INFO	plotxy::::casa
2010-04-23 12:08:11	INFO plotxy:::	###################################
2010-04-23 12:08:11	INFO plotxy::	: ###### Begin Task: plotxy #####
2010-04-23 12:08:11	INFO	plotxy::::casa
2010-04-23 12:08:11	INFO plotxy:::	Switching to GUI mode. All current plots will be reset.
2010-04-23 12:08:11	INFO plotxy:::	Adding scratch columns, if necessary.
2010-04-23 12:08:11	INFO calibrate	e Opening MS: ah847_1-k-selected-flagged-calibd.ms for calibration.
2010-04-23 12:08:11	INFO Calibrate	e Initializing nominal selection to the whole MS.
2010-04-23 12:08:12	INFO	Data to be selected from matches the following:
2010-04-23 12:08:12	INFO +	Baselines: *ALL pairs of* VA01, VA02, VA03, VA04, VA05, VA06, VA07, VA08,
2010-04-23 12:08:12	INFO +	Fields: *ALL* 12190+47182, 12191+48299, 1331+305
2010-04-23 12:08:12	INFO +	Spectral Windows: *ALL*
2010-04-23 12:08:12	INFO +	SPW 0: *ALL Channels* 1 to 1 with a step of 1
2010-04-23 12:08:12	INFO +	SPW 1: *ALL Channels* 1 to 1 with a step of 1
2010-04-23 12:08:12	INFO +	Correlations:
2010-04-23 12:08:12	INFO +	Corr. ID 0 - RR, RL, LR, LL
2010-04-23 12:08:12	INFO +	Corr. ID 1 - *NONE*
2010-04-23 12:08:12	INFO +	Time Range *ALL* 2004/5/22/01:06:05 to 2004/5/22/03:32:25
2010-04-23 12:08:12	INFO +	Scan Numbers: *ALL* 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 1
2010-04-23 12:08:12	INFO +	UVRanges: *ALL*
2010-04-23 12:08:12	INFO	Preparing data
2010-04-23 12:08:21	INFO	Now get the data
2010-04-23 12:08:23	INFO	Done Processing data
2010-04-23 12:08:24	INFO	Now get the data
2010-04-23 12:08:25	INFO	Done Processing data
2010-04-23 12:08:26	INFO	plotxy::::casa
2010-04-23 12:08:26	INFO plotxy::	##### End Task: plotxy #####

D.



Pictures from a typical analysis session

2) enter commands in a MATLAB-like environment

recall previous settings

list present settings
 for given task
(includes parameter
 verification)

ession									
🕞 dpetry@pc014720:~/t	emp/radio-analysis/co	qtau	1+mwc480 - Shell - Konsole 📃 🗆 🗙						
Session Edit View Bo	Session Edit View Bookmarks Settings Help								
	_								
CASA <15>: fluxscale(vis='AT352_A071103	-K'	, caltable='AT352_A071103-K-gain', fluxtable						
='0', transfer='1')									
CASA <16>: applycal(v	is='AT352_A071103-	К',	gaintable='AT352_A071103-K-gain', field='2'						
CASA <17>: tget clean									
> tget(clean)								
Restored parameters f	rom file clean.las	t							
CASA <18>: inp									
> inp()									
# clean :: Deconvolv	e an image with se	lec	ted algorithm						
			name of input visibility file						
			Pre-name of output images						
field =	'2'		Field Name						
spw =		#	Spectral windows:channels: '' is all						
selectdata =	False	#	Other data selection parameters						
mode =	'mfs'		Type of selection (mfs, channel, velocity,						
niter =	500		Maximum number of iterations						
gain =	0.1	#	Loop gain for cleaning						
threshold =	'0.0mJy'	#	Flux level to stop cleaning. Must include						
psfmode =	'clark'	#							
imagermode =		#	Use csclean or mosaic. If '', use psfmode						
multiscale =	[]	#	set deconvolution scales (pixels), default:						
interactive =	True	#	use interactive clean (with GUI viewer)						
npercycle =	100	#	Number of iterations before interactive pro						
mask =	[]	#	cleanbox(es), mask image(s), and/or region(
	[512, 512]		x and y image size in pixels, symmetric for						
	['0.03arcsec', '0	.03	arcsec'] # x and y cell size. default unit						
phasecenter =		#	Image phase center: position or field index						
restfreq =		#	rest frequency to assign to image (see help 🚽						
Shell									



