



How to setup and use the Mayday Processing Framework (MPF)

Florian Battke

Preparations

Essential prerequisites

The Mayday Processing Framework (MPF) requires no external libraries. If you want to embed R functions into processing pipelines, make sure to install the Mayday R Interpreter and its dependencies first.

Installing the MPF

To install the MPF, simply download the plugin from www.zbit.uni-tuebingen.de/pas/mayday/download/download.html and move the file into your Mayday plugin directory. Now start respectively restart Mayday. The MPF will then finish the installation and ask you to restart Mayday again.

Using the MPF

You can use the MPF in two distinct ways: You can apply existing processing modules to your data *or* create new processing modules (pipelines).

Applying processing modules

Starting the MPF

After selecting one or more probe lists as input, invoke the MPF plugin by right-clicking on one of the selected probe lists, open the “Data Processing” submenu and click on “Mayday Processing Framework”. This will open the MPF Applicator window.

Selecting a module

Here you can select a processing module and read the module description (see figure 1). You can also check data regarding the processing job you are about to create: You can see how many probe lists you selected and how many probe lists are expected by the respective processing module. These two numbers are used to determine whether enough input probe lists have been selected and whether the module will be run only once or in batch mode. Click “Next” to proceed.

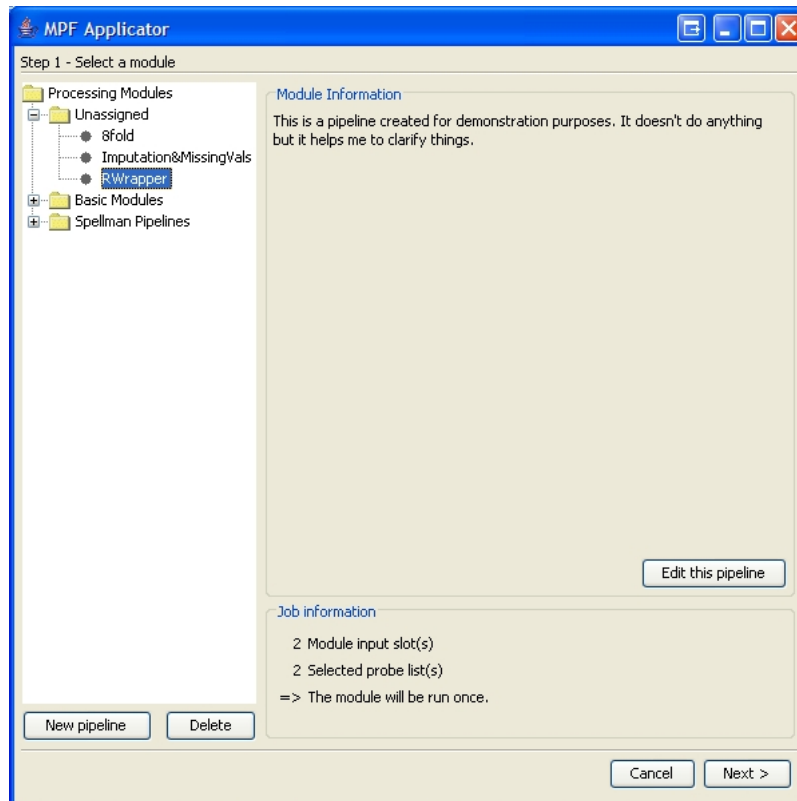


Figure 1: MPF Applicator, Step 1 - Module selection dialog

Assigning input slots

If the selected module expects more than one probe lists as its input, the next screen will show the input assignment dialog. You can now specify how the input probe lists are assigned to the module's input slots. Depending on the module, this can have a profound impact on your results! Click "Next" to proceed.

Module options

The final step is setting module options (see figure 2). The MPF Applicator will present a dialog where you can modify the module's default options. Modules should offer explanatory texts for every one of their options to guide your decisions. As soon as you're satisfied with all options, you can start the processing job.

