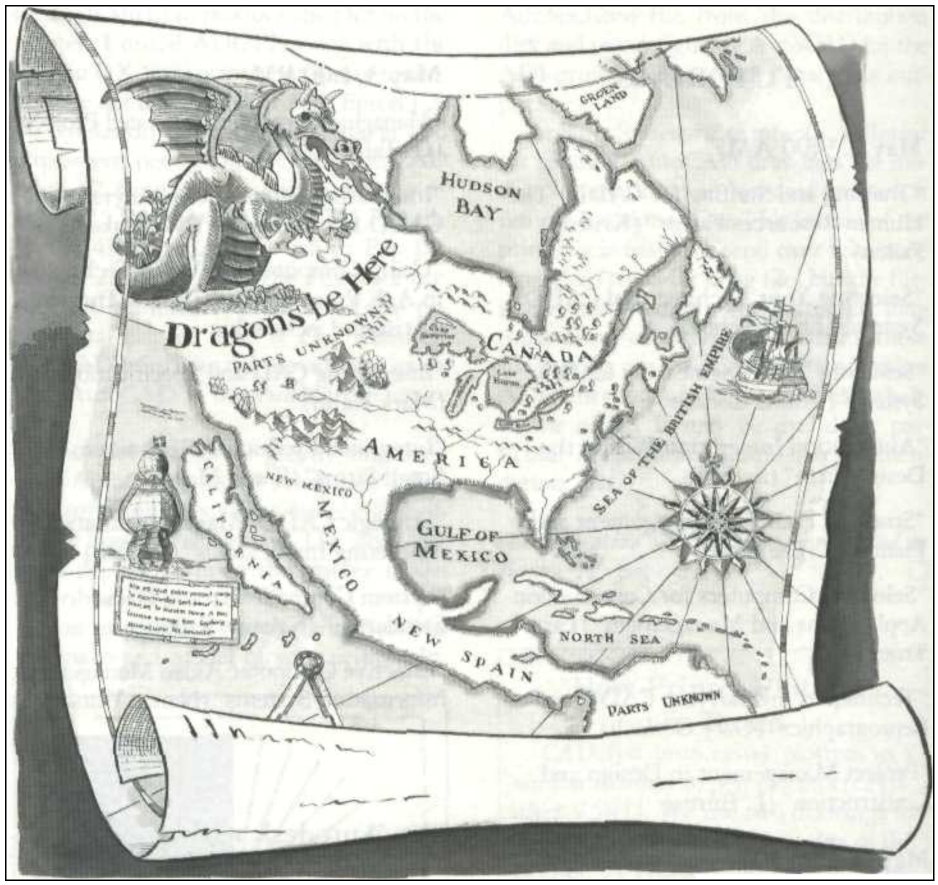
THE GIST OF MAP DATABASES

by Andrew Zolnai

Geographic Information Systems (GIS) link graphic and information databases to produce new, up-to-date maps rapidly.



BEAD NICI JASON

Have you ever noticed dragons, compass roses or inconsequent place names peppering the large empty spaces on a map? The dragons and other fillers stem from our phobia for empty spaces. Meanwhile other areas of the map may be barely legible for the crowded lettering of important entries. This problem is particularly acute in maps which have an intrinsically erratic distribution of data. Maps are traditionally kept on individual sheets in a variety of scales, and the information displayed is a compromise between abundant data and limited space.

Any fixed format cripples map handling if the information is complex, erratically distributed and rapidly evolving, as it is in the resource industries. In the northern oil business, for example, a new and crucial mile-square territory may require a map containing far more information than the maps for a thousand square miles of less prospective tundra. The information it holds may change with every seismic test result, and tomorrow's choice of testing site may depend on correct evaluation of yesterday's tests. With the short season and high costs of the Arctic, a mapping system needs to adapt easily and deliver fast. The trick is to devise a computer mapping system that allows mapmakers to edit and generate maps only as needed, at a scale appropriate to the density of the data.

A Geographic Information System (GIS) provides a particularly flexible means of storage and retrieval of graphic data, by coupling computer databases intelligently with powerful graphic interfaces. The elements of a GIS include a graphic editor, a database management system (DBMS), various input and output systems, and some means to link these together (fig. 1).

Computer Mapping

Early computer mapping systems coped with myriads of points by using huge graphic files on large mainframe systems. The monolithic map files were difficult to maintain and took a long time to process, and the systems themselves were hugely expensive.

