



Getting Started with Visualization in *Mayday*

Nils Gehlenborg

1 Visualization in *Mayday*

Before we get started a word of warning: This How-To describes only fundamental features of *Mayday*'s visualization facilities. This is neither an in-depth introduction nor a tutorial. Be aware of this. Also keep in mind that *Mayday*'s visualization features are continuously being extended and improved, thus the information in this document might not always reflect the current state of the software.

At the moment there exist six visualization plug-ins for *Mayday*:

- A tabular viewer that visualizes the actual values of the expression data in form of a table.
- A box plot.
- A multi box plot, i.e. a viewer that displays several box plots in a grid.
- An interactive profile plot.
- An interactive multi profile plot, i.e. a viewer that displays several interactive profile plots in a grid.
- An interactive enhanced heatmap.

Each of these plug-ins and their common features will be briefly described below.

2 Downloading and Installing the Required Software

Mayday offers functionality through plug-ins that are developed independently of the core software. Thus it is necessary that you download and install the visualization plug-ins for *Mayday* before you can visualize your data.

Visit <http://www.zbit.uni-tuebingen.de/pas/mayday/download/index.html> and download the latest snapshot or release version of the visualization plug-ins. Snapshots are usually uploaded to the website when new features were added and when bugs were fixed, they might not have been tested thoroughly.

Currently all visualization plug-ins are bundled into a single jar-archive file named `mayday-visualization.*.jar`. The full name depends on whether you want the snapshot or the release version.

Download the jar-archive into the plug-in directory of your *Mayday* installation. If you plan to export visualizations of your data to image formats such as PNG, JPEG, TIFF or SVG you also have to download the *Batik* library provided on the *Mayday*



website. Alternatively you can download it from the website of the Apache project (<http://xml.apache.org/batik/>). For the time being it doesn't matter where you install the *Batik* library, as long as you remember where!

If you were running *Mayday* while you installed the plug-ins close your session and restart *Mayday* so the new plug-ins will be loaded. Go to "File → Plug-ins" and select the tab "Visualization". What follows is a slightly tedious process: for every visualization plug-in you have to specify the path to the *Batik* library installation. To do so select the plug-in from the list on the left and click the "Preferences" button. Then click on the "Add ..." button and go to your *Batik* installation. Do you still remember where you placed it? Go to the "lib" subdirectory and select all jar-archives beginning in "batik-" and confirm. Hit "OK" in the preferences dialog. Do this for each visualization plug-in.

3 Some Important Background Information

As for all other plug-ins the basic unit is a probe list. Thus you can only visualize probe lists, i.e. sets of probes, never individual probes¹.

In order to visualize a probe list you have to select it (or more) in the probe list manager. The probe list manager is the list that displays all probe lists of a data set. Click the selected probe list with the right mouse-button (right-click it) to open the context menu. At the bottom of the menu you'll find the plug-in categories. The sub menu "Visualization" lists all plug-ins that you can use to visualize your data. Select one of the visualization plug-ins. The following will (should?) happen: Two windows will open, one contains the viewer you selected in the sub menu and the other one will contain the tabular viewer. Currently it is not possible to open any other viewer without opening the tabular viewer. Nor is it possible to open the tabular viewer individually as it is part of the so-called "visualizer".

Here is why: A visualizer manages one or more viewers and will be created every time you select a visualization plug-in via the probe list manager. This means that you can have more than one visualizer! They keep track of which viewers have been activated and link them to each other. If you want to add a viewer to a visualizer you can do so by selecting one from "Viewers → New → ..." in the window that contains the tabular viewer. The number of the visualizer that a viewer belongs to is displayed in the viewer's window title. You can get an overview of all viewers of a visualizer by clicking on "New". At the bottom of the menu is a list of all viewers that are managed by the visualizer. Select one to bring its window to the front.

As mentioned before, a visualizer links its viewers to each other. For instance, if you have two profile plots that belong to the same viewer and you select the profile of a probe (and thus the probe) in one of them it will also become selected and highlighted in the other profile plot (if it is not hidden, more on that later). This works with all kinds of viewers that allow you to select probes.

¹Well, this is a little confusing since you could create a probe list that contains only a single probe and then you would actually be able to visualize an individual probe.



A visualizer not only manages viewers but also the data that those viewers display, i.e. one or more probe lists. It is possible to add and remove probe lists from a visualizer and in consequence from the viewers of that visualizer. If you close a visualizer all of its viewer will be closed automatically without a warning. Keep this in mind!

4 Common Features of all Viewers

4.1 The Almighty Context Menu

As you will notice quickly there are no menu bars or tool bars in the viewer windows². This is to save space for visualization of the actual data. The features of the viewers can be accessed using a context menu similar to the one of the probe list manager in the main window. The context menu can be accessed by right-clicking anywhere in the area of the plot.

4.2 Color

In the enhanced heatmap viewer, the tabular viewer and profile plots probe identifiers and profiles, respectively, are plotted in the color of their top level probe list. How the top level probe list is determined is too complex to describe here. In general the top level probe list corresponds to that probe list out of all probe lists the probe is contained in, which is assigned the highest priority in the probe list manager. The higher up the probe list is in the probe list manager, the higher the probe lists priority.

