Subject: Re: error with sort
Posted by David L. Keller on Wed, 12 Jan 2000 08:00:00 GMT
View Forum Message <> Reply to Message

David Fanning wrote:

> R.Bauer (R.Bauer@fz-juelich.de) writes:
>
>> Dear David,
>>
>> that's all right, but why get I different results on different Operation
>> Systems?
>
> I'm reliably informed that you get different results because
> IDL uses the system supplied qsort() routine on each platform,
> and as they are different implementations, they are free to
> return different sub-orders for the identical elements.
>
> But, as I say, I can't see how it makes much difference,
> although I did appreciate Wayne Landsman's perspective.
>
> Cheers,
>
> David
>
> --
> David Fanning, Ph.D.
> Fanning Software Consulting
> Phone: 970-221-0438 E-Mail: davidf@dfanning.com
> Coyote's Guide to IDL Programming: http://www.dfanning.com/
> Toll-Free IDL Book Orders: 1-888-461-0155

Answer: When you have 'parallel' arrays. In my field, you might have pressure, temperature, humidity, wind speed. Say the pressure is nearly steady for much of your data. You sort all of the arrays based on the pressure. You then look at values of temperature, humidity, and wind. For subranges where the pressure has the same value, the other arrays have widely varying values. It is nice to have them come up in the same order for the same values of pressure.

For this to make a 'real' difference, you have to be doing something to the other arrays. For statistical work, I might be putting the arrays into 'bins'. Different ordering, different bins. Harder to debug.

Furthermore, the 'steady' pressure values might be in say chronological order. This arrangement might be meaningful in sampling the data. You might wish to sample 'steady' pressures every 3rd hour. You COULD do this if the pressure stayed in chronological order, would be harder otherwise.

And of course, if you are testing any kind of sorting or sampling or statistical algorithm of your own, a consistent sort is nice. Very annoying to run a statistical algorithm, and get different values with back-to-back runs of the same data, same algorithm.

-- Dave --
