Knowledge-based systems and geological survey[1]

T.V. Loudon

British Geological Survey, West Mains Road, Edinburgh EH9 3LA, UK.

E-mail address:v.loudon@bgs.ac.uk

Preprint of article published in *Zeitschrift der Deutschen Geologischen Gesellschaft* (Journal of the German Geological Society) as part of the volume celebrating the 35th Anniversary of the International Association of Mathematical Geology, edited by H. Thiergartner.

Bibliographical reference: Loudon, T.V., 2005. Knowledge-based systems and geological survey. *Zeitschrift der Deutschen Geologischen Gesellschaft* vol. **155 (**2-4), 225-246. (Stuttgart, June 2005)

Abstract

This personal and pragmatic review of the philosophy underpinning methods of geological surveying suggests that important influences of information technology have yet to make their impact. Early approaches took existing systems as metaphors, retaining the separation of maps, map explanations and information archives, organised around map sheets of fixed boundaries, scale and content. But system design should look ahead: a computer-based knowledge system for the same purpose can be built around hierarchies of spatial objects and their relationships, with maps as one means of visualisation, and information types linked as hypermedia and integrated in mark-up languages. The system framework and ontology, derived from the general geoscience model, could support consistent representation of the underlying concepts and maintain reference information on object classes and their behaviour. Models of processes and historical configurations could clarify the reasoning at any level of object detail and introduce new concepts such as complex systems. The up-to-date interpretation might centre on spatial models, constructed with explicit geological reasoning and evaluation of uncertainties. Assuming (at a future time) full computer support, the field survey results could be collected in real time as a multimedia stream, hyperlinked to and interacting with the other parts of the system as appropriate. Throughout, the knowledge is seen as human knowledge, with interactive computer support for recording and storing the information and processing it by such means as interpolating, correlating, browsing, selecting, retrieving, manipulating, calculating, analysing, generalising, filtering, visualising and delivering the results. Responsibilities may have to be reconsidered for various aspects of the system, such as: field surveying; spatial models and interpretation; geological processes, past configurations and reasoning; standard setting, system framework and ontology maintenance; training; storage, preservation, and dissemination of digital records.

Keywords: knowledge systems; ontology; field mapping; complex systems; geological reasoning; spatial models; hypermedia.

---

[1] This paper is published by permission of the Director of the British Geological Survey (NERC). It reflects a personal, not a corporate, viewpoint.

Introduction

The leitmotif of this festschrift dissertation is ‘explore and transform’.

Successful explorers, drawn to what may lie beyond the horizon, are guided by intelligent speculation, but learn as they go and adjust their expectations as they proceed. Therefore Columbus, seeking more convenient access to Asia in 1492, is not pitied for his apparent failure, but celebrated for his navigational skills in locating the obstacle. Thus geologists exploring developments in their science forever scan the horizon: is that a cloud or a hint of undiscovered territory? The horizon, scanning mathematical geology from a personal viewpoint in geological survey, seems to me alive with strange portents. Each on its own might be brushed aside: together they signal a change of direction. Certainly, there is much to reconsider on this 35th anniversary of IAMG. But we are not lost; we are merely adjusting our expectations – exploring and transforming.

Information technology (IT) could transform the wonderful concept of the geological map – a product of geological survey. Geological surveying deals with the spatial characteristics and relationships of geological objects and the procedures whereby we seek to observe, record and explain them, and to communicate and visualise the results (as on a map). It is about interpreting or modelling the objects’ properties and behaviour, and the processes and events that formed and transformed them, for these are essential to understanding the geology and filling the gaps where we cannot or do not observe. Spatial aspects play a large part in geological reasoning. They are the key to disentangling events in geological history, as William Smith realised long ago (Winchester, 2001). His maps, like ours, were tools that transformed spatial pattern by scaling it down. They enable us to visualise an entire region just as we might observe a feature at an outcrop or on a hand specimen. They also record the locations of our observations and interpretations, tie them to the topography of a base map, place them in a stratigraphic sequence, and predict what is present at any location. However, the geological map embodies just one outcome of a system of investigation based on human knowledge. Perhaps it is time to adjust the IT framework for geological surveying by focusing it, not on transforming the map, but rather on supporting transformation of the underlying knowledge-based system. The focus of geological survey could then be on creating, not the geological map, but rather a hypermedia knowledge repository from which maps and many other products could be produced.