Please note that probes cannot be assigned a color individually. Colors can be assigned to probe lists using their properties dialog, which is for instance accessible from the probe list manager.

4.3 Layers

The enhanced heatmap viewer, the tabular viewer and profile plots are organized into layers. Layers are a concept adopted from image editing and desktop publishing software. In *Mayday* each probe list contained in a viewer represents a layer.

Think of a stack of transparencies and you get a good idea of what the layers concept does. Layers can be rearranged to either the top or the bottom of the stack. Further they can be hidden, i.e. temporarily removed from the stack.

In profile plots the layer concept enables you to get an undisturbed view on a probe list of interest by making it the top most layer. In the enhanced heatmap and the tabular viewer rearranging layers changes the order in which the probe lists occur in the respective viewer *if the probes are sorted according to the probe list they belong to*.

²The tabular viewers is a special case because it is integrated into the visualizer.



4.4 Sorting

As mentioned in the section on layers, it is possible to modify the order in which probes appear in the enhanced heatmap and the tabular viewer. It is possible to sort probes according to their assigned identifier, their assigned top level probe list and according to meta information objects. Sorting according to meta information objects is only possible if the meta information objects can be sorted themselves.

4.5 Selection

In a viewer that displays a probe as an individual entity, i.e. all viewers but box plots, you can select one or more probes for further processing.

Probes are selected by clicking on their representation in the viewer. If multiple probes are to be selected the Ctrl key has to be pressed during clicking. The selection can be cleared by clicking in an empty area of the plot³. Selected probes are highlighted depending on the viewer you are using. Another way of selecting a single probe is by using the “Go To →Probe ...” function of the context menu. There you can enter the name of a probe and if it is found it will become selected.

So what can you do with selected probes? For instance, it is possible to look up probe identifiers of selected probes in databases like Ensembl or the Saccharomyces Genome Database (SGD). Database lookup can be accessed via “Look up selection →...” in the context menu. Besides database look up it is also possible to create a new probe list from a set of selected probes. This is done by selecting “Probe List from Selection ...” in the context menu.

If you double-click an unselected probe representation in one of the viewers mentioned above, you will see the properties of that probe displayed in a dialog. In this dialog you can double click on probe lists and meta information objects displayed in the corresponding lists to obtain more information.

It is important to note that the selection is a property of the visualizer, i.e. if you select a probe representation in one viewer of a visualizer, the probe will also become selected in all other viewers of that visualizer. This feature is very helpful for data exploration.

4.6 Export

You can export data displayed in any viewer to an external file selecting “Export ...” in the viewer’s context menu. The tabular viewer is exported to a text file, while the graphical viewers are exported to image files.

Please note that you need to have the Batik library installed and setup for use with *Mayday* to export image files.

³In this case you actually have to click withing the area of the coordinate system not into the gray area surrounding it.



5 Special Features of Graphical Viewers

5.1 Zooming

All graphical viewers can be zoomed. You can zoom in and out of a visualization by selecting “Zoom →Zoom In” or “Zoom →Zoom Out” respectively. Alternatively you can use the corresponding keyboard shortcuts.

If you want to fit the frame of the containing window to the size of your visualization click “Zoom →Fit Frame”.

5.2 Axes, Grids and Labels

Minima and maxima of axes, or of the color gradient in case of the heatmap viewer, are equal to the minimum and maximum values in the whole data set. The color gradient is artificially made symmetrical by using the maximum absolute value.

Labels on axes are only displayed if there is enough room to show them. If you don't see labels on axes you should try to zoom in. If you see labels on axes but actually don't want to see them you can turn them of. See “Settings →Scales →...” for options.

All graphical viewers feature a grid that can be turned on or off. Grid lines can be turned on and off individually per axis. In profile and box plots the grid can be displayed on top of everything else contained in the plot. Features concerning the grid can be accessed via “Settings →Grid →...”.

5.3 Caption and Legends

In order to include information about probe lists and settings in exported viewers you can add legends to your graphical viewers and set a descriptive caption. For options see “Settings →Legends →...” and “Settings →Caption →...” respectively.

5.4 Multi Plots

Multi profile and multi box plots are viewers that contain a grid of several profile plots respectively box plots. If you open a new multi plot you will be asked to define the dimensions of the grid and which probe lists should be displayed at which position in the grid.

6 A Little More Detail on the Enhanced Heatmap Viewer

6.1 General

The heatmap viewer implemented in *Mayday* has several unique features. For instance, the heatmap viewer can integrate meta information about probes into the visualization using various graphical representations. More information about the



concept and theory of the enhanced heatmap viewer can be found in Nils Gehlenborg, Janko Dietzsch and Kay Nieselt, A Framework for Visualization of Microarray Data and Integrated Meta Information, *Information Visualization* (2005) 4, 164 - 175 (doi:10.1057/palgrave.ivs.9500094). See the Meta Information How To for more information on how generate and load meta information in *Mayday*.

