Subject: Re: IDLv5.4
Posted by Craig Markwardt on Sun, 31 Dec 2000 15:26:58 GMT
View Forum Message <> Reply to Message

davidf@dfanning.com (David Fanning) writes:

> Craig Markwardt (craigmnet@cow.physics.wisc.edu) writes:
>
>> On the political side, I think all RSI guaranteed was that SAVE files
>> would be *forward* compatible, so they haven't broken any promises.
>
> This is what I remember thinking was the case, too. But
> here is what the documentation for the SAVE command says:
>
> "Note also that save files containing routines may not be
> compatible between different versions of IDL, but that
> files containing data are always backwards-compatible."
>
> Do you suppose "backwards" really means "forward", which
> would mean...

After looking a little further, I think this may indeed be a bug in IDL, or at least a misimplementation of a design spec.

So far the only difference I found was that they have changed one record type to allow for 64-bit file offsets, as opposed to 32-bit. Presumably this is to allow people to save files larger than 4 gigabytes. However there are several tragic flaws in the implementation.

Every record in an IDL SAVE file of the past had a 16-byte header. Files are composed of a sequence of records and each one has a small header with the record type, and a pointer to the next record in the file. Sometimes this structure is called a "tagged" format (like TIFF is the Tagged Image File Format). This is an easy way for even old versions of IDL to read newer files: if they encounter a record type they don't recognize, they just skip to the next one. Since the header always had the same layout, this ensured compatibility.

In IDL 5.4, the header layout has changed incompatibly, *sort of*. The new record header has *20* bytes instead of 16 bytes, to accomodate 64-bit instead of 32-bit file pointers. However the IDL people seem to have recognized this might be incompatible, so they implemented a new record type (type 17) which signals a format change. Before record 17 appears in the file, the old 16-byte header layout is used; afterwards the new 20-byte layout is used.

This is tragic for several reasons:

* Record 17 appeared even in the short files that David and Jeff sent me. They could have easily only written this record for large datasets, ones that were known to exceed 4 gigabyte limit. That at least would have limited the incompatibility to large files that IDL 5.3 couldn't store anyway.

* Even more tragic is the fact that the old 16-byte layout appeared to have eight unused bytes. They could have allocated some of these existing bytes, and kept the header size unchanged that way. The current approach has variable-sized headers, which guarantees incompatibility.

* The original IDL SAVE protocol was so elegant because it was largely state-free. That is, one record really didn't depend on another. This allowed old versions of IDL to ignore records they didn't understand, with little or no consequences. This isn't true anymore.

With all this said, I think it is fair to say now that RSI had the opportunity to make this new format backward compatible, but they dropped the ball somehow. Perhaps they intended to maintain compatibility but didn't test enough to discover the problem. In that case there is hope it will be corrected in a new release. Another remote possibility is that they are pushing a new file format ala Microsoft to encourage upgrades. Let's hope this isn't the case.

Craig

--
 -- --------------
Craig B. Markwardt, Ph.D. EMAIL: craigmnet@cow.physics.wisc.edu
Astrophysics, IDL, Finance, Derivatives | Remove "net" for better response
 -- --------------