IS HALF THE LIFE OF KNOWLEDGE OF A DIGITAL CARTOGRAPHER NOW LESS THAN ONE YEAR?

A.R. Boyle, University of Saskatchewan, Saskatoon.

ABSTRACT

The inter-relation of the present changes occurring in cartography, hardware and software will be discussed. Often the present situation for cartographers, computer scientists and vendors can be traumatic. Perhaps the least affected at this time are the hardware designers and producers, as their now accepted rapid advances leave a trail of advantages and disadvantages that are continuously changing. The work of the writer is particularly concerned with the useful production application of cartographic/geographic systems in the real world, and it is difficult and interesting to accept the continuous analysis of the changing economics necessary to take advantage of the present trends; one year makes a major difference. Cartographers are now at last responding and changing their attitudes; the problems mainly arise in the availability of adequate software often due to poor original planning not following good basic principles to allow for the changing environment.

KEYWORDS - Cartography, Geographic Information Systems, Digital Cartography

For hundreds of years, cartographers have developed their skills to provide a graphic representation of the earth for human acceptance, maintaining the greatest information content with the greatest ease of use.

Gradually compromises have developed and have become frozen in accepted technique. On a closer analysis and with hind-sight, it is sometimes a little difficult to assess if this accepted technique was more for the cartographer or for the user. Until now most maps have been supplied by governments at a fraction of the cost of production; now that many governments are making their mapping establishments charge an economic price, customer resistance and demands become more obvious.

Over the years cartographers have cleverly developed symbolizations and colours for points, lines and areas; these have become accepted standards. Perception psychologists have examined the problem in some detail or given weight to the choices. Due to the needs of colour printing, different types of data have generally been allocated to different separation sheets and again accepted techniques have grown up. Much care has been given to the clear placement of names on maps and relating these to the optimum styles and point sizes.

Cartographers have always taken great pride in the best presentation of their map data. Appearance and clarity have been pre-eminent and sometimes 'truth' has had to be distorted to make the data readable. Fortunately the tuned mind of the reader, knowing these perception 'tricks', is usually able to accomodate them.

A few years ago the carefully constructed edifice started to crumble, and more recent changes have caused consternation to the well trained mind of the traditional cartographer. The reason is the need for base map data in digital form; this form of data is mainly for use in Geographic Information Systems. No longer is appearance and visual clarity required of the compiler, at least in the first stages of his work. These traditional skills only come to the fore when there happens to be a need for a map to be created from the digital data. The requirement from the compiler is for the precise (or as near as possible to the precise) truth because the compilation will be read and stored by a machine, not by a human mind. This implies that there should be no symbology of points, lines or areas and no created departures from the truth (usually referred to as generalization). While the cartographer has usually tried to reduce the number of separations to reduce printing costs, the need generally now is for more. The cartographer does not know which way is up and feels his job and his skills are threatened. This, in fact, is far from the truth, but the changes are certainly traumatic enough. In reality, there is no change from previous work. The cartographer/ compiler has always first drawn the truth and then symbolized and generalized it. Now the compiler finishes at the first stage, hands it to the machine for digitizing and storing and then only completes the second stage when a graphic map output is needed. Of course, the second stage of map compilation is done on a CRT interactive display, but that is merely a change in the tool used.

The greatest difficulty for the cartographer is to accept that the map- the graphic map- is no longer the be-all and end- all of his work. There are many other uses for cartographic data in computation and analysis. The traditional cartographer has only recently begun to see this.

We now have to look at the compiler/cartographer's work in the present age, as against the traditional. To do this let us assume that the cartographer has come to accept the new double system - first the compilation and then using the machine to produce the symbolized and generalized map, after pre-map edit on a display. The procedure could, in fact, allow far more creativity to the cartographer in presentation, as much of the previous tedious work has been removed. Let us assume this change of heart and examine in detail what the first part compilation process involves.

With contour separations, lines should not be broken to insert numerical values. In fact, numerical values should rather be put on a different 'name' sheet or, if it has to be on the contour sheet for some reason, then it should be in a different colour. Contours should not be broken lines where they are in doubt and, in general, they should be made to stop whenever they approach the near vertical cliff situation, rather then allowing them to coalesce. With modern scanners, it may, or may not be useful to use various line weights for height value discrimination.