Pictures from a typical analysis session

3) where needed, tools have GUIs:

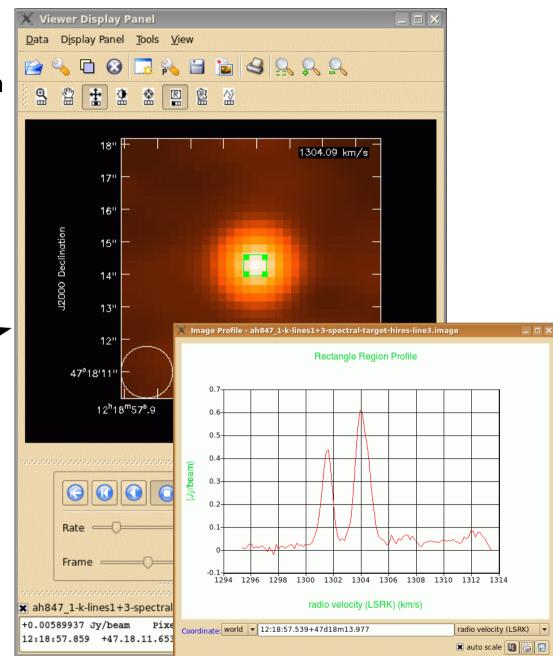
plotxy, plotcal, browsetable, viewer, clean

(started in separate threads)

The *viewer* is a powerful multifunction tool for data selection and visualization.

Uses Qt widget set (but 80% independent)

Rendering based on pgplot





A typical analysis session

3) where needed, tools have GUIs:

plotxy, plotcal, browsetable, viewer, clean

(started in separate threads)

browsetable permits you to explore any CASA table, e.g. Measurement Sets

Also Qt-based.

		50	/ 🕤 🍸	- 🚜 💈	1					
IEL	D	SOURCE SPECT		ordertest-ngc4	826.tutorial.ms				(
		UVW	FLAG	FLAG_CATEGOR	WEIGHT	SIGMA	ANTENNA1	ANTENNA2	ARR	
	0	[-6.2148, -9	[1, 64] Boole	[0, 0, 0] Boo	[0.103487]	[3.10854]	0	2	0	
	1	[-42.2483,	[1, 64] Boole	[0, 0, 0] Boo	[0.0862398]	[3.40523]	1	2	0	
	2	[18.8527, 33	[1, 64] Boole	[0, 0, 0] Boo	[0.189394]	[2.29782]	0	3	0	
•	3	[-17.1808, 0	[1, 64] Boole	[0, 0, 0] Boo	[0.157829]	[2.51714]	1	3	0	
	4	[-9.80468, 3	[1, 64] Boole	[0, 0, 0] Boo	[0.186832]	[2.31353]	0	4	0	
	5	[-45.8382, 2	[1, 64] Boole	[0, 0, 0] Boo	[0.155693]	[2.53434]	1	4	0	
	6	[25.0675, 42	[1, 64] Boole	[0, 0, 0] Boo	[0.115879]	[2.93763]	2	3	0	
•	7	[-3.58988, 4	[1, 64] Boole	[0, 0, 0] Boo	[0.114311]	[2.95771]	2	4	0	
	8	[-28.6574, 1	[1, 64] Boole	[0, 0, 0] Boo	[0.209203]	[2.18633]	3	4	0	
	9	[21.9573, -0	[1, 64] Boole	[0, 0, 0] Boo	[0]	[0]	3	8	0	
	10	[50.6147, -2	[1, 64] Boole	[0, 0, 0] Boo	[0]	[0]	4	8	0	
	11	[5.72437, -0	[1, 64] Boole	[0, 0, 0] Boo	[0]	[0]	3	5	• • •	
	Restore Columns Resize Headers									
F	AGE	NAVIGATIO Firs	st <<][1/	6]>> La	ast 1		Go	Loading 1000	row	



