Transactions in GIS, 2004, 8(4): 401–405

Guest Editorial

Search Procedures for Geographers, By Geographers and Using Geographers

1 Introduction

This article attempts to highlight some of the methodological approaches, by now well established within the GIScience community, that may be harnessed to address a specific research question. The illustrative task, if you will, is to find the location and altitude of the highest point in the Cairngorms – a range of hills in the Scottish Highlands. The answer is in fact *Ben Macdui* which, at 1,309 m (4,296 ft), is the second highest mountain in Scotland, but that is entirely irrelevant to what follows.

Many problems of analysis that geographers regularly encounter, including classification, wayfinding, clustering, prediction and selection of the next desirable location for the GIScience conference can be re-stated in terms of a *search*. By doing so, the problems become amenable to an array of statistical, mathematical and machine learning tools, that can, if nothing else, add significantly to the trendiness of articles we write (McNoleg 1996). The search is typically of all the possible configurations that some method could take, to find the one that produces the outcome that best supports our own point of view, or failing that, the one that best supports the data we have.

To illustrate the range of emerging search and analysis methods now available to address the problem of (say) finding the highest mountain we use here an analogy that may be only distantly familiar to many GIScientists – that of the field excursion.

2 Computer Science: Gradient Search – Steepest Ascent

Drop 1,000 blindfolded geographers over the Cairngorms from the back of an airplane flying at a height of 20,000 feet (you may want to provide parachutes for those whose specialty you like!) This will ensure a random point pattern of geographers (or ex-geographers) on the landscape at the start of the search procedure. Upon landing, each geographer must walk in the direction of steepest ascent (the one that takes the most effort) until they can go no further without going downhill. Then they should stop, take off their blindfold, record the position and height readings from their GPS, eat their sandwiches and walk back to base camp.

These unsophisticated search rules will lead to problems with local minima (a geographer reaches a local peak from which there is no uphill route without first going downhill). Hence, it is entirely possible that the highest point will be unreachable using these rules, so that an optimal solution cannot be guaranteed. To increase the chances

© Blackwell Publishing Ltd. 2004. 9600 Garsington Road, Oxford OX4 2DQ, UK and
350 Main Street, Malden, MA 02148, USA.

of finding the desired solution, it is necessary to use a large population of geographers. Further improvements are possible by enhancing the algorithm, for example by allowing the geographers to see some way into the distance so they may plan their route more carefully. The easiest and kindest way to provide limited optical cues that would mimic a search radius is to conduct the experiment at night, with each geographer issued with a weak flashlight or torch (another less humane option is to make the geographers use GIS intensively until their long range vision deteriorates sufficiently).

3 Geocomputation: Genetic Algorithm

As each geographer travels, searching for the highest point, they make maps of where they have been, carefully recording the position and heights of all hills they encounter. At selected intervals, any geographers with bad maps are killed (there, now you wish that you had paid more attention in cartography class)[1]. Those left alive then rip up their maps into sixteen square quadrants and make a pile of the pieces (in Morton order, just for fun). Then they need to find a partner, and working with their partner they assemble two new maps from their combined pieces by drawing an equal number of quadrants from each pile, but choosing which pile at random. Each geographer now completely forgets what they knew before (experience suggests this should not be difficult), and begins again with their new map. Or if they prefer, they can procreate two new geographers, then die and pass the maps on to their offspring – the choice is theirs, but it takes a lot longer this way. With any luck (luck is a necessary condition of all heuristic research) this approach will not only find the highest point but also other significant peaks in the region.

Both genetic algorithms and gradient ascent have the advantage that a relatively small number of geographers can search a much larger space, but they cannot guarantee to find the highest point definitively. For example, it is possible that the highest point might be missed because of: (1) poor initial randomization of geographers upon the landscape, (2) local perturbations in the surface morphology trapping the geographers in some local maximum, (3) inclement weather (McNoleg 1996), and (4) hill walking accidents.

By contrast, the following two approaches are either very geographer-intensive, or more time-intensive, but they do guarantee to find the best solution.

4 Urban Simulation: Cellular Automata

Tessellate the entire Cairngorms region into 10 m by 10 m squares (cells), and place a large pole securely in the center of each cell. Select a geographer, then using a chain of length 4 m; fasten one end to the leg of that geographer, and the other to the pole. Repeat this procedure for all cells. Even though geographers generally have two legs, do not attempt to use the same geographer for two cells[2]. Geographers then essentially play Chinese Whispers, passing messages to their immediate neighbors about high places identified by themselves or by others 'upstream'. Message passing between the tethered academics is synchronized by a coordinated network of fog-horns or air-raid sirens, sending out regular blasts at one minute intervals. Ensure all geographers are well-protected against the elements since the experiment may fail completely if just a single geographer succumbs to exposure.

