Automated Cartography in a Bush of Ghosts

William A Mackaness
Institute of Geography
School of GeoSciences
The University of Edinburgh
Drummond St, Edinburgh EH9 8XP
William.mackaness@ed.ac.uk


Abstract
Research into automated cartography has been going on for about fifty years and it is interesting to see how the techniques and methodologies have changed during that time. That richness in variety, and the exciting breadth of cognate disciplines drawn into this area of research reflects 1) the challenges of modelling the art and science of cartography, and 2) the changing way in which we use maps. Some of our ideas have foundered but this has always led to new innovations and the lateral application of new methodologies. This talk will review the past as a way of explaining current thinking, and argue that cartography is has much to do with reasoning about space and that 'design is all about making sense of things'. Perhaps these should be the goals of automated cartography?

1.0 Automated Cartography in a Bush of Ghosts
The title of my talk is inspired by the book of Amos Tutuola from 1954 'My Life in a Bush of Ghosts', in which he writes of an underworld odyssey, in which an eight-year-old boy, abandoned during a slave raid, flees into the bush, "a place of ghosts and spirits". In case you nod off during my presentation, I begin with a three point central message:

- I think we've built our mapping systems on sand - that sand is the paper map.
- I don't think we have properly understood what automated cartography is trying to achieve.
- There are however lots of exciting research activities that are getting into the soft underbelly of the art and science of cartography.

I want to talk about efforts to automate cartography. Before I start, I want to begin by saying that in discussing ideas of automated cartography and in particular generalisation, I feel it pertinent to mention where my current research interests lie. The focus of my research is on the derivation of very small scale mapping automatically from very large scale mapping (Figure 1).
The idea is that we record only at the fine detail, and taking advantage of the
interdependence of geographic phenomenon we can somehow automatically derive higher
order phenomenon. I am therefore interested in modelling large changes in the level of detail
with which we portray geographic information. My interests lie in modelling what the
wonderful Jean Claude Muller referred to as ‘conceptual cusps’ in map generalisation - which
are points in the representational continuum where the role of the map and the way in which
phenomena are represented - fundamentally changes. And I think trying to find answers to
this type of problem will help shed light on problems of small changes in scale/where there
are relatively small changes in the level of detail (Figure 2).

Given the complexity of this task, and the poor skills of the user, it is assumed that the
human is largely absent from the process. In terms of an automated solution, what is
required is an autonomous self evaluating system able to create maps in digital or analog form
that vary widely both in scale and thematic content.

2.0 Failure
I want to try and explain why attempts at automated cartography (as defined by my research
goal) have by and large, failed. I want to use those explanations to reason about what I feel
we should be doing in order to move on from where we currently are. For some, failure is an
emotive term,- but there is no need for any gnashing of teeth. On the contrary, it is often
through failure that we learn most. So why do I have this feeling? My first observation is
that cartographic methodologies have been poorly incorporated within commercial GIS. It
seems to me that even after 40 years of development, most GIS systems deal with the
cartographic component as some sort of 'end process' via simple data reduction algorithms,
colour ramps, symbol defaults, and choropleth mapping. (Choropleth mapping being a fine
example of a digital fossil - that is to say an outdated way of mapping sadly made permanent in every GIS system on this earth). The incarceration of these inadequate and anachronistic ways of doing cartography within GIS systems to me signals our failure to convince the community of the importance of cartography and a failure to innovate cartographic science.

I think we have also failed to design interfaces and interaction models that make it easy for users to specify their map requirements; we’ve not ‘google mapped’ cartography; Cartography has lost its placeholder in GIS - becoming somewhat marginalised by the linguistic intoxication associated with ‘intelligent visualisation’ and complex interaction models that afford real time, high dimensional fly throughs in both space and time. I think the problems have arisen for two main reasons:
- The first is that cartography is a poorly understood art and science; and when it came to automation we failed to develop methodologies that captured the process of design, (the idea of creating, and evaluating different cartographic solutions).
- secondly we failed to anticipate the disruptive nature of technology and the changing context in the use of maps; maps as interface, maps as product.
Let me elaborate a little with a brief trip down memory lane.

