

## PyMca HDF5 support

Version 4.4.0 of PyMca introduced HDF5 (<http://www.hdfgroup.org/HDF5/>) file format support using Andrew Collette's h5py library (<http://code.google.com/p/h5py/>) as binding to the HDF5 library.

A simple analogy of an HDF5 file is that of a hard disk. A hard disk can contain files that can be into folders that in turn may contain other folders. An HDF5 file contains datasets (your data) that can be arranged into groups that in turn may contain other groups. The analogy goes till the point that you can create links between datasets or groups and that to access a dataset or a group you have to provide the path to it.

Obviously, from a graphical user interface point of view, the logical access to an HDF5 should be provided by something similar to a file browser. The HDF5 file browser used in PyMca is based on a contribution by Darren Dale.

The data in an HDF5 file provide information about their size and type but they do not provide information about what they represent. Therefore, the approach followed by PyMca to properly visualize the data is cumbersome (at least when used for first time) but simple. The approach is based on creating a selection table with the datasets of interest. This can be achieved by double clicking the relevant datasets or via a right-button mouse click. The nice feature is that the table provides a context menu (right-button mouse click) allowing the user to save or load selection tables therefore reducing the need to repetitively browse the file. In addition, the selection table is saved among the PyMca settings (File Menu -> Save ->PyMca Configuration or File Menu -> Save Default Settings).

Once the datasets of user interest are in the table, he can select what datasets are to be used as axes (first table column containing checkboxes), as signals (second column containing checkboxes) and eventually as monitor (third column with checkboxes). The only selection that is mandatory to generate a plot is the one corresponding to the signal. In case of selection of several axes, the order in which the check boxes were selected determines the dataset to be used as first, second or third axis.

## PyMca NeXus support

NeXus (<http://www.nexusformat.org>) provides a set of directives to share data among different facilities. It provides an API supporting an HDF4 backend, an XML backend and an HDF5 backend. PyMca does not use the NeXus API and therefore only supports the HDF5 backend. By other hand, HDF5 is the most common NeXus backend used at large scale research facilities.

NeXus HDF5 files can be handled in the same way as standard HDF5 files. In addition, PyMca will try to make as much use as possible of metadata, default plots and application definitions provided by NeXus to reduce user interaction.