Lingering map metaphors

Digital cartography partly automated the geological drawing office, also interesting some managers (intrigued by its high initial costs) but only a few geologists. Map-making in general is big business, justifying a large investment to develop robust commercial systems that are indifferent to what is depicted. Success came slowly, but after some forty years of development, digital cartography in geology has now met the initial aims of providing well-printed maps more quickly at less cost. If supported by a good archive, sheet revision is simplified. It is a giant step for printing maps, and a tiny step for geoscience. That is a pity, because information technology is most successful when enabling things to be done differently, rather than in its horseless-carriage mode of imitating earlier procedures.

Mathematical methods made an early contribution to subsurface geology in automatic contouring. Take an algebraic function, say a power series, and fit this to the data, say of elevation measurements of a formation top at well locations. This process, of generating a surface from which the location of contours can be mathematically calculated, is easily explained to a geologist, perhaps along the following lines. Imagine a horizontal plane going through the middle of the three-dimensional pattern of elevation measurements. Adjust it by adding on a proportion of elevation values from another plane that is tilted a bit, to bring the composite surface generally closer to the data. Add scaled proportions of more surfaces (called basis functions) that are curved to some extent in a north-south direction and possibly to a different extent east to west. Just use a few of these, because otherwise it gets confusing, for each time you add another one you have to recalculate all the others. They are curved like that because they show the variation of the square or cube or higher power of a number as it steadily increases in an eastward or northward direction. The proportions of the basis functions are chosen to bring the composite surface near to the data by moving each point of an originally horizontal surface vertically up or down by a carefully graduated procedure. The result of all the additions is known as a fitted surface.

If that lacks geological appeal, one might invite the geologist to take a statistical view of this process of bringing the surface near the data points. (If this were a multi-media presentation, I would begin at this point to intersperse short clips from a sequence of John McEnroe discussing tennis with an umpire.) The computer adjusts the function by a statistical procedure that determines the proportions of the various basis functions that will give a least-squares fit. To picture this, think of vertical lines joining each data point to the fitted surface represented by the function; then imagine each line as the side of a square. If you add up the areas of all the squares, the least-squares fit is given by the function that gives the smallest total area. The result might be regarded (on the basis of various statistical, if geologically implausible, arguments) as giving the best estimates of elevations at points where the true value is unknown. The estimates are uncertain, but lead to confidence intervals within which the surface is expected to lie.

The results can be (and have been) manipulated in various ways. For example, estimates from the fitted surface can be subtracted from the original measurements to give 'residuals', and new functions fitted to the residuals to depict smaller-scale variation. In other words, the same rituals are applied to values of the discrepancies between the original measurements and the surface generated by a rather arbitrary composite function fitted by a dubious criterion. Then again, major discrepancies in the original data might be identified, defined as anomalies and removed from the surface, which can then be recalculated without them, thereby selecting the data to fit the model. But the real surfaces are geological, not mathematical, and even superficial examination of a good exposure or a present-day landscape suggests that they do not really look like polynomial surfaces. Nor do they entirely resemble products of more recent versions of the surface-fitting process, such as geostatistics, kriging, splining or patch blending. Furthermore, there is no clear match between the mathematical procedures and geological reasoning processes. The use of statistics implies uncertainty. This cannot be attributed to the surface, for the surface is not uncertain: it is where it is. More plausibly, perhaps, the uncertainty can be attributed to random or stochastic elements in the geological processes that formed and transformed the surface, of which more shortly.

The discomfort with such methods is not just geological. Here is a philosophical and statistical viewpoint that applies to regression and surface fitting in general (Howson and Urbach, 1993, page 294): "The classical arguments in favour of least square estimation of regression parameters are quite untenable. That based on intuition is inconclusive, vague, and lacking in epistemological force. The Gauss-Markov justification rests on the linearity criterion, which itself rests on nothing. And the maximum-likelihood defence is simply fallacious." On page 297: "the two standard interpretations of confidence intervals [categorical-assertion and subjective confidence] are both incorrect. Hence confidence intervals [on a regression line] do not constitute estimates: for the same reason they cannot properly function as predictions."