3.0 The Roots
The challenges of automated cartography were clear from the very beginning of the digital era when efforts were first made to model its various components. It was clear that cartographic solutions of any form, required among other things an understanding of scale, of aesthetics, as well as an understanding of geography (Figure 3) and all of this building upon a long tradition of paper map making extended back through time.

![Figure 3: Requirements for cartographic design.](image)

Attempts at automation quickly revealed cartography to be both an art and a science - indeed efforts to model that artistic component continues to attract a great deal of research effort to this day. It became clear that trying to capture the subtleties of the draughting process was going to be hugely challenging; I quote from the International Yearbook of Cartography from a 1965 article entitled the records of the discussion on the papers of
Tabler Bickmore and Boyle: 'Imagine a gorge with a river and a road and a railway. First we plot the river, then we display the road. The railway is displaced further and finally the contours are moved. This presents a very difficult problem for the machine to solve.'

But undaunted by the absence of techniques for handling such problems, and without really understanding what the paper map quite meant, massive programmes were undertaken to digitise paper maps of various scales. The paper map was a wonderfully constrained thing - thankfully. It was a scale dependent model - a particular abstraction of reality, it was a characterisation, a compromise, an interpretation, a cartoon of the real world. These digitising programmes chose to ignore this fact, capturing in the process all the nuances, tricks and 'lies' of the human cartographer in an instant. In a very real sense the ghost was in the machine!

I don't think we ever answered the question as to how precise these digital products were - and we failed to acknowledge that they reflected a compromise of design. Once in the database the cartographic caveats were discarded and attention focused on truth - truth in terms of locational precision and topological correctness. And listening to some of the papers and discussions of this conference, I think it fair to say that we've been panel beating these datasets ever since, trying to remove the lies of the cartographer. I remember Scott Morehouse explaining to me the concept of software fatigue - in which each additional software work around made it harder and harder to maintain and understand the behaviour of a system. By analogy what NMAs have is data fatigue.

The paper map became central to the whole data capture process. A map centric view heavily influenced our conceptual models of what GIS should be. We borrowed heavily from cartographic paradigms that persist (unhelpfully) to this day. Some examples of this being our 'layered' based approach to modelling geographies, our 'points/lines and areas view of the world and scale based datasets. The overriding objective of data capture smothered debate on a number of key issues: The cartographer had been the gate keeper - How were they to be incorporated into the system? How might we model the scale dependent nature of geographic phenomenon? How might the technology affect how we use maps? What were the links between the database and the map?

As the art and science of cartography drove into the oncoming digital juggernaut, we began to try and find answers to these questions. Firstly the science of map generalisation was born, and secondly a view emerged about how best to connect together a variety of technologies - from the capture process at one end, through to the interaction and display at the other. Before I go any further it is appropriate to give a brief definition of the term generalisation - since the remainder of my talk will focus around this term. The definition by Rob Weibel will do fine: 'a process of producing maps at coarser levels of detail, while retaining essential characteristics of underlying geographic information.'

Typically we talk about a range of methods used to achieve this - some relate to selection and reclassification of the data (this is termed model generalisation), others relate to giving emphasis and clarity to the data via exaggeration, displacement, and aggregation (this is termed cartographic generalisation). As for that connected view: the mantra has become: 'capture once, use many times'. The idea is of a database at the very finest level of detail: a
single maintenance point, one that could be continually updated via a process of automatic change detection...via data captured from remote sensing platforms. Attached to this was a set of generalisation methods that would enable automatic derivation of any map product (digital or analog) in any theme, at any level of detail. The idea of the ephemeral map - produced in an instant and then thrown away, was here to stay (Figure 4).

