Subject: Re: Algorithm for lat/lon searching
Posted by Gordon Sande on Fri, 18 Aug 2006 15:18:09 GMT
View Forum Message <> Reply to Message

On 2006-08-18 11:50:56 -0300, Paul van Delst <Paul.vanDelst@noaa.gov> said:

> Hello,
>
> I want to implement a global *land* surface emissivity database (as a
> LUT) into a radiative transfer code. Fir simplicity the database is
> simply gridded by lat/lon (land and sea). Due to memory limitations, I
> want to only keep the land gridboxes in my lookup table. Obviously,
> doing this complicates searching for the actual lat/lon element since
> they're no longer stored on a grid.
>
> What I'm looking for is a simple and/or quick method for searching a
> somewhat irregularly spaced database for particular points. In the IDL
> newsgroup there was recently a discussion above finding unique number
> pairs (lat->"high" portion of 64 bit int, lon->"low" portion) and I was
> thinking that would provide a searchable database. By converting the
> lat/lon pair to a unique number, e.g.
>
> JD Smith wrote:
> <IDL code follows>
>> epsilon=1.e-7 ; difference in degrees for equality
>> lat_lon = ulong64((lat+90.)/epsilon) + ishft(ulong64(lon/epsilon),32)
>
> the resultant lat_lon array being simple to search.
>
> An additional problem is that, since this data will be used for
> satellite data assimilation and satellites tend to scan "diagonally"
> across lat/lon, adjacent/close-by *geographical* grid elements will be
> accessed and it's not clear to me that the above lat/lon orgainisation
> will put elements separated by a short physical distance anywhere near
> each other in the lat_lon array.
>
> I will be playing with and testing this over the coming days, but I
> wanted to pick the brains of folks out there in advance.
>
> Thanks for any suggestions/advice,
>
> cheers,
>
> paulv
>
> p.s. Since the final code needs to be Fortran95, I set followups to
> comp.lang.fortran

Welcome to multiple key searching.

The granddaddy technique goes by the name of kd-trees. As in K Dimensional trees. When k=2 they are called quad trees. When k=3, oct trees. When ...

The problem is also called nearest neighbour searching with many geographers using natural neighbours as a variant. Also called associative searching or even content directed searching.

This has a large literature with much of the terminology very graph theoretic. Triangulation is an important problem for many so there is much discussion of that. Regular spatial arrangements are called crystals which is a whole field in physics. Geographic databases are pretty common.

If you like combinatorics there are a variety of space filling curves that can be used to keep things which are close in both (real) indices close in their single (referencing) index. The problem you are asking about.

And here you thought it was going to be a simple answer to a simple question!

Isn't this the sort of thing that outfits like NOAA are supposed to be experts in? Unfair question as you have to cross speciality boundaries and wade through arcane terminology. But seriously, there should be folks around there who know this sort of stuff.
