Subject: My final solution..thanks for your help! Posted by m.doyle on Tue, 13 Jan 2004 14:17:54 GMT

View Forum Message <> Reply to Message

Hello everyone, and many, many thanks for all your helpful suggestions.

I managed to get the runtime for this problem down to 35 seconds, from what I previously estimated was going to take about 4 days using for and if loops! Good old IDL!

I ended up combining many of the solutions posted previously, as follows;

My original files were in the format below:

```
> file1:
> 1954 12 31 23  90  11  4 366  0.00
>
> file2:
> 1954 12 31 23  2.80  2.10  2.20  95
```

With the intended result:

```
> 1954 12 31 23 90 11 4 366 0.00 2.80 2.10 2.20 95.21
```

I used Ben's suggestion and concatenated the first 4 columns of each file resulting in a "field ID" if you like:

```
> file1_ID= (file1[0,*]*1000000D) + (file1[1,*]*10000D) + (file1[2,*]*100D) + >file1[3,*] 
> file1_ID_final = round([file1_ID])
```

result: 1954123123

I then used the match() routine from the NASA library:

http://groups.google.co.uk/groups?selm=331C553A.41C67EA6%40a strosun.tn.cornell.edu&oe=UTF-8&output=gplain

This program allowed me to output 2 vectors of indices indicating matching pairs of "field ID's". These outputs were suba for file1 and subb for file2. For example, if suba[0] = 2 and subb[0] = 5, then file1_ID[2] EQ file2_ID[5].

I then concatenated the 2 files based on these indices:

```
> endresult = [file1(*,suba(*)), file2(4,subb(*)),file2(5,subb(*)), file2(6,subb(*)), file2(7,subb(*))]
```

> printf, 3, endresult, format = finalformat Once again, many thanks for all your helpful suggestions, Best wishes, Martin.. m.doyle@uea.ac.uk (Martin Doyle) wrote in message news:<d33d6a4b.0401080227.1a588e88@posting.google.com>... > Hello all, > > I really hope someone out there can help me with this....I am tearing > my hair out as my code is so slow! > I have 2 files of data (hourly met data) with one file containing one > set of parameters, and the other file containing another set of > parameters. What I am trying to do, is to match the data based on the > YY, MM, DD and HH values and then write BOTH sets of parameters to a > seperate file. For example; > > file1: > 1954 12 31 23 90 11 4 366 0.00 > > file2: > 1954 12 31 23 2.80 2.10 2.20 95.21 > > intended result: > 1954 12 31 23 90 11 4 366 0.00 2.80 2.10 2.20 > 95.21 > NOTE: Both files have no order to them, so a simple concatenation > won't work > I have written some code, but it is wrist slashing-ly slow!; > I read in each variable as a seperate array... > b=0L> REPEAT BEGIN > c=0L> REPEAT BEGIN

and output!

> If (year(b) EQ year2(c)) AND (month(b) EQ month2(c)) AND (day(b) EQ

```
> day2(c)) AND (hour(b) EQ hour2(c)) THEN BEGIN
>
> printf, 3, year(b), month(b), day(b), hour(b), winddir(b), windsp(b),$
> present(b), visib(b), mslpres(b), airt(c), dewt(c), wett(c), relh(c),$
  format = finalformat
> endif
> c=c+1
> ENDREP UNTIL c EQ lines2-1
> b=b+1
> ENDREP UNTIL b EQ lines1-1
  I'm sure there must be a better way than this.
>
 Please help me!
> Many thanks in advance, Martin..
```