6.2 Browsing the Heatmap

The heatmap viewer is organized into pages that each hold a certain number probes. You can define the number of probes per page depending on your need. See the “Settings → Probes Per Page →...” sub menu for options.

Pages can be browsed using either the “Pages →...” sub menu in the context menu or the following shortcut keys:

- Ctrl + Page Up: Go to previous page.
- Ctrl + Page Down: Go to next page.
- Ctrl + Home: Go to first page.
- Ctrl + End: Go to last page.

6.3 Enhancing the Heatmap with Nominal Information

You can add additional columns containing nominal, i.e. categorical, meta information to the enhanced heatmap viewer. See the Meta Information How To for more information on how to obtain and handle meta information in *Mayday*.

Assuming you have loaded nominal meta information for at least some of the probes contained in the heatmap, you can add an additional column via “Enhance → Additional Columns ...” in the context menu. Select “Add Nominal Column” in the dialog and one or more meta information object groups. You can change the colors assigned to the categories by clicking on the corresponding cells in the table. You can also change the name of the new column.

Once you have added the new column you can change its settings (colors, name) either by holding down the Alt key and double-clicking the a cell of the corresponding column or via “Enhance → Additional Columns ...” in the context menu.

6.4 Enhancing the Heatmap with Scoring Information

If you want to enhance the heatmap viewer with scoring information you need to have loaded scoring meta information objects for at least some of the probes contained in the heatmap. See the Meta Information How To for more information.

Additional columns Adding further columns containing scoring information to the enhanced heatmap works exactly like adding nominal columns to the heatmap viewer. However, instead of one color per category a color gradient is employed



that ranges from 0 (not relevant at all = lowest score) to 1 (most relevant = highest score). You have to define only the colors for 0 and 1.

Height Scaling Meta information for scoring probes can also be used to scale the vertical height of heatmap rows. You can access this feature via “Enhance →Height ...” in the context menu. Using the threshold you can determine how many rows will be scaled up and how many will be scaled down. Further you can set a maximum scaling factor and a minimum scaling factor. Remove or edit the height scaling by clicking “Remove” or “Edit ...” respectively, in the meta information selection dialog.

Additional Color Gradient Another way to include meta information into the heatmap viewer is through use of an additional color gradient. The additional color gradient is orthogonal to the basic green-black-red color gradient. It can be either be adding a blue component or an alpha (transparency) value to the original gradient. Probes with high scores (close to 1) will either be represented with a large blue component (i.e. colors are from a cyan-blue-magenta gradient) or very opaque (i.e. not transparent or faded out). Access this feature via “Enhance →Color →...”. Remove the additional color gradient by clicking “Remove” in the meta information selection dialog.

6.5 Changing the Color Gradient

If you don't feel comfortable with the default green-black-red color gradient you can change it via “Settings →Color Gradient ...”. Besides colors you can chose between two different types of gradients, either a linear gradient or a sigmoid gradient. Further it is possible to reduce the resolution of the gradient and thus the of the heatmap, which is sometimes useful to highlight different classes of expression profiles.

6.6 Accessing Data

The values underlying the color-encoding in the heatmap can be accessed by holding down the Alt key and right-clicking a cell of interest in the heatmap. A tool tip like display will show the value encoded by the corresponding color. This works both for regular cells as well as for cells in additional columns.

7 More Visualization Options

Further visualization options in *Mayday* are offered by the hierarchical clustering plug-in, which provides several variants of a tree view of a hierarchical clustering. The R interpreter plug-in offers a manifold of options for visualization of data, however, most of those visualizations are not interactive.



8 Known Issues

8.1 CPU Usage

Profile plots and in particular multi profile plots are very expensive to compute. Depending on the size of your data set and your machine it might take up to several minutes to compute a multi profile plot. The plots have to be recomputed every time you interact in any way with the visualization. Currently there is no progress indicator and the machine might seem to be stalled. It is not! Keep this in mind if you work with large data sets.

8.2 Memory Usage

Currently some viewers cause “memory leaks”. That means that they do not return the memory resources they were using even after you close them. This will make your machine slow and eventually cause the Java Virtual Machine to reach it’s maximum heap size. We are working on this problem and it will be fixed as soon as possible.

8.3 Multi Plots

There are several known issues with the multi profile plot and the multi box plot.

- Zooming into the visualization sometimes causes the visualization to be partially drawn outside the visible area. A workaround for this bug is to zoom in one more time and then to zoom out.
- Exporting multi plots is generally possible, however there are problems with defining exact dimensions. In some cases the exported images were found to be much smaller or larger than defined. If you want to export a multi plot just zoom it to the size you would like to export and then export it without defining new dimensions.

8.4 Keyboard Shortcuts

On some computers the keyboard shortcuts displayed in the context menus work, on others they don’t. They don’t work properly on Macs and they don’t work properly with some keyboards with an American keyboard layout. You can always use the corresponding menu entries, even though that might be annoying in some cases. That’s why we introduced shortcuts in the first place ...

This Mayday How-To was written and edited by Nils Gehlenborg. If you have comments or questions please contact the author via email, nils@gehlenborg.com. The latest version of this document can be found at <http://www.zbit.uni-tuebingen.de/pas/mayday>.