The next step was to apply to graphic files the same principles of traditional database management: storing the graphic elements separately within an appropriate framework, controlling file size, and providing links to combine the files in various ways. The individual graphic files are tagged with suitable locators such as grid references or latitude/longitude to facilitate combining them by machine. By linking individual elements into composite graphic files that remain manageable in size and legible in output, one can then create maps of target areas and "zoom in" on areas of greatest density.

More powerful microcomputer graphics and software, now free from the burden of handling huge files, have opened up the field to affordable systems.

Unit Area vs. Unit Density

Unit map area is typified by constant-area map sheets all drawn to the same scale—where, for example. 96 map sheets representing one square mile might be required to depict a twelve-by-eight mile region at 1:50.000 (fig. 2a).

Unit data density suggests that if one-half of the region (the west side of fig. 2b] contains only four features of interest, a single map sheet at say 1:250.000 may be sufficient to portray it. Another area (the northeast quadrant of the same figure), almost as lightly detailed and containing seven features, could be rendered on two sheets at 1:125,000.

Conversely, if a single square mile is crowded with sixteen features, dividing that map into four separate sheets at 1:10,000 will come closer to creating map sheets of equal data density (ie. about four features per sheet). In this way many fewer maps (30 sheets in this example), organized by unit data density, can be used to contain the same information as the 96 sheets organized by unit map area.

To link unit data density maps, a computer DBMS provides the obvious tool. Each map sheet file requires a DBMS entry to define its location, extents and scale. The map sheet files may then be combined by a graphics program that can use the DBMS entries to control the process. The sheets can then be correctly placed in relation to each other at the same scale. The problem of text and symbol size in combining files held at different scales must be solved by the graphics program.

Graphic Entities Database

Taking the principle of unit data density one step further, it is possible to make greater reductions in the empty space, and the data storage requirement, by converting the graphic entities themselves into entries in a database file. Each entity requires one entry with only the necessary data to allow the graphics program to render it fully. Thus a DBMS entry for a straight line might contain the location of its endpoints, its visual linetype (eg. dashed) and a key to its nature (eg. property boundary). A "symbol" entry might need only its symbol name and location coordinates; the graphics program renders the symbol's appearance from a single copy in its image library.

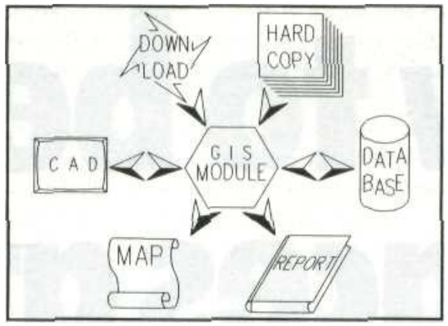


Figure 1: Synopsis of a GIS,

With this approach, the problem of scale can be eliminated. The actual geographic data may be held at full scale, and is reduced automatically as needed by the graphics program. Symbols and text must be sized appropriately to some constant ratio to the sheet size, but this is also conveniently done by program as the symbols and lettering are generated.

If in fact all the mapped territory's graphic entities can be reduced to such DBMS entries, then it may no longer be necessary even to subdivide them by unit data density map sheets. The data reduction is so significant that they may simply be held in a single uniform database of moderate size. By linking this DBMS to a suitable graphics engine, the required graphic entities may be extracted from the DBMS and used on demand to generate

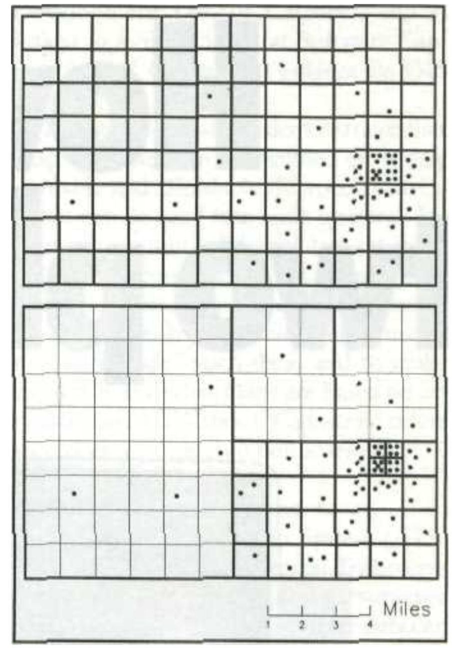


Figure 2: An eight-by-twelve mile region represented on maps of (a) Equal unit area, and (b) Equal data density The dots suggest a typically uneven distribution of features of interest.