© Blackwell Publishing Ltd. 2004

This method is extremely geographer-intensive, with the number of geographers required rising as the square of the length of one side of the study region.

5 Exhaustive search

Using this approach a single geographer is tasked with traversing the entire region and recording heights as they go. A strict route must be designed and adhered to in order to ensure all possible locations are surveyed. A Peano curve has exactly the right properties for such a route, assuming that the geographer will want to minimize the distance to be traveled, but the path may appear rather amusing to other hikers encountered along the way[3]. The search process will quite literally exhaust the geographer, hence its name.

This method is time intensive and degenerates into a random walk if the geographer is allowed to choose the directions to be explored, or if they drink a little too much emergency brandy from their hipflask. This, in turn, introduces the possibility of missing the highest hill and settling for a local maximum, or even a ditch, instead.

6 Spatial Information Theory: Qualitative Spatial Reasoning

Design an algebra that is far too complicated for anybody else to understand. Include within it rough-fuzzy qualifiers for constructs such as '*Not-quite the smallest mountain I've seen in the last hour*', '*Over the hill and far away*', and '*Oh dear, I appear to be quite lost*'. Train geographers to use this language to communicate their findings as they are allowed to roam freely over the landscape.

7 Remote Sensing: Digital Image Processing

Download the latest imagery from the Shuttle LIDAR mission flown in 2000 (http://www.nasa.gov/). Write an application to automatically segment out the relevant orbital swaths, then stitch together the required coverage from the different passes obtained via a system of quadratic match-filters running on a Beowulf Linux cluster. Geo-register to 1,000 ground control points using the bi-cubic convolution method. Correct for atmospheric distortion, platform wobble, sensor drift, sun spot activity and UFOs. Run through a neighborhood smoothing filter to remove localized high-frequency anomalies, then use a multi-outflow distributed surface hydrological algorithm to verify the surface is feasible, connected and contains no obvious local minima (pits). Using a trace-back of the algorithm run-time stack, find the set of pixels with no upslope neighbors and from these select the one with the highest Z value; this represents the highest point. But why stop there when there is still so much science to do . . . Next, conduct a sophisticated error assessment to check that no other pixels might be candidates for the highest point, bearing in mind the error of the original data and the accumulated uncertainty derived from the analysis techniques used thus far. Find that in fact all pixels are candidates, so abandon error assessment and instead compute a fuzzy viewshed analysis from the highest point located and render the scene in 3D using an immersive visualization environment such as a Cave.

© Blackwell Publishing Ltd. 2004

8 Hydrological Engineering (practical)

Assemble a large construction crew and travel to the Cairngorms. Build a very high and thick impermeable barrier around the entire region, and wait until the inside fills with water (earlier research into Scottish rainfall patterns (McNoleg 1996) suggests that thus should not take long) and only a single peak remains unsubmerged (a practical instance of the invasion-percolation approach). Build a cantilever bridge from the side to this peak and plant a flag on the top. Re-engineer the barrier to include a hydroelectric power station, and forget to conduct an environmental impact assessment; consequently, silt up the Cairngorms until they are all the same height and the original question becomes moot.

9 Pragmatics

Use an experienced field guide equipped with a GPS receiver or type the request into Google [keywords: Scotland, Cairngorms, Highest Mountain].

10 Cartography

Look up the answer on a topographic map.

11 Conclusions

Perhaps the most important characteristic of GIScience illustrated here is that it encourages almost infinite complexity, supports new theory creation and experimentation to obtain results that could easily be known by browsing an existing resource or conducting a small amount of field work. Why is it that so much of our research attempts to replace common sense with numerous novel analysis methods, three graduate students, two research grants, four smart new computers and several trips to professional meetings in exotic locations? Oh, right!

Oleg McNoleg
Brigadoon University of Longitudinal Learning
School of Holistic Information Technology
Noplace, Neverland

Author's note

No geographers were physically harmed during the research described in this paper, though at least one was mentally challenged.

Endnotes

1. Actually, by bad maps here, we really mean maps that contain no significant hills.

© Blackwell Publishing Ltd. 2004

2. In the spirit of this article, readers are encouraged to think of their own caveat for use in this sentence: "unless____________________________________".
3. Depending on your perspective, not looking silly may be countered by sporting very expensive and totally over-engineered equipment such as a garish 100 liter rucksack, gortex under-garments, flashy wrist altimeter, and so forth. Your standard plastic mac and grubby, cotton 'SDH 1994' conference bag will not cut it with most of the people you will meet.

Reference

McNoleg O 1996 The integration of GIS, remote sensing, expert systems and adaptive co-kriging for environmental habitat modelling of the Highland Haggis using object oriented, fuzzy logic and neural network techniques. *Computers and Geosciences* 22: 585–8

© Blackwell Publishing Ltd. 2004