I-TRAIN #1: Common questions and issues

- ● Questions related to the use of the imaging pipeline

Question:
Why would you want to do pipeline imaging as opposed to just using/making the images produced in the `scriptForPI.py`? `scriptForPI.py/scriptForImaging.py` *already has commands to complete imaging...*

Answer:
In the most recent ALMA cycles more than 90% of datasets are reduced and imaged with the ALMA Pipeline and in fact are not delivered with a `scriptForImaging.py`, unlike the first ALMA cycles where manual reduction scripts were made by analysts and the raw CASA commands scripted. The scripts now provided from the ALMA Science Archive download will be the ALMA pipeline task runs and will allow the reproduction of the images as made for QA2 analysis using the Pipeline automatically. This can be done using the `scriptForPI.py` after changing `pipererun` to 'True'. The imaging pipeline tasks run will be those we explored in this tutorial and explained how to run 'interactively'.

Working with the imaging pipeline has some advantages with respect to using the types of `tclean` commands as were found in the older style 'scriptForImaging.py':

1- It saves you the work to specifically extract each command, as per a `scriptForImaging.py` for each project & target. This is especially convenient if you have a large number of projects that you want to image. For the latest ALMA datasets there is no `scriptForImaging.py` delivered and so you can navigate into the delivered CASA-logs to find the relevant `tclean` command to reproduce a more traditional `scriptForImaging.py`-like task list.

2-The pipeline heuristics have improved significantly since first development and now it does a much better job, including auto-masking, without you having to "babysit" the process of imaging. It is simply a different way of working in which you launch the process/many processes and check the results afterwards, then adjust parameters as required and launch again - when compared to the usual interactive imaging (that relies on the CASA viewer).

3- The imaging weblog will be made for the images you produce, so you already get easy means to compare many images (without the CASA viewer) and are provided with statistics on noise and peak values, along with a representation of the image spectra within the clean masks for image cubes.

Question:
Why not to import the online flux catalog in the "`importdata`" step by selecting `dbservice=False`?

Answer:
Users don't have access to it. This is specifically for QA2 reduction. When any user downloads an archive dataset and opts to '`restore`', the data with the exact same run as was made for QA2 is restored. For imaging, you don't need the flux catalogue either - the data after restore is fully calibrated.

For more information within the downloaded package you should have a file "flux.csv" which is the one the analyst created at QA2. These were the best fluxes for the calibrators at the time from the QA2 run.

Question:
When should `h_save()` be done?

Answer:
In principle `h_save()` needs to be only done as the very last command before closing the CASA session. This will save the entire state of the pipeline run and the context correctly should you wish to later resume from where you stopped off in the future. This is the intended use case. It is good practice to do this often in an interactive pipeline session, in case you are likely to leave the session open and might accidentally close it after a few days.

Should CASA crash or (a pipeline task get killed) some stages after the save, it is somewhat possible to restore the session working from where the last `h_save()` was made. However, there will likely be corruptions to the weblog itself which will have stored stages subsequent to those after saving up to the point CASA was closed, and which will get filled again when analysis begins after a restore from a crashed session. Recovery after a crashed session is not the intention of `h_save()` nor should it be treated as such.

Question:
What are the advantages/differences between running the pipeline versions of the tasks and the "normal" ones. e.g. `hif_uvcontsub` *vs* `uvcontsub`*?*

There is no difference between the tasks for the same CASA version. They both do the same, but the "`hif_*`" task needs to be used for pipeline runs. The actual output is the same result, i.e. no advantanges/differences. The meanings of the several options `hif/hifa` … were explained during the training. You may find more information at the ALMA Science Pipeline User Guides corresponding to your CASA version.

Question:
If I didn't copy the file "`cont.dat`" to the working directory and run the `hif_findcont` *task. Can I rerun* `hif_findcont` *after moving "`cont.dat`"?*

Answer:
Yes, you can run `hif_findcont` again and now the file "`cont.dat`" will be used for the continuum in the subsequent steps (but this must be done before hif_uvcontsub) . In summary: the findcont task will assess the continuum ranges and report the ranges in the `cont.dat` file. However, if the `cont.dat` file already exists in your working directory before you launch findcont, this task will not create the continuum ranges but will use the ranges reported in the `cont.dat` file.