If at any time you want to change your settings, simply use the "Back" button to review all options.

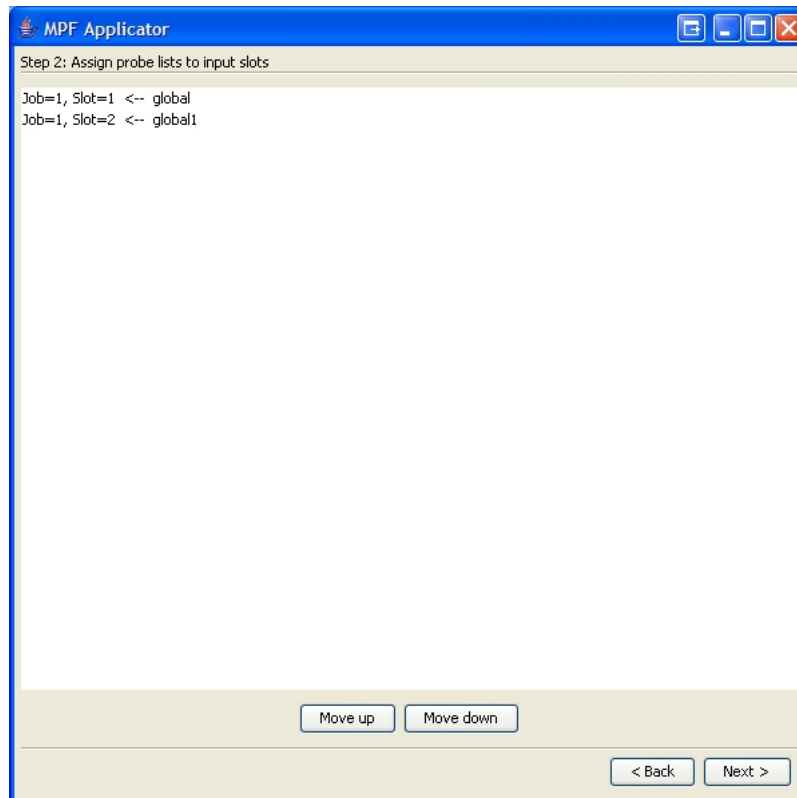


Figure 2: MPF Applicator, Step 3 - Module option dialog

Canceling execution

You can always cancel a running processing job by clicking on the “Cancel” button. Depending on the modules used, it may take some time before the job responds to the cancel request. This delay is mostly due to memory cleanup, so please be patient.

Log messages

When your processing job is finished, you should take a few moments to review the messages in the log (see figure 3) window before closing the MPF Applicator. The log window is used to alert you to any problems that arose during processing. Most often, problems are due to malformed input (i.e. probes with missing values).

Creating new processing modules

Starting the MPF Designer

To create a new module, start the MPF Applicator by right-clicking on a probe list and selecting “Mayday Processing Framework” from the “Data Processing” submenu. In the Applicator, click on “New” to invoke the MPF Designer.