With drainage separations, they will be similar to those used at present, but again dashed lines for intermittent streams should be avoided. River names should not be over-written, unless in a different colour to allow scanner discrimination. Any location symbols such as flow arrows or dams should be in another colour, or on another symbol sheet.

With man-made culture separations there must be a division into data types. One will be for all location symbology and, if this location symbology is to be read by scanning, must follow certain acceptable OCR rules. The second, for line data representing roads, railways, telephone lines etc, must all be non-symbolized and as single centre-line entities. Of course, for operator reference while drafting or when labelling, written notes can be added on top of any lines to be scanned, as long as they are in an acceptable colour difference.

An area sheet for forest stands will be defined by a boundary line only. Also areas such as for marshes, which have traditionally been defined by a pattern of location symbols, will be now defined by a similar boundary line.

In all cases, the technical needs of scanning, are paramount, but depend to some extent on the scanner. For efficient work these rules must be defined and obeyed. This is well within the abilities of cartographers and will not impose any restraints that they cannot accept.

Certain new additional data may have to be digitized to meet the needs of modern GIS thinking. For example this may require a line to be added down the centre of a lake to act as the river flow connection, the lake being a superimposed polygon on that line. Some hydrology techniques may require the watershed lines to be added. This type of thinking is at present in its infancy.

However, in all of the above, it will be noted that we are pandering to the graphic tendencies of the cartographer, whereas we should really be trying to obtain a 'digital base map'. But what is this digital base map? Certainly we do need the coordinates of points and cartographic line 'spaghetti', but it is also absolutely critical that we label them, and label them in a very special way.

We must appreciate that the drawn map was made for human assessment and that the reality of the presentation comes about because of a symbiosis between the drawing and the human brain, with all its complex relationship abilities. It is no good simply digitizing the map lines and adding a simple label to say that a line is a road or a river, or even to say a certain type of river or road. A machine computer at this time can only work on relationships of a spatial kind when it is told exactly what to do - and even then it has major difficulties. Attempts have been made using large computers and hours of CPU time. Thus, in some way we must organize, specify; and enter notes about the spatial relationship. Do we have to add something at the time of compilation- not symbolization, but something else? Do we need an inordinate amount of interactive labelling to make use of the capabilities of the human mind at the digitizing time? We come back to 'What is a digital base map?'. Nobody is yet clear. Of course if we only want to reproduce a graphic, in a complex Xerox machine way, then we do not have these problems, and many cartographic establishements have been misled into thinking that this was an acceptable route. There have been many failures and a few successes, but the only successes are usually for the pure graphic manipulation procedures.

The tendency at the moment is to treat the lineal and aereal units quite differently in the computer. The first is generally useful as properly labelled spaghetti and the second case as polygons. A polygon area is specified by the surrounding 'arc' lines with a label at a selected central 'centroid' point. The compiler has to think of these lineal feature and polygon boundary lines as different entities. However, a line such as that of a river may be both at the same time; this raises difficulties.

It will be seen that the cartographer has to have a much deeper understanding of spatial relationships and be able to define these; it can no longer be left to the viewer's mind. The cartographer is now starting to appreciate this attitude, but is still fighting against it; the lead is still being taken by system designers and such users as foresters and hydrographers. The title of this talk might sound to some - especially to cartographers - to be a little facetious; it is not. Matters are changing very rapidly indeed. Part of the change is in the needs of cartography and I have already discussed some aspects of this; part is in the rapid growth of technology to meet these needs. Cartographics has always been difficult for graphic equipment, because of the very large amount of data at such a very high resolution. One map separation alone, can require a resolution of $10^4 \times 10^4$ in pixel form. As a result, it is always pushing the hardware state-of-the art to obtain good results. Each six months new abilities appear in CPUs, displays and storage media and these have to be taken with both hands immediately they appear.

Of course, there is still a long way to go; personally, I hope that the rate of advance accelerates. We do need to have large area high resolution displays; at this time we can only look at part of a maps, unless we plot it out. We do not have large enough storage with fast access and of an archival nature. The answer to the first problem appears to be in advances in large flat screen displays; for the second, the best approach appears to be the laser written optical digital disk. Very fast multi-processor 32 bit machines are also required for fast large-area coordinate handling.