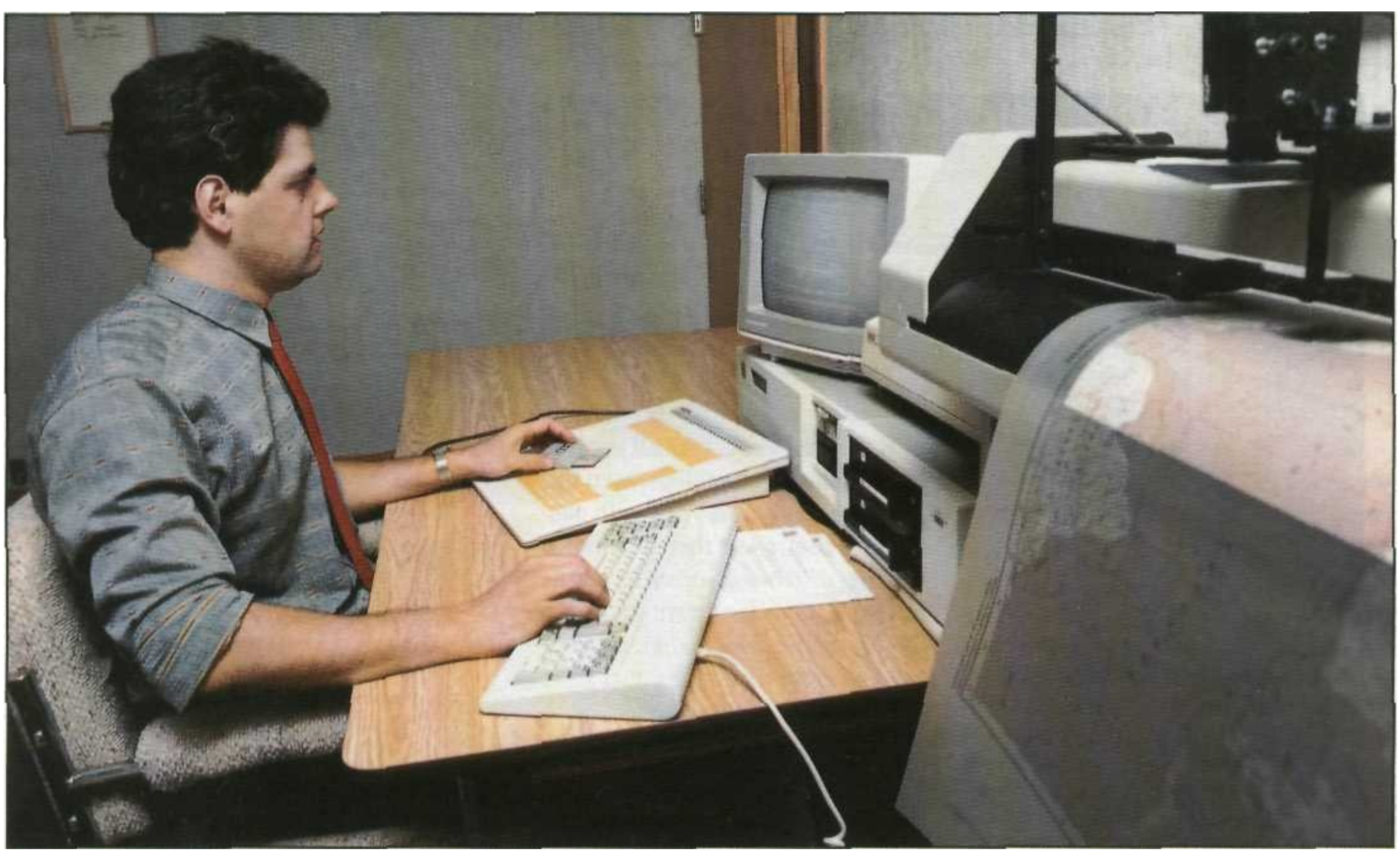


Figure 3: Video-tracing workstation.

any map at any scale- This is the current State of GIS development.

AutoCAD and the GIS

One implementation of such a GIS uses AutoCAD as its graphics engine. Georef, by Software Support Ltd. of Edmonton, Alberta, allows users to input and query data as AutoCAD drawing attributes, and offload the attributes into a DBMS. This system permits users to convert or to extract alphanumeric data as AutoCAD drawing entities. The database link may be to a micro-based DBMS like R:Base System V or Paradox, or the micro-to-mainframe uniform-structure database called Oracle. Software Support's Munmap system, an application package, is specific to the needs of municipal map users, who often have their data already on a DBMS *[see Municipal Mapping Made Affordable, p. 29}.* My company customised this package further to create Minmap and Oilmap for the resource industries.