Webster (2000) defends the traditional statistical approach for mapping soil types. He points out that a typical geostatistical analysis of soil data proceeds on the assumption that the properties of interest are the outcome of random processes, and asks whether the assumption is reasonable, using arguments that could apply to many areas of geoscience. He concludes that: "The processes that have together created the soil are many, interact with one another, and can amplify even small differences in the starting material by positive feedback. They and the fact that accurate knowledge of the soil is so fragmentary make the result indistinguishable from the outcome of random processes. In this situation, we may model the soil as if it were random… For practical purposes, especially local prediction, we need statistical models of the soil… Our questions concerning randomness and stationarity can refer only to our models. They do not have straightforward answers. All we can ask are whether the models are reasonable in the circumstances, and whether they are profitable in that they lead to accurate predictions."

This line of argument might suggest that the random element of the model identifies and quantifies areas of ignorance but does little to enhance our geological understanding. Hinze et al (1999, page 52) prefer to use hand-drawn cross-sections as the basis for their computer model of a wide mesh of boreholes through Quaternary deposits of complicated irregular shape. "Interpretation of the data by an experienced geologist well acquainted with the regional geological setting is needed to delineate reasonable boundaries. 'Reasonable' means in accordance with the principles of sedimentation, erosion, stratigraphy and tectonics. No mathematical algorithm, as well defined as it may be, can construct reasonable geological boundaries from a limited number of randomly distributed points."

Of course, surface fitting is useful in geology, and many thousands of geologists draw contour maps to prove it – with paper or plastic, pencil, eraser, pen and ink. Accounts of the rationale behind manual methods of geological contouring may range from the limp to the bizarre and the results may be inconsistent, but they work. Computer methods lack the flashes of human inspiration, but can produce a clean and tidy product, and on large routine tasks are faster and more convenient, and may give some better predictions and more consistent results. Interactive contouring programs may allow the geologist to adjust the calculated lines, thereby improving the appearance but abandoning any claim to mathematical consistency.

Computer contouring may be another qualified success. But greater gains are likely when IT encourages geologists to consider more clearly the reasoning underlying the maps and how new tools can improve its representation. There are encouraging smudges on the horizon, discussed later. At present, automated contouring succeeds largely by replicating features from earlier technology, again like a horseless carriage. (In a multi-media version, I would include a clip of the early automobile thus referred to. An optional extra was the whip holder within easy reach of the driver. The whip was not for disciplining the non-existent horse, but was needed to signal a left or right turn by twirling it in a standard pattern. Until the 1940s, the British Highway Code illustrated the pattern, largely for horse-drawn vehicles.) Lingering metaphors may be essential for backward compatibility, incorporation of legacy systems and retention of the customer base. But they can lead to awkward imitation of incidental and ultimately irrelevant features of the system, which should not be allowed to block or distort our view of future prospects.

Multifaceted survey

To get beyond mere imitation of the present-day products of surveying, we need to think more deeply about the underlying objectives. This has the potential advantage of moving from a view (or conceptual model) of geological surveying based on cartography to more flexible models based on IT. They may not liberate the information content from its containers, but in the spirit of a Tupperware party (this is after all a 35th anniversary) should at least offer an attractive range of new containers to organise our ideas and observations more agreeably.

In field mapping, geologists must weigh evidence from many sources, as Harrison (1963, page 227) pointed out in his outline of the steps in preparing a geological map. At each stage, geologists must weigh many inputs: a framework of pre-existing definitions, concepts and procedures; observation and measurement; comparison, analogy and correlation; expectations of rock properties and processes; interpretations of Earth history; and reconciliation with topographic evidence and geoscience models from other studies. They must take a top-down view throughout, looking carefully at

