Subject: Looking for IDL Structures to Python compatible HDF5 Examples
Posted by eben.pendleton on Tue, 01 Sep 2015 17:33:03 GMT
View Forum Message <> Reply to Message

Hi all,

I'm looking for code examples of writing ODL structures to HDF5 tables that can be read in Python (2.7). Filenames are removed

I've tried the following approach using the IDL coyote guide to create the file in IDL 7.1:

```
pro write_idl_struct_to_hdf5,struct,dir_out,outfile

  file = dir_out+outfile+'.h5'
  fid = H5F_CREATE(file)

  datatype_id = H5T_IDL_CREATE(struct)
  dataspace_id=H5S_CREATE_SIMPLE(1) ; not so simple..

  dataset_id = H5D_CREATE(fid,'o',datatype_id,dataspace_id)
  H5D_WRITE, dataset_id, struct

  H5S_CLOSE, dataspace_id
  H5T_CLOSE, datatype_id

  H5F_CLOSE, fid

  print,'---Finished Writing IDL struct to HDF5---'
end
```

Python code and Error

```
import pandas as pd
dir_in=r'C:\D_Drive\Python\hdf5'
infile=...

mynewdf = pd.HDFStore(dir_in+'\\'+infile)
print(mynewdf)

print(mynewdf.select('o',columns='TASK_NAME'))
```

resulting in:

```
<class 'pandas.io.pytables.HDFStore'>
File path: ...
/o          frame_table [0.0.0] (typ->generic,nrows->1,ncols->63,indexers->[index],dc->
[TASK_CODE,TASK_NAME,EVENT,SYS_LOC_CODE,SUBFACILITY_CODE,ARE
A,RIVER_STREAM_MILE,MILES_FROM_NC_MOUTH,X_COORD,Y_COORD,COOR
```

D_TYPE_CODE,REFERENCE_POINT,ELEV,ELEV_UNIT,ELEV_COLLECT_METH
OD_CODE,ELEV_DATUM_CODE,WATER_COLUMN_DEPTH,WATER_COLUMN_DEPT
H_UNIT_CODE,SYS_SAMPLE_CODE,PARENT_SAMPLE_CODE,MATRIX_CODE,L
AB_MATRIX_CODE,SAMPLE_TYPE_CODE,SAMPLE_DATE,START_DEPTH,END_
DEPTH,DEPTH_UNIT,DEPTH_INTERVAL,SEDIMENT_NATIVE,GROUP_DESC,C
HEMICAL_NAME,CAS_RN,FRACTION,ANALYTIC_METHOD,LEACHATE_METHOD
,PREP_METHOD,RESULT_VALUE,LAB_QUALIFIERS,VALIDATOR_QUALIFIER
S,INTERPRETED_QUALIFIERS,DETECT_FLAG,TARGET_UNIT,REPORTABLE_
RESULT,VALIDATED_YN,VALIDATOR_REASON_CODE,BASIS,METHOD_DETEC
TION_LIMIT,REPORTING_DETECTION_LIMIT,DILUTION_FACTOR,LAB_NAM
E_CODE,LAB_SAMPLE_ID,ANL_SHORT_NAME,PARENT_LOC_CODE,DATA_REL
EASE_STATUS,DATA_ANALYSIS_HIERARCHY,RESULT_VALUE_MDL,RISK_SC
REEN_USABILITY,BASELINE_RA_USABILITY,CAS_RN_4PROUCL,CHEMICAL
_NAME_4PROUCL,TREATMENT_FLAG_4PROUCL,CHEMICAL_NAME_4LABEL,US
ABILITY_HIERARCHY])

The error generated is
ValueError: Wrong number of items passed 283774, placement implies 1

The 283774 number is the number of rows in the structure and it seems that each tag in the dataframe has one row with 283774 elements.

Is there some layout adjustment that's needed to read a structure correctly in HDF5?