**Figure 4: The essential ingredients of the ‘capture once - use many times’ model.**

This ambition has not changed much over time—LIDAR technology and stereoscopic imagery makes it ever more likely that in the near future, capture will include the third dimension. But I should mention that this is a contested view - there is a feeling that there are too many ghosts in the machine, and various approaches are being debated that reflect a more pragmatic approach.

There is a new term that is being banded around - that of 'Geographic Exploration Systems' - the best example of GES apparently being 'google map and google earth. What is worth emphasising about GES is that it is a system that enables the user to move seamlessly between different levels of detail, exploring different datasets.

4.0 The Vision

*Cartography and Decision Making*

So, how close are we to this idealised view of capture once, use many? From early on it was clear that design was a complex decision making process and thus attempts were made to model some of the tasks performed by the human using artificial intelligence techniques. Rules were derived via a process of protocol analysis (which involved asking the cartographer at each stage why they did something, what information they used) and rules were drawn from inspection of map specifications. Three problems arose when attempting this:

- Firstly, Expert systems technology proved unwieldy for modelling a process that seemed packed with exceptions;
- Secondly, Cartographers drew heavily from trial and error and found it hard to articulate the decisions that they made; they always wanted to articulate their ideas via examples - and examples are hard to formalise into rules);
- Thirdly, and perhaps most importantly, was the thought that perhaps the human cartographer was not the best person to ask. I think I can articulate this idea best via an example.

Imagine a person travels from A to B by horse. An inventor decides to design a mechanical device by which they can achieve the same task, only more quickly. In order to do this they study the movement of the horse and by a process of reverse engineering, design a mechanical device that mimics that movement. This is precisely something that the inventor Rygg did (and patented) in 1893 (Figure 5).

![Figure 5: The Mechanical Horse (patented by Rygg in 1893).](image)

Nowadays we have many mechanical devices for getting from A to B - cars, helicopters, hovercraft, but somehow the mechanical horse never caught on. I don't want to overwork the metaphor but I simply ask, are we making the same mistake when we observe the human cartographer at work? How useful is it to know the human way of doing things when it comes to designing automated solutions?
The human operates within all sorts of constraints and opportunities. But these constraints and opportunities are not the same in automated environments. Carroll and Kellogg talk of the cognitive task artefact - the idea is simple - that the introduction of technology fundamentally changes the way we go about doing a task. Type writers and word processors are a very good example of this. Originally handwritten documents were sent to a typing pool, and the typed document was edited and then labouriously completely retyped. The development of word processors fundamentally changed how people designed and prepared manuscripts. The flexibility afforded by the technology did away with the notion of the typing pool. There is evidence to suggest that cartography has not taken advantage of technology in the same way - indeed I sometimes see cartographic techniques ossified within technology, rather than adopting and adapting new techniques and approaches. Three other issues impeded developments in automated cartography: Problem subdivision; Over simplistic models; Evaluation.

Subdivision of the problem
Firstly our paradigms of problem solving were inappropriate. When grappling with complex concepts, the strategy within computer science, has been to subdivide the problem into manageable chunks. This surely has been the curse of automated cartography? In breaking down the problem we lost context, we overly focused on the easy bits, we failed to model the interdependence of the myriad of generalisation methods that were created. We disappeared down holes - exploring bite sized problems that have little relevance to design. Research on text placement and Line simplification algorithms being prime examples of such false horizons.

Overly simplistic models
An issue that follows on from this stepwise refinement of the problem relates to the data model. The data models proposed were inadequate for the tasks; Our endeavours were constrained by simplistic Cartesian models - where a map was considered to be made up of lines, points and polygons. The poverty of these models highlighted the importance of context in giving meaning to a map, and so attention focused on the database - we began to ask what level of sophistication was required in the data model, in order to support the cartographic design process. (another topic I want to return to a little later).

Evaluation
A further challenge in the design of autonomous solutions was in the development of evaluation techniques. It appeared that automated environments required the equivalent of eyes with which to design and assess solutions. The task of evaluating the quality of output from such systems proved very problematic, - how do we know 'what is a good generalisation solution?' Just what we are trying to do when we generalise a map - how do we start from the fine, and work to the coarse whilst taking into account the user's needs, and their level of knowledge.