- Issues encountered when executing the tutorial commands

Question:

The pipeline folder was created after running `h_init()`, but there is no message shown in `Out[2]` giving the name of that folder, what is wrong?

Answer:

Sometimes the task only returns a message saying "`setting plot level to 'default'`" when creating a new context. Later when you save it, there should be a message indicating the context folder name (this could be machine-setup and CASA version dependent).

Question:

`Imageprecheck` has finished but the QA score is 0 and it says: `'Invalid Beam. hifa_imageprecheck: ImagePreCheck has failed with the beam'`
Is it okay to ignore and continue?

Answer:

Yes everything should still run even if `imageprecheck` had an issue.

The reason for `imageprecheck` is for QA2, where it reads the data to find out what the PI asked for in terms of a beam size and then pipeline creates parameters to try to achieve that. CASA 5.4.0 and 5.6.1 should return

```
ImagePreCheckResults:
     robust=0.50
     uvtaper=[]
```

Several participants had this issue with `imageprecheck` but the underlying reason has not been identified.

Question:

I obtain severe error messages when imaging `specmode='cont'` in steps 10-11 of the tutorial.

Answer:

Several participants had the same error messages:

```
ERROR: Cleaning failure for field RU_Lup
spw 16,18,20,22,24,26 specmode cont
SEVERE
pipeline.hif.tasks.makeimages.makeimages::pipeline.hif.tasks.makeimages.makeima
ges:: Cleaning
```

```
failure for field RU_Lup spw 16,18,20,22,24,26 specmode cont
```

We have reproduced the error, which also seems to be CASA-version dependent. Unfortunately we have not found the underlying reason yet (there are some hints that spectral window 16 may be problematic, TBC). This will need to be investigated for future tutorials that use this same dataset.

- ● Questions related to the web browser and weblog

Question:
Where is the "index.html" located?

Answer:
The file "index.html" is located in the folder `working/pipeline-today's data + other numbers/html/`.
There will be two different pipeline folders — make sure you choose the one you just created today (there will be one associated with when the calibrated ms was/were created - with an older date and time).

Question:
I have problems opening the weblog in Firefox. It opens the first page and then it does not open up the secondary links, why does it not open properly?

Answer:
Make sure you have applied the security fix that is recommended in this article.
Alternatively, you may want to try the following: First, type this in a terminal: " python -m http.server " …. then, in your firefox, you can go to the web "localhost:8000"

Question:
So Firefox is the suggested browser over Chrome or Safari?

Answer:
Yes, as explained in this article Firefox is the recommended browser for full functionality of pipeline weblog viewing, as Firefox is the browser which you can intentionally modify

the specific security settings to view the weblog. Chrome is explicitly not supported nor tested and thus not guaranteed to work '100%' - but probably it still does in some cases.

Question:
The weblog is not updating/the weblog doesn't get updated.
When trying reopening it does not seem to work. What could be wrong?

Answer:
Make sure that you have navigated and opened the correct pipeline folder - the one made at the time you run the tutorial (pipeline folders are named according to the date & time they are created). If you are in the wrong location in a pipeline folder with an old date, you will only see a `restore` data command - as this was from the task we ran to restore the data. Also bear in mind that the weblog only gets updated after each task/stage is completed. Let any running task finish as it will only write to the context directory and weblog when it is completed.

Question:
The weblog sent an error right from the beginning: Error loading file:///Users/.../2018.1.01201.S/science_goal.uid___A001_X133d_X2c85/group.uid___A 001_X133d_X2c86/member.uid___A001_X133d_X2c8b/calibrated/working/pipeline-202 01204T101503/html/stage1/t2-4m_details-container.html: 0 error. Did I set something up wrong?

Answer:
Make sure you are in the calibrated directory and running the commands there -- not in the weblog directory. You should have CASA open and running in one terminal (and not move from that working directory) and explore the weblog separately.

- Questions on automasking

Question:
How can I change automasking parameters?

Answer:
You can find a guide on automasking in the CASA guides: CASA Automasking Guide.