Much resource data is only available on map sheets. The only practical way to enter this information into the database is to digitize it. In order to streamline the attribute input, we used a digitizing system by Brighter Images of Lafayette CA, which synchronizes video and AutoCAD graphics for on-screen tracing of drawings from an image of the original.

Crucial in terms of volume and time spent is the consideration that many hardcopy documents in various forms (paper, mylar: maps, sections, etc.) must be entered into this GIS. It is important to be able to digitize onto an existing grid, and to control imported data against that grid. Otherwise many an oil well is mapped on the wrong side of the road. Interactive editing capacities are desirable, as it helps to be able to compare original hardcopies to digitized input rapidly.

Video-tracing systems accomplish such comparisons by projecting a video image of the plotter-mounted hardcopy on an AutoCAD screen (fig. 3), matching and synchronizing them via AutoLISP. Digitizing is done by AutoCAD tracing and drawing over the video. The interactive WYSIWYG process (pronounced Whizzywig for "What you see is what you get") and the automatic quality control offered by image superimposition, both help to increase speed and accuracy over tablet digitization. Direct vector digitization bypasses the raster-to-vector translation problems associated with scanners, and allows direct conversion into ASCII formats for post-processing. Attributes can be entered at the time of digitizing. The resulting data may be easily extracted into RiBase and merged with downloaded data.

For GIS hardware, we ourselves use an AT compatible for processing speed midway between an inexpensive PC and the powerful 386 machines. Control over file sizes, through database management, can allow large maps to be handled on 640K PCs if needed. Our particular machine can be linked via Ethernet to a SUN 3/160 network file server (at Teknica Resource Development Ltd. of Calgary), and is linked to a SUN 3/60 (at Software Support Ltd of Calgary) to enlarge its GIS capacity.

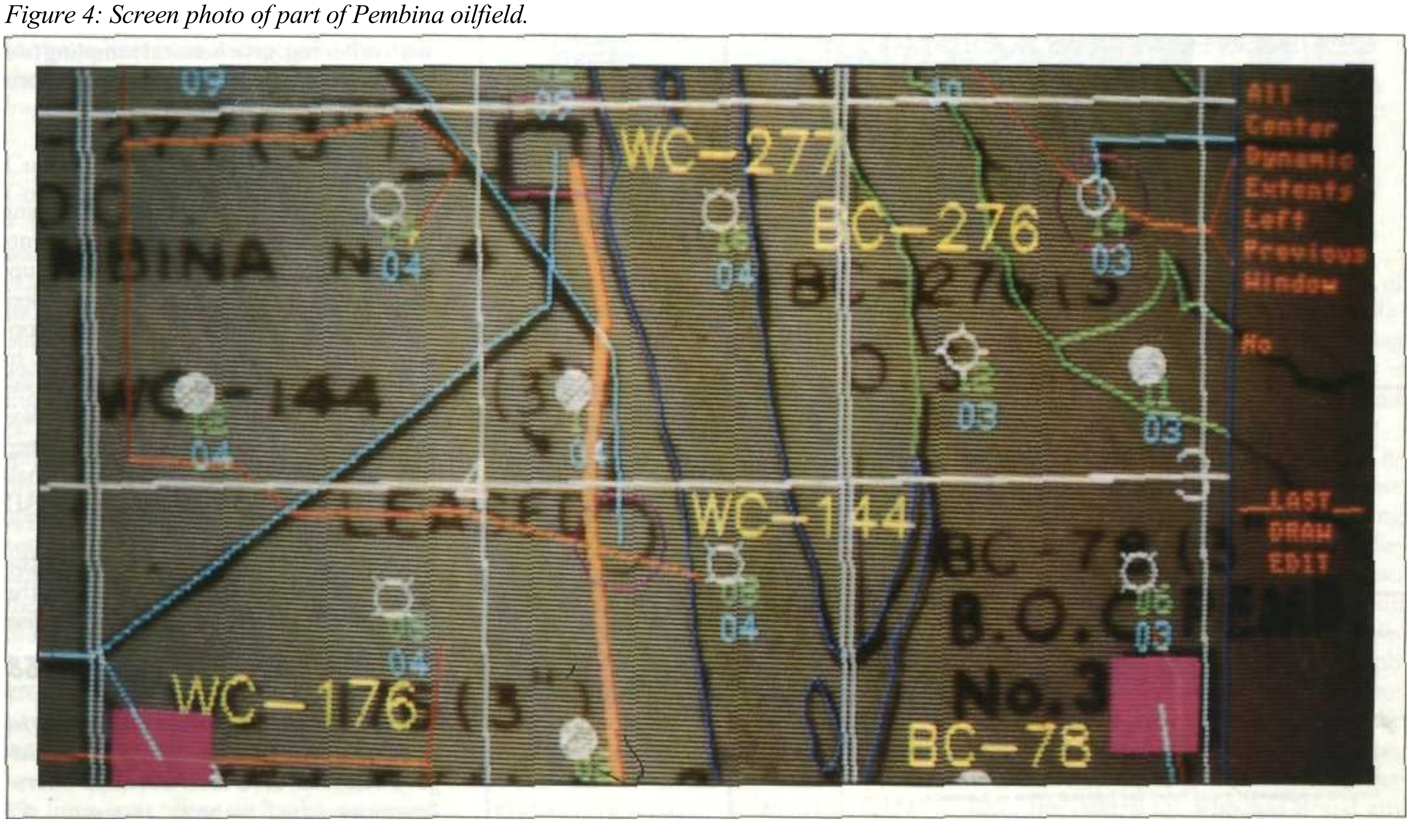
This array of systems demonstrates that map databases can be handled on widely available and affordable microcomputers. What can be read into such systems? Increasing amounts of digital data are available from commercial or government sources, or in-house as computerization evolves. The key to their proper use is to "high grade" such data, and to make especially sure that maps are based on accurate cadastral bases. ("Cadastral" refers to property ownership boundaries.)