Much effort has been devoted to developing methodologies that can answer this question. The assumption is that if a generalised form is derived from a more detailed one, there must be a basis from which we can compare the two. A second assumption is that a good solution is one where there is as little change between the two, in the location and shape of the phenomenon, and as little change as possible in the topology between objects being mapped.
and various cartometric and shape measures have been devised to measure 1) changes in content, - 2) changes in topology, and 3) of location and of shape of the mapped objects.

My thesis is that when it comes to large changes in the level of detail, I think both the assumptions and the methodology of this approach are flawed. I would like to argue that conveying ideas with precision and accuracy does not translate into a requirement for precision in space (metric nor topological). On the contrary, I would argue that topology necessarily changes. Indeed that many properties of the map change. And that the priority of characteristics that we give emphasis to changes with scale. Various design constraints become much more relaxed as we change scale.

A map necessarily reflects many compromises. Monmonier said that when you make maps you have to lie to tell the truth - and there is much truth in this statement! We necessarily lie when we answer geographic questions: We know spatial query to be scale and context dependent - the generalised statements that we make about the world necessarily put to one side both detail, and exceptions. Hence the wry comment that 'Generalisations are dangerous and so is this one'. I would argue that in attempting to show something different about the world, that generalisation is necessarily dangerous!

In generalisation, I do not think the goals should be to retain precision, nor preserve shape, nor minimise changes in topology. These might be useful constraints, but they are not valid goals.

I would prefer to see a generalisation methodology that has the intent of deliberately making change. A methodology that is all about revealing different patterns and qualities among the features represented, and is a methodology that is all about revealing the meaning inherent among the phenomenon being represented. To me, this is what cartography is all about - giving meaning to things. It connects with a line of thought argued by Klaus Krippendorff Professor of Communication at the University of Pennsylvania, when he said that 'design is about making sense (of things)'. Cartography has everything to do with design.

And if we agree that meaning is intrinsic to cartography - then perhaps we can begin to think differently about the underlying structures necessary to support automated cartography. Going beyond phenomena centred databases, to data models that support the process of reasoning about space. And here I want to tell another story - not about horses but about pebbles and beaches.

5.0 Pebbles on a Beach
For a long time, I have been doing battle with a quote attributed to Minsky (one of the fathers of AI), who is quoted as saying: 'you cannot tell you are on an island by looking at the pebbles on the beach' (Figure6). I shall ignore an argument from a geomorphologist friend of mine, who said he thought he could.
But if we jump in a hot air balloon and sail aloft, we see a different property of the pebbles - their connectedness as a perimeter ring providing the answer to our question (Figure 7).

What we see are two fundamental different views on the world (we’ve crossed one of Muller’s conceptual cusps). In each view, we see different properties, different patterns and associations - that enable us to answer different questions. It is not the case that any one scale of observation contains less or more information: much more the case that they contain different, albeit related/and connected pieces of information.

In today’s GIS world, our database stores the pebbles (it has no concept of ‘Island’), but the idea is that (given the interdependent nature of the pebbles and the island) generalisation processes can be used to ‘create’ the island. Herein lies the conundrum – how do we develop
generalisation techniques that abstract out from the data source (the pebbles), and reveal different properties of the phenomenon sufficient that we can identify and then display something that is called an 'island'?

If we go back to my earlier criticisms of generalisation methodologies, I am arguing that this change from pebble to island is so significant, that there is little value in trying to measure and preserve the metric qualities of the pebbles. I don't believe we can derive meaningful strategies of evaluation based on analysis of the pebble.