A typical analysis session

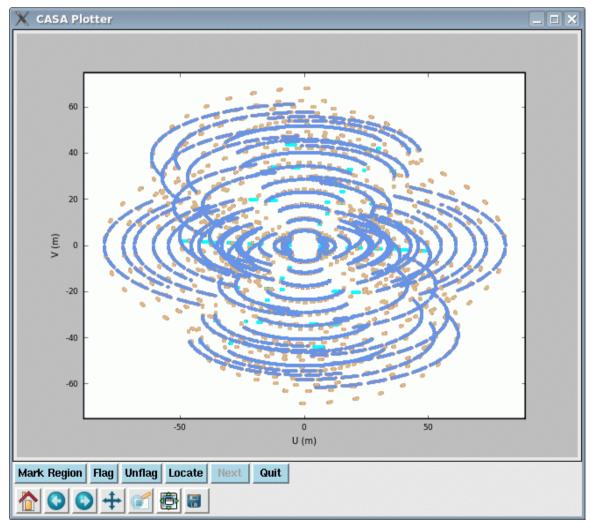
3) where needed, tools have GUIs:

plotxy, plotcal, browsetable, viewer, clean

(started in separate threads)

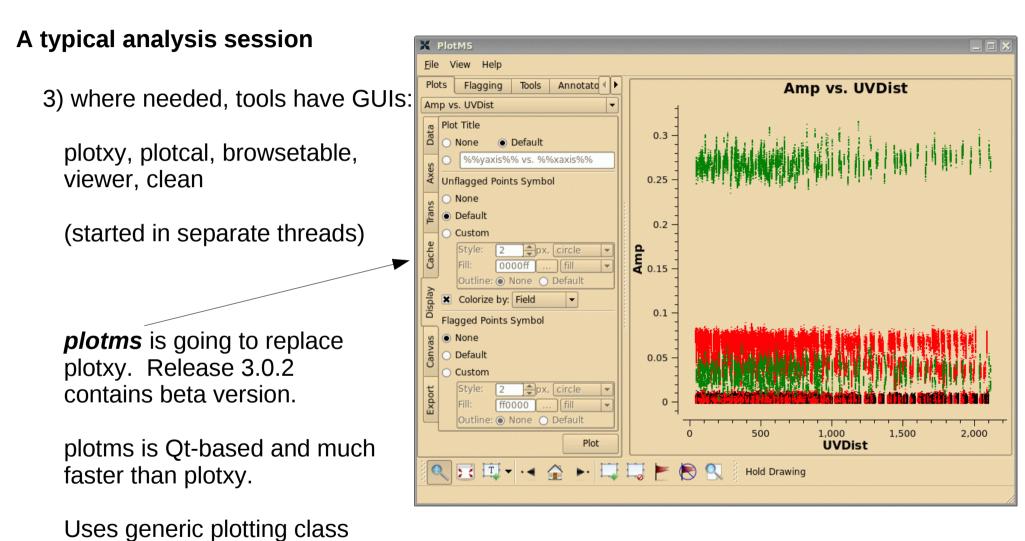
plotxy is a specialized tool for diagnostic plots and data selection

To be phased out.



D. Petry, MWA School, Amsterdam, June/July 2010





D. Petry, MWA School, Amsterdam, June/July 2010

which in turn uses Qwt.



Summary

- The standard science data analysis package for ALMA and EVLA is CASA
- Data from other observatories can also be processed, e.g. VLA, BIMA, ATCA, ...
- CASA derives from AIPS++ (partially survives in casacore)
- approx. 20 people are working on CASA in North America, Europe, and Japan
- CASA is a toolbox with
 - MATLAB-like user interface
 - $\boldsymbol{\cdot}$ GUI tools for data selection, browsing, and image processing
- the heart of the science analysis code is the *Measurement Equation*
- the internal data format are CASA Tables
- the *Measurement Set* is the CASA data format for visibility data (it is technically a Table with several well-defined sub-tables)
- CASA is publicly available under GPL for Linux and Mac OS X
- The first public release of CASA (version 3.0.0) became available in December 2009
- The latest release is version 3.0.2