Eben

Subject: Re: Looking for IDL Structures to Python compatible HDF5 Examples
Posted by markb77 on Tue, 01 Sep 2015 17:49:24 GMT
View Forum Message <> Reply to Message

hi Eben,

HDF5 Tables are a specific format of HDF5 file, and I suspect that's what you're running into here - you are not really creating a "table" according to the HDF5 Tables API. IDL has no native interface for writing HDF5 tables, however I have written an IDL library to do this.

you can look here:

http://www.github.com/superchromix/wmb_lib

the library does not have great documentation, but look in the directory

source/hdf5/h5tb

for the file wmb_h5tb_examples.pro, and this will show you how to use the library to create HDF5 compliant tables from IDL.

Another method of testing your HDF5 files: try to open them with HDF View, a freely distributed program for viewing all types of HDF5 data.

Mark

Subject: Re: Looking for IDL Structures to Python compatible HDF5 Examples
Posted by Alain Kattnig on Wed, 02 Sep 2015 07:20:34 GMT
View Forum Message <> Reply to Message

Le mardi 1 septembre 2015 19:49:26 UTC+2, superchromix a écrit :
> hi Eben,
>
> HDF5 Tables are a specific format of HDF5 file, and I suspect that's what you're running into here - you are not really creating a "table" according to the HDF5 Tables API. IDL has no native interface for writing HDF5 tables, however I have written an IDL library to do this.
>
> you can look here:
>
> http://www.github.com/superchromix/wmb_lib
>
> the library does not have great documentation, but look in the directory
>
> source/hdf5/h5tb
>
> for the file wmb_h5tb_examples.pro, and this will show you how to use the library to create HDF5 compliant tables from IDL.
>
> Another method of testing your HDF5 files: try to open them with HDF View, a freely distributed program for viewing all types of HDF5 data.
>
> Mark

You might also use the superior, in my opinion, Python-based Vitables

Best

Subject: Re: Looking for IDL Structures to Python compatible HDF5 Examples
Posted by eben.pendleton on Wed, 09 Sep 2015 20:28:57 GMT
View Forum Message <> Reply to Message

Thanks Mark and kallisthene. The code and table viewer you have provided work great. Many thanks.

One question however: I'm having trouble with the chunksize when using an iterator to restore the

data in Python. The python error is ValueError: Shape of passed values is (1, 141887), indices imply (1, 283774)
whereas my structure is 283774 and my chuck size is 141887 (283774/2) Do I have to adjust something when creating the databuffer to get the indices to work?

Thanks,
Eben

On Wednesday, September 2, 2015 at 3:20:38 AM UTC-4, kallisthene wrote:
> Le mardi 1 septembre 2015 19:49:26 UTC+2, superchromix a écrit :
>> hi Eben,
>>
>> HDF5 Tables are a specific format of HDF5 file, and I suspect that's what you're running into here - you are not really creating a "table" according to the HDF5 Tables API. IDL has no native interface for writing HDF5 tables, however I have written an IDL library to do this.
>>
>> you can look here:
>>
>> http://www.github.com/superchromix/wmb_lib
>>
>> the library does not have great documentation, but look in the directory
>>
>> source/hdf5/h5tb
>>
>> for the file wmb_h5tb_examples.pro, and this will show you how to use the library to create HDF5 compliant tables from IDL.
>>
>> Another method of testing your HDF5 files: try to open them with HDF View, a freely distributed program for viewing all types of HDF5 data.
>>
>> Mark
>
> You might also use the superior, in my opinion, Python-based Vitables
>
>
> Best

Subject: Re: Looking for IDL Structures to Python compatible HDF5 Examples
Posted by markb77 on Thu, 10 Sep 2015 21:04:38 GMT
View Forum Message <> Reply to Message

does the table open correctly in HDF View? If so, could this be a bug in the Python code?