And from a data modelling side of things, I have not been able to fathom a way by which we get from the pebbles to the Island via a bottom up approach. (..that is to say that a series of observations made about the pebbles will lead us to say 'this is an Island'). We can only get to the island if have some pre conceived notion of what an Island is. I can therefore see a solution if we adopt more of a top down approach. An approach in which we understand what an Island is, and how to search for them within a database of pebbles. In order to get from our pebbles to the Island, I think we need the following ingredients: 1) A prototypical view of what an island is, 2) a way of identifying properties inherent among those pebbles and 3) some hierarchical structure that enables us to store the newly found relationship between our pebbles and our Island. Let me talk about each of these in turn.

Prototypical definitions of things
There are many ways by which we can define 'Island' - properties such as size, remoteness, and that they are disjoint from other things (surrounded by water). Beyond these physical properties we can also define our island functionally or in terms of thematic attributes: that they are used to bury treasure, as safe anchorage or a nice place to holiday. Some of these properties are modelled more easily than others. In the interests of computational efficiency, we want a minimum set of measures sufficient to accurately identify 'islands'. Given that our database contains pebbles, it would be useful to define our Island based on this class of object. An Island could be defined as 'a set of touching/connected pebbles that encircle a small land mass.'

Methods for extracting properties
My second ingredient (or requirement) is that we have 'a way of extracting properties among these pebbles'. We use statistical methods to measure variation in individual size, or pattern recognition techniques to extract gestaltic properties such as their distribution patterns or topological properties (how are they connected). Again, in terms of computational effort, we want a minimum set of methods sufficient to enable us to discern whether or not, these pebbles do in fact constitute an Island. If we have a prototypical definition of island, a hierarchical structure that connects different scale dependent phenomenon, and ways of searching the database for these different prototypical forms, then I think we will get from our pebbles to the island!

At the beginning of my talk, I suggested that there are some exciting areas of research that I feel will provide insight into solutions to this problem. In particular I want to give emphasis to two areas of research that I have been reading about. The first is Scale Space - the second is mereology. I begin with a beginners guide to Scale Space.
6.0 Scale space
Scale space - a term introduced by Witkin in 1983 has its roots in image processing. It is a formal theory of how we might describe image structures at different scales. It shares many of the ambitions with cartography though it works in a more constrained environment of raster imagery (the sort of view a robot might have on the world). The objective is the suppression of fine scale noise, and exaggeration of characteristics defining the typicality of objects in the image - signal suppression in search of deep structures - something that Witkin and Tenenbaum intriguingly refer to as 'scale space smoothing'. Figure 8 is an example of an attempt to illicit the structures from the photograph through automated techniques.

Figure 8: Signal suppression in search of deep structures.

The idea is to try and find the best scales for handling given problems and seeks to provide a framework for modelling image structure at multiple scales. Linberg talks about 'trying to link filters of different types and at different scales', and goes on to argue that the 'scale parameter is just as important as the structure'

I like this language, and I like the ideas - for me it has very strong resonance with the goal of generalisation and connects with ideas of a theory of scale - a model by which we can link together scale dependent phenomena - linking our pebbles to the Island.

Primal Sketch
Indeed we come full circle at this point since interestingly Witkin refers to Marr's vaguely suggested idea of the 'primal sketch'. The scale-space primal sketch defines a hierarchical data structure that links together different scale dependent phenomenon. The idea is that the primal sketch can be used to guide later stage processes in scale space smoothing.

7.0 Mereology
I think we can improve on this primal sketch by including ideas from research in mereology. Mereology is a research field concerned with formally defining the relationships between parts and their respective wholes. By defining objects in terms of being 'part of' something else, we can describe the parthood structures among different phenomenon. In essence it is a classification system that enables us to define the connections or functional associations between different things. For example from cells → twig → branch → tree → forest or we might functionally define all the components that define 'school' - that the play area, the staff parking area, the gym and the teaching spaces are all 'part of' school.
The other classification system we use is taxonomic - in which we classify things within the same class (home, business office, shop - all examples of buildings). As we alter the level of detail, we revert to higher orders in the classification (less distinction between different types of building).