There is a problem in forming a group of people to properly assess the problems and the changes. Cartographers themselves are often too inbred and do not have an adequate academic knowledge of thinking in terms of the needs of digital databases and complex graphic systems. On the other hand, computer scientists can produce software design but have difficulty in appreciating, or being interested in, the cartographic/geographic problems. Geographers interested in GIS manipulations are far too few in number. The odd engineer like myself can only tackle part of the problem.

In the last decade, the main thinking was from vendors often trying to expand their CAD/CAM systems to GIS. In general, they understood the graphic part (as do cartographers) and some made gallant attempts, but did not appreciate the real problems of the software and data formats for good GIS manipulations. They offered to supply and actually delivered software and hardware, but the real results were disappointing, and any geographic manipulation they claimed tended to be excessive in time (and cost) so that it was of little real advantage to the user.

Over the same period a few vendors aimed specifically at cartography and GIS as against CAD/CAM, but these were small companies and made little money, without the necessary large and exotic sales forces associated with the much larger market of CAD/CAM; as a result their efforts were limited in scope and effect.

For these reasons the main design load of GIS appears to have been taken by a few groups of consultants, each consisting of a mix of geographers, cartographers, computer scientists, engineers and others. Their efforts are now much sought after by the large government mapping organizations and the users, such as foresters, where GIS applications seem to be nearer useful fruition than in many other areas.

If cartographers themselves are to maintain their position they must now take the lead. They have the necessary organization but must learn, and continuously re-learn methodology and advances in cartographic, GIS and hardware. If they do not do so, new establishments will arise and replace them. It is almost certain that in 1983, more maps will be made by computer systems (without cartographers being used) than by the traditional cartographers.

We have to examine how cartographers should go about this. Certainly they have to understand the true modern reason behind the production of cartographic data. They certainly do not need to learn programming. They need to understand the manipulation functions of cartographic/geographic data that are required in the next decade. Some of these foundations of needs have been laid in recent papers(1), but many others have still to be thought out and defined. They do not need to understand hardware, any more than most people understand their car engines, but they must appreciate the capabilities and shortcomings. They should appreciate the future possibilities and not become hardware-bound in a limiting way.

They also have to appreciate that much present (and also future) software is poor in results both as regards functions and speed. They must be able to assess these short-comings relative to expectations.

Mainly they must be able to organize their work. Their future lies in the build- up of GIS for manipulation, and they must be able to enter data and store it with a true understanding of future needs. This is difficult, and there are many systems around today that rapidly made themselves useless by entering data in a way that has entirely limited any future usefulness. Graphic information system philosophy is difficult to define and perhaps most difficult for the traditional cartographer. He still tends to believe that a human mind must be involved in the analysis and that the best visual communication means is by a series of discrete lines on a graphic. In fact he has developed this line aspect to a fine technique in contouring. Contours are accepted and understood by most people, but they are only one philosophical interpretation of the earth's shape and, of course, do not really exist on the earth's surface. A spectrum of colours fading into each other might well have been used instead, if this had fitted with the printing colour technology at that time.

Cartographers, in being so bemused with line representation, have understood digital vector line data but not its topological shortcomings. The storage of data in raster format is much more difficult to comprehend, although they see it every time they view a weather map on TV. They feel that somehow rasters are a disorganized approach and do not appreciate that it inherently retains much of its spatial information that is lost in vector forms.

Their first thought has always been that data must be entered line by line in vector form and stored serially, one label and line after another. As one vendor once said to me, "if it is too slow in retrieval, we will simply supply more hardware because you can do anything with hardware!". Up to a certain limit this is perhaps true (say 10,000 coordinates total), and this is usually the case in engineering drafting, but search times increase from seconds to minutes, and even to hours, for high resolution contour sheets which may well contain a million coordinates when in vector form.

Much GIS work is relative to areas and, if line data is used as the standard method, then areas must be described as polygons limited by their boundary lines. Some problems arise in this. For example, is each boundary line duplicated or is it maintained only once and that pointed at by a polygon directory? Many vendor systems utilize the duplicated full-polygon approach, and thus end up with double the already very large amount of data. In fact, it is worse than that, as polygons themselves cannot be updated successfully. It is not possible to update one polygon without affecting adjacent ones, so that all updates must in reality be made on the original boundary line data followed by new polygon re-forming. This updating problem means that without a polygon directory structure, data can become triplicated!