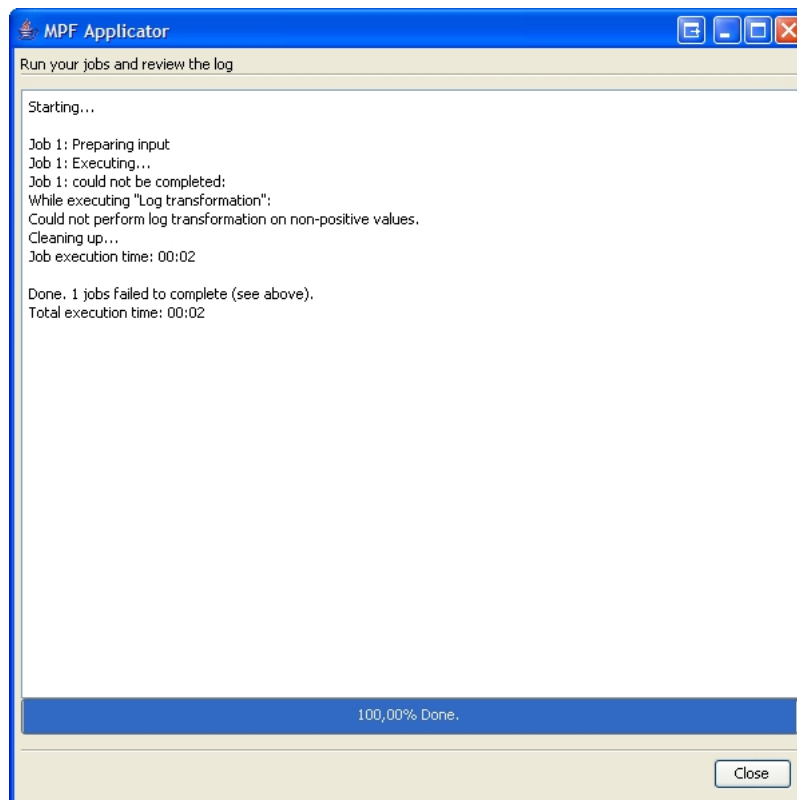


Figure 3: MPF Applicator, Reviewing log messages

Creating a new processing pipeline consists of a few basic steps:

- Adding modules to the pipeline
- Setting module options
- Connecting the modules
- Configuring global pipeline options

Adding modules

To add a new module to the pipeline, simply select the module from module list in the “Add module” menu (in the Designer menu bar). The new module will appear in the pipeline view. You can set the module’s options by clicking on its “Options” button.

Connecting modules

Once you have more than one module in your pipeline, you can create the pipeline graph (see figure 4) by connecting modules together. To connect two modules A and B, start by clicking on an output slot of module A. Then click on one of the

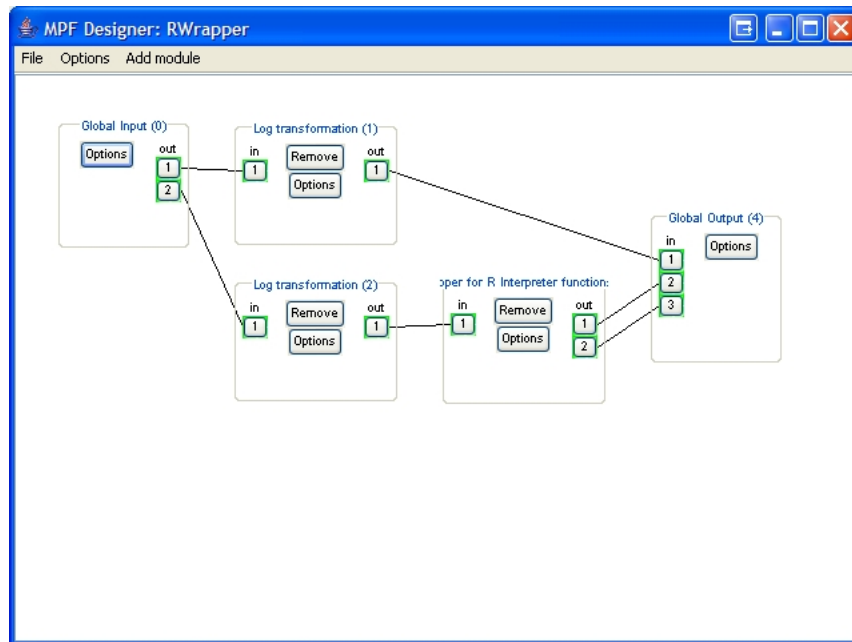


Figure 4: MPF Designer showing a pipeline

input slots of module B. This will link the two modules and the pipeline graph will be updated to reflect the new logical ordering of the modules, i.e. “A before B”. To disconnect the two modules, simply create a new connection either starting at A’s output slot or ending at B’s input slot.

