Subject: Sort without loops Posted by lan Dean on Wed, 17 Aug 2005 09:08:07 GMT

```
View Forum Message <> Reply to Message
Hi All.
  I have a large string array (~100000 elements) that need sorting on two
fields within each string.
e.g. array=['F;100', 'ABC;6', 'DE;2', 'DE;10', 'DE;1']
Order required is a) sort items to left of ';' followed by b) sort items
numerically to right of ';'
This would produce:
  ABC:6 DE:1
                   DE:2 DE:10 F:100
A simple sort (sort(array)) procudes:
  ABC:6 DE:1
                   DE:10 DE:2 F:100
The only way I've found is to conmvert the RH part to I4.4 format within a
loop and search on the new values:
  for j=0, n elements(array)-1 do begin
     parts=strsplit(array[j], ';', /extract)
     RH=string(fix(parts[1]), format='(I4.4)')
     new[j]=parts[0]+RH
  endfor
  order=sort(new)
i.e
  new array is F;0100 ABC;0006
                                     DE;0002 DE;0010
                                                             DE;0001
which is then sorted correctly.
Is there a clever way of sorting on two fields like this without using a
loop. The above algorithm is faaaar slower than just using sort.
```

I hope I have made this as clear as mud.#

In expectation, lan

Subject: Re: Sort without loops
Posted by Ian Dean on Thu, 18 Aug 2005 13:28:09 GMT
View Forum Message <> Reply to Message

Thanks to you all it works well.

"lan Dean" <lan.d.dean@baesystems.com> wrote in message

```
news:4302fc84$1_1@glkas0286.greenlnk.net...
> Hi All.
     I have a large string array (~100000 elements) that need sorting on
two
  fields within each string.
  e.g. array=['F;100', 'ABC;6', 'DE;2', 'DE;10', 'DE;1']
>
> Order required is a) sort items to left of ';' followed by b) sort items
> numerically to right of ';'
> This would produce:
     ABC;6 DE;1 DE;2 DE;10 F;100
>
 A simple sort (sort(array)) procudes:
     ABC;6 DE;1 DE;10 DE;2 F;100
>
>
  The only way I've found is to conmvert the RH part to I4.4 format within a
  loop and search on the new values:
  .....
>
    for j=0, n_elements(array)-1 do begin
       parts=strsplit(array[j], ';', /extract)
>
       RH=string(fix(parts[1]), format='(I4.4)')
>
       new[j]=parts[0]+RH
>
     endfor
>
     order=sort(new)
>
> ...
> i.e
     new array is F;0100 ABC;0006 DE;0002 DE;0010
                                                              DE;0001
  which is then sorted correctly.
  Is there a clever way of sorting on two fields like this without using a
> loop. The above algorithm is faaaar slower than just using sort.
>
  I hope I have made this as clear as mud.#
>
  In expectation,
     lan
>
```

Subject: Re: Sort without loops
Posted by David Fanning on Thu, 18 Aug 2005 13:41:21 GMT
View Forum Message <> Reply to Message

Ian Dean writes:

> Thanks to you all it works well.

Did you happen to run a speed test on this, Ian? I have a new section on my Tips page called "The IDL Way". I generally don't put things on my page that I don't understand, but I'm making a few exceptions for JD's legerdemain. It is nice to put at the end of the article "By using the IDL Way I sped the algorithm up by a factor of 1,345,723."

Cheers,

David

P.S. If anyone has their favorite IDL Way story, I'm accepting all the help I can get. :-)

--

David Fanning, Ph.D.
Fanning Software Consulting, Inc.
Coyote's Guide to IDL Programming: http://www.dfanning.com/

Subject: Re: Sort without loops

Posted by peter.albert@gmx.de on Fri, 19 Aug 2005 08:01:23 GMT View Forum Message <> Reply to Message

Hi David,

using strmid and avoiding the for-loop increased the speed on my test dataset by roughly a factor of 10.

Cheers,

Peter

Subject: Re: Sort without loops
Posted by MarioIncandenza on Fri, 19 Aug 2005 14:57:30 GMT
View Forum Message <> Reply to Message

David,

I started using IDL about 14 months ago, and now I have a Ph.D. to show for it. Your help has been essential to this.

My "IDL Way" story has to do with resampling a 3d xyt grid to match a list of irregular samples in space and time. By using HIST_ND and creative array manipulation, I changed a procedure that took six+ hours

into one that takes 30 seconds. My largest bottleneck is now loading the data off the disk.

Cheers,

Edward Hyer.

Subject: Re: Sort without loops

Posted by David Fanning on Fri, 19 Aug 2005 15:05:40 GMT

View Forum Message <> Reply to Message

Ed Hyer writes:

- > I started using IDL about 14 months ago, and now I have a Ph.D. to show
- > for it. Your help has been essential to this.
- > My "IDL Way" story has to do with resampling a 3d xyt grid to match a
- > list of irregular samples in space and time. By using HIST_ND and
- > creative array manipulation, I changed a procedure that took six+ hours
- > into one that takes 30 seconds. My largest bottleneck is now loading
- > the data off the disk.

Whoa! Do you remember the details. I have a guy bugging me about this kind of thing right now!

I'll send you a template and you can write it up. I'm pretty sure you will be famous. :-)

Cheers,

David

David Fanning, Ph.D.

Fanning Software Consulting, Inc.

Coyote's Guide to IDL Programming: http://www.dfanning.com/