Subject: Please HELP!!!!!
Posted by phil on Thu, 17 Aug 1995 07:00:00 GMT
View Forum Message <> Reply to Message

Howdy,

I have a problem that is only loosely related to IDL that I hope one of you can help me out. I want to take the bytes from a 32-bit floating point number and determine the floating point value of it. I can do this in IDL by the following:

```
IDL>a = bytarr(4)
IDL>openw,1,'temp'
IDL>writeu,1,160.0
IDL>close,1
IDL>openr,1,'temp'
IDL>readu,1,a
IDL>close,1
IDL>print,float(a,0)
160.0
IDL>print,a
67 32 0 0
```

Now, here is the problem. I'm trying to get the same answer by hand and can't seem to do it. I have been doing all sorts of interations to get the sign, exponent and mantissa for this and nothing seems to work. I know this is simple, but I just seem to be missing something. Could one of you people out there humble me and show me what I am doing wrong.

TIA,

Phil
--
/** **********/
Phil Williams
Postdoctoral Researcher "One man gathers what
MRI Facility another man spills..."
The Ohio State University -The Grateful Dead
email: phil@peace.med.ohio-state.edu
URL: http://justice.med.ohio-state.edu:1525
/** **********/

Subject: Re: Please HELP!!!!!
Posted by Mike Mathews on Tue, 22 Aug 1995 07:00:00 GMT
View Forum Message <> Reply to Message

phil@peace.med.ohio-state.edu (Phil) wrote:
> Howdy,
>
> I have a problem that is only loosely related to IDL that I hope one
> of you can help me out. I want to take the bytes from a 32-bit
> floating point number and determine the floating point value of it. I
> can do this in IDL by the following:
>
> IDL>a = bytarr(4)
> IDL>openw,1,'temp'
> IDL>writeu,1,160.0
> IDL>close,1
> IDL>openr,1,'temp'
> IDL>readu,1,a
> IDL>close,1
> IDL>print,float(a,0)
> 160.0
> IDL>print,a
> 67 32 0 0
>
> Now, here is the problem. I'm trying to get the same answer by hand
> and can't seem to do it. I have been doing all sorts of interations
> to get the sign, exponent and mantissa for this and nothing seems to
> work. I know this is simple, but I just seem to be missing something.
> Could one of you people out there humble me and show me what I am
> doing wrong.
>

Try the following for the 32bits

s eeeeeeee fffffffffffffffffffff

msb lsb msb lsb

l/msb is least/most significant bit

if 0 < e < 255 then v = (-1)^s * [2^(e-127)] * (1.f)
if e = 0 and f <> 0 then v = (-1)^s * [2^(-126)] * (0.f)
if e = 0 and f = 0 then v = (-1)^s * 0
if e = 255 and f = 0 then v = (-1)^s * inf
if e = 255 and f <> 0 then v = NaN (not a number)

I haven't check your numbers but I hope this helps

Mike

---------- mailto:fskmjm@pukfsk.puk.ac.za ----------

Subject: Re: Please HELP!!!!!
Posted by pjclinch on Wed, 23 Aug 1995 07:00:00 GMT
View Forum Message <> Reply to Message

Phil (phil@peace.med.ohio-state.edu) wrote:
: Howdy,

: I have a problem that is only loosely related to IDL that I hope one
: of you can help me out. I want to take the bytes from a 32-bit
: floating point number and determine the floating point value of it.

: Now, here is the problem. I'm trying to get the same answer by hand
: and can't seem to do it. I have been doing all sorts of interations
: to get the sign, exponent and mantissa for this and nothing seems to
: work. I know this is simple, but I just seem to be missing something.
: Could one of you people out there humble me and show me what I am
: doing wrong.

First of all, the format will depend on your system. VAXen use a
unique floating point format, I think, but other than that it's probably
IEEE 753-1985, which is the main standard.

format for a 32 bit float is:

bit 31: sign bit, 1 -> negative, 0 -> positive
bits 30-23: exponent
bits 22-0: significand

the problem interpreting these is usually in the significand, which is,
not very intuitively at first sight, a binary fraction greater or equal
to 1 and less than 2. I think (but can't be totally sure offhand) that
all the significand bits set at 0 is equivalent to an actual value of 1,
while all set at 1 is not quite 2, with fractional points in between.

the other fun bit is the exponent, which you get by subtracting a bias
value of 127. This allows you to get negative exponents, so an actual
bit value of 130 in bits 30->23 would mean an exponent of 3 for the
actual number represented.

Have fun...,
Pete.
--
Peter Clinch Dundee University & Teaching Hospitals NHS Trust

Tel 44 1382 660111 x 3637 Medical Physics, Ninewells Hospital
Fax 44 1382 640177 Dundee DD1 9SY Scotland UK
net p.j.clinch@dundee.ac.uk http://www.dundee.ac.uk/MedPhys/