Subject: Re: I have the craziest idea...
Posted by Martin Schultz on Fri, 16 Feb 2001 11:02:31 GMT

View Forum Message <> Reply to Message

tbowers wrote:

>
> I've written a widget program that I've tried to design as a general data
> import interface. What I mean by general is that it's a wrapper widget to
> import various types of data and present it back to the user in some kind of
> standardized format for easy use. E.g., I have some 3D gridded data in ASCII
> text files of one format, 2D scattered data in ASCII text files of another
> format, etc. etc. What I did was write the widget as a function call so all
> I have to do is call it and I always have access to ALL data files that I
> know how to read. And, more importantly, the data is passed back to me in a
> format that's consistent and easy to process and visualize in IDL.
>
> So, what do all of you think about the crazy idea of me putting this up on a
> server and all of us passing this around, adding code where we can so that
> we all have a centralized function that'll read many data formats? Seems to
> me that we all have our specialty, and when one of us wants to read in a new
> type we just come to this newsgroup for the answer anyway. I call it the
> ngDataImporter. I usually put my initials in front of the name of each
> function so I don't ever step on other functions. Since this would be a
> group effort by us all, I just fronted it with ng for newsgroup, the IDL
> newsgroup users.
>
> From the intro comments: [snipped][/color]

Excellent idea, but from glancing over it, I am not sure that the
structure is general and flexible enough. I started with a similar
effort based first on the netcdf format and its ideas of metadata.
Unfortunately I haven't gotten too far yet (because people don't let
me spend so much time on IDL programming here), but I do have a
working version of a netcdf reader, all object oriented, and with a
class hierarchy that would allow just what you propose. It is based on
a generic "basefile" object (of course with initials up front ;-) and
that object already does a lot of the work like opening the file,
closing it, searching for it, etc. The idea is then to inherit from
this object and implement only the file format dependent stuff and put
the data into the generic format that is used by the generic object -
and these are, yes: variables implemented as objects. The concept is
most powerful, but apparently relatively hard to realize, especially
when I work on it only one or two evenings in a period of 2-3 months.
;-(If you are interested, I can send you the current state of the
development (as I say: reading netcdf already works, and I also have
the beginnings of a graphical viewer that can take the variable
objects that are generated by the data reader and plot them as
contours etc. (only 2D data so far)).

A few more general issues in terms of such a project: If you really want a tool that almost everyone on this newsgroup can use, you need to make sure that you can access portions of the data in the file upon request. For example, here we often deal with model output on the order of 2GB (and it could be larger were there not the Linux filesystem limitations), and you certainly don't want to read the complete file before starting to work with these data. Now, this requires a program structure where you analyse the headers first, let the user select what he/she wants to read, and then go out and grab it. This idea is what prompted me to mirror the netcdf attribute structure in my variable objects, so that these objects contain all the information about the data (the "metadata"), and you could even reproduce a netcdf file (or convert it to an HDF file) with the information you have stored. And then it's only a small step to have a file converter in addition to the reader program, which, I am sure, many of us need almost as badly!

... glancing over this again, I must be out of my mind to give you such a hard time when you are proposing such a good thing. I guess, at this point, any tool that is easy to handle (so that even a scientists can use it ;-) and flexible enough, would already be most welcome.

Good luck,

Martin

--
 [[[[[[[[[
[[Dr. Martin Schultz   Max-Planck-Institut fuer Meteorologie   [[
[[Bundesstr. 55, 20146 Hamburg   [[
[[phone: +49 40 41173-308   [[
[[fax: +49 40 41173-298   [[
[[martin.schultz@dkrz.de   [[
 [[[[[[[[[