Here's a question. I have some strings with non-ascii characters, e.g.

If I convert these to bytes, I see they are multibyte encodings:

IDL> c=byte(a)
IDL> print,(c)
 208 154 209 131 208 186 209 131 209 136 208 186 208 176

and I can happily convert those numbers back to the original chars...

IDL> print,string(c)

Ok, now I have the same string, encoded (I believe) in UTF-8, from an html page:

IDL> b=" Кукуш& #x43A;а "

With some string splitting, I can convert these to bytes...

IDL> print,ch
  4 26 4 67 4 58 4 67 4 72 4 58 4 48

which unfortunately are not the same as those I had before. There's an almost simple relation, but I can't quite figure it out:

IDL> print,c-ch
 204 128 205 64 204 128 205 64 205 64 204 128 204 128

I was hoping one of these transparently named things might do the job, but no luck so far:

I18N_MULTIBYTETOUTF8

I18N_MULTIBYTETOWIDECHAR

I18N_UTF8TOMULTIBYTE

I18N_WIDECHARTOMULTIBYTE

Anyone care to enlighten me?

Greg