Lab Session 6: HDF5 Cheatsheet

This sheet provides a concise overview of HDF5 libraries for Python (h5py) and R (rhdf5). The content is limited to functions that are necessary for solving the tasks of Lab Session 6. You can also use additional functions or HDF5 functions from different libraries (e.g., pandas).

1 HDF5 Library for Python

This section documents essential functions of h5py.

h5py.File(filename, 'w')
Creates a new HDF5 file called `filename`. `'w'` creates a file with write access, overwriting existing files. Other relevant parameters are `'r'` for read-only access and `'r+'` for read/write access for an existing file.

f.create_group(name)
Creates a new group in HDF5 file `f`. Subgroups can be specified by giving the full path, e.g. 'meas/hypersnort/foo'. Returns a group object.

g.create_dataset(name, data=<data>)
Creates a new data set called `name` in group `g`. The data of the date set is specified in `data`. Pay attention to use an appropriate type, e.g. a numpy array. Returns a data set object.

f[path]
Reads the specified path of HDF5 file `f`. This can return groups, as well as, data sets.

x.attrs[attr] = val
Assigns the value `val` to attribute `attr` in group/data set object `x`.

2 HDF5 Library for R

This section documents essential functions of rhdf5.

h5createFile(filename)
Creates a new HDF5 file called `filename`.

h5createGroup(filename, name)
Creates a new group called `name` in the HDF5 file identified by `filename`. Subgroups can be specified by giving the full path, e.g. 'meas/hypersnort/foo'.

h5write(data, filename, path)
Writes `data` in the HDF5 file identified by `filename` in the group specified by `path`.

H5Fopen(filename)
Opens the HDF5 file identified by `filename`. Returns a file ID.

H5Gopen(fid, path)
Opens the group identified by `path` in the HDF5 file identified by file ID `fid`. Returns a group ID.

H5Dopen(gid, path)
Opens the data set identified by `path` in the HDF5 group identified by group ID `gid`. Returns a data set ID.

h5writeAttribute(attr=<attr>, h5obj=<id>, name=<name>)
Writes the attribute(s) called `<name>` with value(s) `<attr>` to the HDF5 object identified by `<id>`.

h5readAttributes(filename, name=<name>)
Reads the attribute from the HDF5 File identified by `filename` in the group or data set specified by `<name>`. Note that this is not the same parameter as in `h5writeAttribute`.