Some GIS Applications

The Alberta Township System (ATS) is a comprehensive Universal Transverse Mercato(UTM)grid, recently resurveyed and calculated, to which all other data can be referenced. It includes grid subdivision and road allowance alignment. Software Support's product Atsplot makes the ATS data available now in an AutoCAD DXF format compressed and stored in under 10Mb. Cultural data may likewise be added at under 30Mb. An ATS grid with road allowances, pipelines, road access, wells and morphology for one township in the Pcmbina oilfield in Alberta forms a roughly 180Kb drawing (fig. *4).*

This GIS can produce a map with as many layers as there are data types, broken down into areas of coherent and manageable data density. It enables us to clip

Figure 4: Screen photo of part of Pembina oilfield.



together various townships (or portions thereof) out of roughly 300 held on file. We can assemble them into target quadrangles with the desired layers of information. Any area of interest can be clipped out and worked upon, but coordinates remain in that single UTM grid. The Georef module allows us to import other data into the AutoCAD maps, such as oilwell locations or geophysical survey stations.

Similarly, the province of British Columbia makes historical mineral occurrence data available to the industry on floppy diskettes. Georef allows us to query and input these data from R:Base into AutoCAD. Oilmap and Minmap are other modules that streamline the input of attribute data onto the maps within AutoCAD, referenced to a UTM grid province-wide, or a property grid locally. Contouring (using algorithms that grid, or triangulate, and map surfaces with common properties) can furthermore be performed for volumetric resource assessment such as oil, gas and mineral ore reserves.

Another resource application concerns bore hole measurements, They reflect rock and mineral types underground, and used to be output as hardcopy logs, a wiggly trace on a strip of paper. Once digitized as a polyline, these can be sampled by a LISP routine into files, and merged with more recent digital logs for contouring. When done, the contour maps can be re-entered into AutoCAD for overposting correction, superimposition onto an existing grid and possibly further digitization of data not handled by contouring or other algorithms. These can be merged with other (eg. geochemical sampling) databases to create comprehensive reports and high quality maps.

Tracks of the Dragon

The challenge of computerized mapping is to enhance bulky hardcopy documents into a graphic database, to keep them up-to-date, and to create custom maps or sections suitable for each purpose. The concept is to move from a map area unit to a data density unit. The answer is a powerful graphic engine coupled with a flexible database. The tool is a GIS to keep files manageable in size, amenable to small system processing, yet capable of full CAD and database management.

Once, dragons may have roamed the empty spaces of ancient maps. Now, the principles of a GIS breathe fire into a new era of map management.

Andrew Zolnai is a professional geologist and the owner of CAL-CAD Ltd. in Calgary, Alberta. CAL-CAD develops microcomputer based mapping systems. □