
Subject: Re: comparing and concatenating arrays...please help!!

Posted by [wmconnolley](#) on Fri, 09 Jan 2004 09:50:37 GMT

[View Forum Message](#) <> [Reply to Message](#)

Martin Doyle <m.doyle@uea.ac.uk> wrote:

> My files are _huge_. File 1 is about 250000 lines and file2 about half
> that. which is why my code was taking about 4 days!

Here is my initial entry in the obfuscated solution using perl.
However, I think I misread you because in whats below I've assumed
that dates for each ob exist in each file. If you only want output
when both match, and throw away what doesn't have equiv's in both files,
then use the alternative last line.

```
#!/bin/perl
```

```
@f1=split("\n",`cat f1`);  
@f2=split("\n",`cat f2`);
```

```
for (@f1) { ($y,$m,$d,$h,$r)=(/((....) (..) (..) (..) (.*)/); $f1{"$y $m $d $h"}=$r };  
for (@f2) { ($y,$m,$d,$h,$r)=(/((....) (..) (..) (..) (.*)/); $f2{"$y $m $d $h"}=$r };
```

```
for $k (sort keys %f1) { print "$k $f1{$k} $f2{$k}\n" };
```

```
# Alternative output line
```

```
# for $k (sort keys %f1) { if ($f1{$k} and $f2{$k} ) { print "$k $f1{$k} $f2{$k}\n" } };
```

This should be fast (a few mins at most) if all the data fits into your memory at
once, as it ought to. You can write "reverse" in front of sort if you want it
backwards.

-W.

--

William M Connolley | wmc@bas.ac.uk | <http://www.antarctica.ac.uk/met/wmc/>
Climate Modeller, British Antarctic Survey | Disclaimer: I speak for myself
I'm a .signature virus! copy me into your .signature file & help me spread!