Cartographers take advantage of our conceptual understandings of space and our tendency to organise the things we see within classification structures or frameworks that are both taxonomic and partonomic. Cartographers take advantage of such structures in the design of maps and so my thesis is that automated environments need equivalent structures.

![London](image)

Figure 9: A dot with the word London next to it.

So when you look at Figure 9, we immediately understand the significance of that dot and all that it constitutes. The meaning we give to that dot is one of ‘city-ness’ and what that city constitutes and means. I would therefore argue that Marr’s Primal Sketch needs to have an equivalent sophistication. This seems logical to me since cartography is very much about visualising these different taxonomic and partonomic relationships. From an understanding of these relationships, the user is able to make inference about the interdependence of geographic phenomenon.

The existence of partonomic and taxonomic structures within the primal sketch is what would enable us to separate our rider from the horse, and to understand the connections between the two. It seems logical to me, that the modelling that underpins automated cartography needs to support both these types of classification. As Tversky argues ‘these partonomic structures provide a ‘conceptual skeleton’ linking appearance and function.’ So we can try and bring these things together and form an idealised view of a system necessary to support automated cartography (Figure 10) (and to get from our pebbles to the island, from our island to the archipelago).
Some of you may be sitting there feeling that this talk is more about data modelling than it is about ideas of automated cartography. My response to this anticipated criticism, is that every time I have tried to find solutions to the problems of automated cartography I have found myself having to deal with issues of modelling design, and of modelling meaning. There appear to be many dependencies that connect the underlying model with methods of representation. It seems that much of the cartographic process is dependent upon context and meaning – taking into account the associations implicit in the juxtaposition of the phenomenon being represented. As bizarre as it may sound, I end up feeling that cartography has very little to do with drawing! Indeed I am left wondering if more progress could have been made in automated cartography if we had not viewed it as a drawing exercise, and if we had not viewed it as an exercise in trying to automate the horse!

8.0 Conclusion

The importance of scale

It is time to conclude. Research in this field is very important - objects in the world only exist and make sense over limited range of scales. There is a tight weave between scale, query and meaning, and an intrinsic link between analysis, representation and exploration. “All geographical processes are imbued with scale” wrote Taylor in 2004. To that end we need to continue our efforts to develop a theory of scale, borrowing from other paradigms such as scale space, and mereology. We need to see stronger connection with research into semantic reference systems that attempt to model this link between meaning and scale.

The Goal

I think we need to have clear in our minds just what it is we are trying to achieve when we generalise a map. It seems generalisation methods lie along a continuum. At one end are methods that make subtle artistic changes; at the other are those that enable us to transition Muller’s conceptual cusps. Where it remains unclear to me is how these different methods fit and work together in the overall process of design.

New Ideas about maps

I think it important that we continue to evolve our ideas about maps - and explore further their relationship with respect to exploration, and meaning, and the context in which they are
used. It relates to this idea that viewing things at different scales is not about seeing less or more information, but different information and that generalisation is a pro active process of abstracting and emphasising particular scale dependent patterns inherent in the data.

The importance of interaction
This pro-active process links through to human computer interaction. I believe that interaction plays a significant role in how we give meaning to things. I no longer think that autonomous systems (as outlined at the beginning of my talk) are what we really want. In the situation where part of the cartographic decision making shifts to the system (away from the human), we need to identify what the optimum balance of decision making is between the human and the system.

I began my talk by discussing some of the impediments to developments in this field. But I feel there is a convergence of opinion on the importance of semantic reference systems as a necessary underpinning to automated cartography. This is what we should build our systems on, not the sand of paper maps.

Final Note
Irrespective of whether you agree with my views (or have even understood them), what I find immensely exciting about research in this field is how attempts at automation reveal the subtleties and complexities inherent in the art and science of cartography. That is something we can all celebrate. Thank you.

9.0 References
Anon (1965) Records of the discussion on the papers of Waldo, RT Bickmore DP and Boyle AR. International Yearbook of Cartography 5:30-33