Removing modules

To remove a module from your pipeline, click the appropriate “Remove” button. The pipeline view will be update to reflect your changes.

Validating the pipeline

If you’re done connecting the modules, you should check whether the pipeline graph is valid. In the Designer menu bar, click on “Options”, then on “validate”. Only valid pipelines can be used to process data.

Pipeline options

Finally, save your new pipeline by selecting “Save as...” from the “File” menu in the Designer menu bar. This will bring up the global pipeline options window (see figure 5). Here you can name your pipeline, write a description of what your pipeline does and set the pipeline’s version number. Please read the provided help texts for more information on these options. Click on the “?” button next to an option’s name to read its help text.

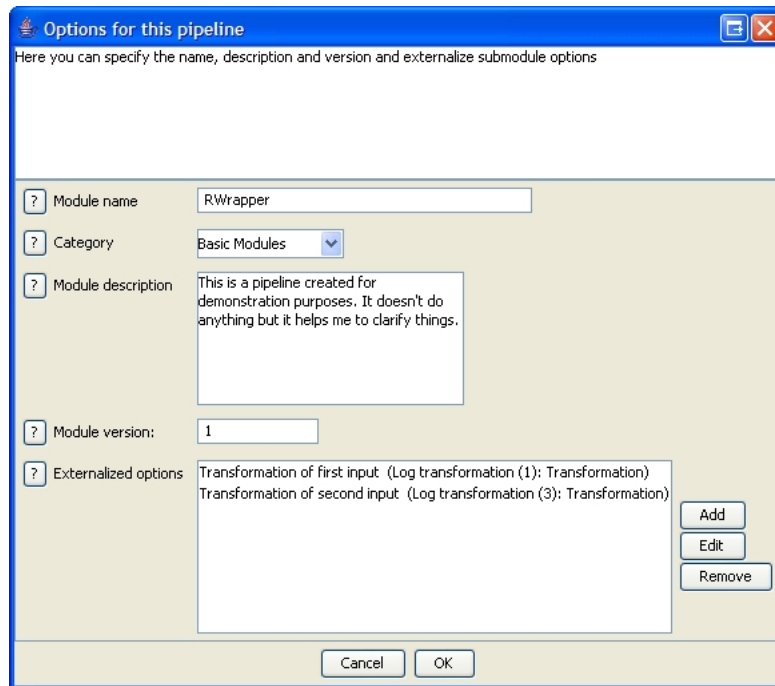


Figure 5: MPF Designer showing the pipeline options window

The global pipeline options dialog can also be opened from the Designer menu bar by clicking on “Options” and then selecting “Module options”.

Externalized options

An important decision when designing a pipeline is which internal options should be exposed to users of the pipeline. Such options are called “externalized” options within the MPF. Options can be externalized from the global pipeline options window. For each externalized option, a new name and description as well as a default value can be set (see figure 6).

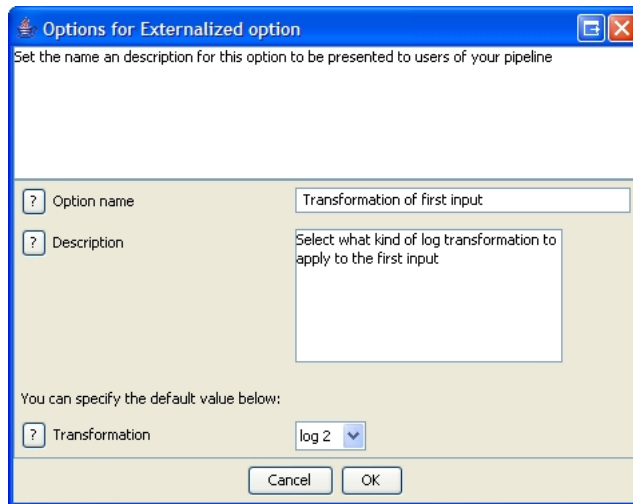


Figure 6: Configuring an externalized option