H5edit Phase 2 Plan

(Rev: July 24, 2012 by Albert Cheng)

1   Purpose

This describes the new features of the phase 2 implementation of the H5edit tool. It describes the phase 2 features in the order of recommended implementation priority.

2   Phase 2 features

2.1   Atomcity option

2.1.1   Description

Commands are applied to the data file in an atomic manner, that is, the file will not be modified only partially if H5edit encounters error during the execution. Currently, H5edit recommends users to make a backup copy of the data file before launching the command. If there is any execution error, the user may recover the data file from the backup copy.

The Atomicity option will first make a backup copy of the data file before applying the command input. If the tool encounters any error, the user may recover the data file from the backup copy.

The Atomicity option, --atomic, supports 3 levels, *yes*, *no*, and *inc*.

> --atomic=yes (default) applies the input commands in an all or none manner. That is, either all input commands are applied to the data file, or none is applied.
>
> --atomic=no will not make the backup copy and apply the input commands as much as possible. The user may use this if he is sure of the validity of the input commands or he is not concern if the data file is partially modified or corrupted.
>
> --atomic=inc will apply the changes in an all or none manner but at an command level. If H5edit encounters an error during execution, the data file can be restored back to the last command applied successfully.

2.1.2   Benefits

The Atomicity option provides the user an automatic backup of the data file in case of any execution errors. This is an important error recovery feature and provides an insurance for the users' data files.

2.1.3   Implementation time estimate

For design, implementation, testing and documents: 72 hours

2.2   Dryrun option

2.2.1   Description

Verify input commands are correct without executing them. (No change apply to the data file).

2.2.2   Benefit

This is nice to have, especially before the Atomicity option is implemented. On the other hand, it cannot be 100% sure since it can check on the validity of commands that depend on the result of previous command. For example:
CREATE /Dataset1/Attribute1 …; DELETE /Dataset1/Attribute1;
Since the creation of Attribute1 is not applied, the data file does not have Attribute1. When the second command tries to DELETE it, the Dryrun option will not validate it.

2.2.3   Implementation time estimate

For design, implementation, testing and documents: 26 hours

2.3   More H5edit Commands

2.4   RENAME

This changes the name of an existing attrbute.
Implementation time estimate: 16 hours

2.5   COPY

This creates an exact copy of an existing attribute under a different name.
Implementation time estimate: 20 hours.