Maybe polygons should not be used, because the real problems show up in the expensive CPU times of overlaying. Almost all real world geographic area manipulations require some form of overlaying, thus creating a new area where certain specified conditions overlap. Maybe grid cells should be used. Grid cells are small square areas labelled with an average attribute. This can be very effective in overlaying, but seems to lose its advantage when going to the very fine grid cells that approximate in resolution to high resolution polygon definition. The grid cell approach is foreign to cartographers but not to geographers. Personally I believe that this is due to the earliest digitizing work being done by geographers at low precision to meet their budgets. The digiting process was to overlay the map with a transparent grid and key in the average attribute as seen in each square, in turn.

Recently a group at Denver(3) has come out with a hexagon based grid cell method. This is a revival in computer form of an old idea in geography. The advantage is that it provides a 'conformal' ability, or the fact that distances between centers of all grid cells, either along the axes or at 45°, are the same.

Cartographers have heard of DBMS and think that it is only necessary to buy and employ one of these in their computer for all to be well. The problems of fast search on two axes simultaneously are not appreciated. The search procedure cannot be pre-specified in a way that is acceptable in linear searches for which most DBMS are defined and, as a result, most DBMS are completely useless on geographic data. The more knowledgeable wonder if they should go to the quad-tree and octa-tree sturctures as proposed at the University of Maryland. Many new structures for GIS need to be devised and tested, at first in an academic atmosphere. It is a pity that we cannot find more academic cartographers and geographers prepared to do this.

Knowledgeable cartographers may also have heard of the advantages of raster structures (2) They realise that this is something like grid cells, but more efficient. They have heard it tends to be more bulky than vector data, but then they have also been told that compaction methods such as run length encoding can be almost as efficient as Freeman chains and delta techniques in vectors. Because of the recent implementation of scanners for digitizing and plotting, they know that scan data is with them, but again they have heard of the large computer times involved in vector to raster and raster to vector conversions. Does this mean they might maintain their data in raster at all times? They are not so averse to raster formats as might be expected, because they have grown used to scanners in the graphics arts world and printing. The difference between the use of an analog intermediary situation and a digital one is not clear to them but does not confuse their thinking either.

It will be seen that the cartographer must be confused by all the changes he sees and indeed, by the half truths he is sometimes told. Each time he learns something new, the state of the art of hardware or software designers, computer scientists, geographers or vendors moves still further forward, but still with no perfect end in sight. For those of us in these changes and indeed creating them, this situation is normal, but to the cartographer and to GIS administrators it is difficult to accept. Some say they will wait until it is all finally defined and is presented to them as A, B and C. Others are trying to catch up and find they have to run even to stay in the same relative position. I think that we must accept responsibility at this meeting to help such real-world groups as cartographers in every way we can.

I do not believe that the problems are being created irresponsibly. Traditional cartography is not meeting today's need. It is too slow and costly in production and cannot be used by machines - machines of all types, examples being for direct printing and even more importantly, for navigation. The needs are real, but those of us trying to fulfill those needs do not yet know efficient and understandable answers. We do not yet even know all the questions! However, it is up to us to try to help the cartographers in their dilemma, which is much wider and all- embracing and multi-disciplinary than any of us have in our more equably defined disciplines, (cartography until recently was also like that). Canada is at this time a leader in these changes to GIS and we should see that this lead is maintained.

REFERENCES

1. R.F. Tomlinson and A.R. Boyle, "The State of Development of Systems for Handling Natural Resources Inventory Data", Cartographica Vol. 18, No. 4, Winter 1981, pp. 65-95.

2. D.J. Peuquet, "An Examination of Techniques for Reformatting Digital Cartographic Data, Part 1: The Raster-to-Vector Process", Cartographics, Vol. 18, No. 1, Spring, 1981, pp. 34-48 and "Part 2: The Vector-to-Raster Process", Cartographica, Vol. 18, No. 3, Autumn, 1981, pp. 21-33.

3. Interactive Systems Corp., Littleton, Calif.