Mark

Subject: Re: Looking for IDL Structures to Python compatible HDF5 Examples
Posted by marcomottola87 on Wed, 17 May 2017 10:41:19 GMT
View Forum Message <> Reply to Message

Il giorno martedì 1 settembre 2015 19:33:06 UTC+2, eben.pe...@gmail.com ha scritto:
> Hi all,
>
> I'm looking for code examples of writing ODL structures to HDF5 tables that can be read in Python (2.7). Filenames are removed
>
> I've tried the following approach using the IDL coyote guide to create the file in IDL 7.1:
>
> pro write_idl_struct_to_hdf5,struct,dir_out,outfile
>
> file = dir_out+outfile+'.h5'
> fid = H5F_CREATE(file)
>
> datatype_id = H5T_IDL_CREATE(struct)
> dataspace_id=H5S_CREATE_SIMPLE(1) ; not so simple..
>
> dataset_id = H5D_CREATE(fid,'o',datatype_id,dataspace_id)
> H5D_WRITE, dataset_id, struct
>
> H5S_CLOSE, dataspace_id
> H5T_CLOSE, datatype_id
>
> H5F_CLOSE, fid
>
> print,'---Finished Writing IDL struct to HDF5---'
> end
>
>
> Python code and Error
> import pandas as pd
> dir_in=r'C:\D_Drive\Python\hdf5'
> infile=...
>
> mynewdf = pd.HDFStore(dir_in+'\\'+infile)
> print(mynewdf)
>
> print(mynewdf.select('o',columns='TASK_NAME'))
>
>
> resulting in:
> <class 'pandas.io.pytables.HDFStore'>
> File path: ...
> /o frame_table [0.0.0] (typ->generic,nrows->1,ncols->63,indexers->[index],dc->[TASK_CODE,TASK_NAME,EVENT,SYS_LOC_CODE,SUBFACILITY_CODE,AREA,RIVER_STREAM_MILE,MILES_FROM_NC_MOUTH,X_COORD,Y_COORD,COOR

D_TYPE_CODE,REFERENCE_POINT,ELEV,ELEV_UNIT,ELEV_COLLECT_METH
OD_CODE,ELEV_DATUM_CODE,WATER_COLUMN_DEPTH,WATER_COLUMN_DEPT
H_UNIT_CODE,SYS_SAMPLE_CODE,PARENT_SAMPLE_CODE,MATRIX_CODE,L
AB_MATRIX_CODE,SAMPLE_TYPE_CODE,SAMPLE_DATE,START_DEPTH,END_
DEPTH,DEPTH_UNIT,DEPTH_INTERVAL,SEDIMENT_NATIVE,GROUP_DESC,C
HEMICAL_NAME,CAS_RN,FRACTION,ANALYTIC_METHOD,LEACHATE_METHOD
,PREP_METHOD,RESULT_VALUE,LAB_QUALIFIERS,VALIDATOR_QUALIFIER
S,INTERPRETED_QUALIFIERS,DETECT_FLAG,TARGET_UNIT,REPORTABLE_
RESULT,VALIDATED_YN,VALIDATOR_REASON_CODE,BASIS,METHOD_DETEC
TION_LIMIT,REPORTING_DETECTION_LIMIT,DILUTION_FACTOR,LAB_NAM
E_CODE,LAB_SAMPLE_ID,ANL_SHORT_NAME,PARENT_LOC_CODE,DATA_REL
EASE_STATUS,DATA_ANALYSIS_HIERARCHY,RESULT_VALUE_MDL,RISK_SC
REEN_USABILITY,BASELINE_RA_USABILITY,CAS_RN_4PROUCL,CHEMICAL
_NAME_4PROUCL,TREATMENT_FLAG_4PROUCL,CHEMICAL_NAME_4LABEL,US
ABILITY_HIERARCHY])
>
> The error generated is
> ValueError: Wrong number of items passed 283774, placement implies 1
>
> The 283774 number is the number of rows in the structure and it seems that each tag in the dataframe has one row with 283774 elements.
>
> Is there some layout adjustment that's needed to read a structure correctly in HDF5?
>
> Eben