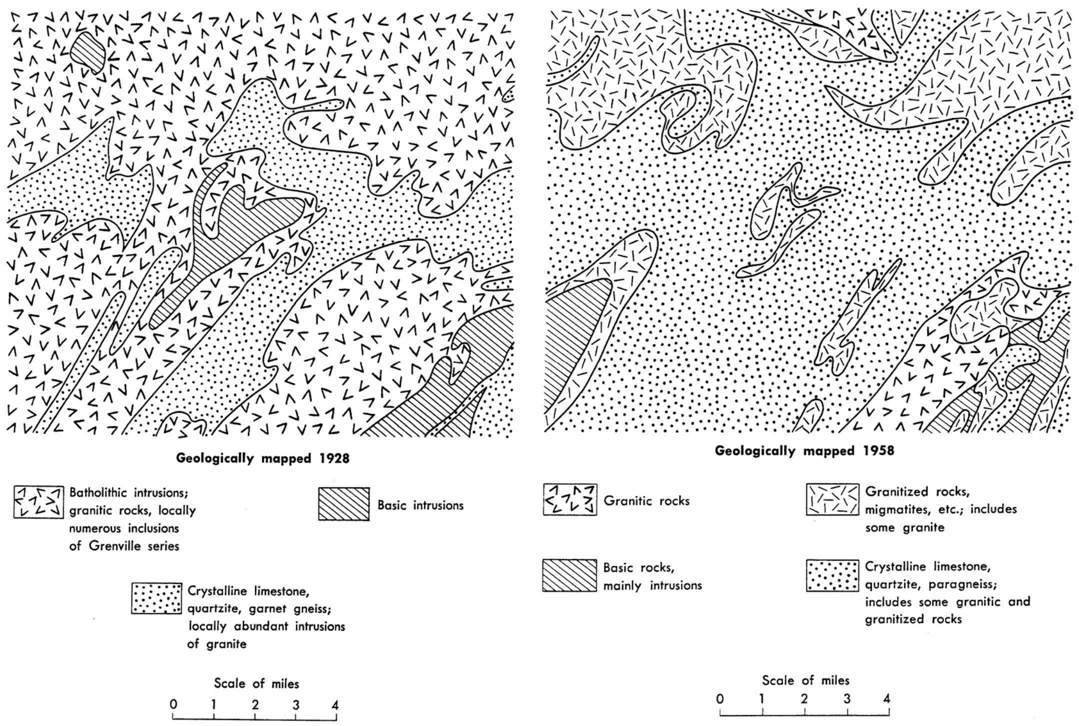

Figure 1: The two maps refer to the same area in the Canadian Shield at the same scale. The rocks did not change: views on granitisation did. Published with permission, from Harrison (1963).

all they know so far, testing and improving their current interpretations, imagining how the situation might be in its entirety, and deciding how best to complete the picture by filling the space from the available fragments. This holistic view is needed because items in isolation cannot provide a full understanding, as their significance may depend on their place, role and function in the system as a whole. Harrison provides a startling illustration of the consequences of adjusting a top-down view (figure 1).

Nevertheless, a Geological Survey map has several distinct parts. The map face, depicting the geology with colours, lines, ornaments, comments and symbols, is overlaid on an independent topographic base and a geographic grid. The margins may show vertical cross-sections and a stratigraphic key of formations arranged in time sequence. There may be a generalised vertical section that illustrates the sequence of formations, their variation in thickness across the map sheet, and the stratigraphic relationships between them. In addition, the map probably relies on a text explanation published as a separate document. The holistic view is in the surveyor's mind: the records are fragmented.

During field survey, the geologist might record some items of information (such as measured sections) relevant to the depositional model represented in the generalised section, and some (such as bedding orientations) that are relevant to the effects of the structural model illustrated on the map face. In the field, different aspects of the geology (such as the depositional and structural models) may be studied at the same time and place but interpreted and depicted separately. At a more detailed level, different models, say of fold geometry and location of folds, may be developed simultaneously but eventually depicted separately on stereograms and maps. The cross-sections in the map margin may depend on borehole descriptions obtained from other sources and archived separately. Other information may be recorded in text form in the map explanation, such as detailed descriptions and chains of reasoning involving sequences of geological processes and events. The explanation may refer to other maps, perhaps adjacent sheets, or geophysical or geochemical maps providing additional information that influenced the interpretation, or perhaps derived or related maps of, say, geological hazards or soil types. Thus, many interrelated facets of the geology are considered and reconciled in making a map.

The object-oriented methods of IT can organise the facetted information more flexibly than is possible within the more restrictive concepts of cartography. But appropriate organisation depends on a better understanding of the reasoning processes of geologists (Pshenichny, in press). The reasoning that connects the facets relies on the abstract concept of the general geoscience model (figure 3). This refers at all levels of detail to the three-dimensional location, form and configuration of the geoscience objects at the present day, to their observed and interpreted properties and composition, and also to their history of transformation throughout geological time, including the processes that created them and are crucial to their interpretation, and their relationships and interactions with other conceptual objects. In other words, it refers to what is there, what it is called, where it is, how it is arranged, what its properties are, what it is made of, where it came from, how it got there, and how we know. The general model does not exist in its entirety, because it would be too vast to comprehend or represent. But geologists' understanding of it (their conceptual framework) is essential to making sense of the individual aspects of geology shown on a map. The geologist must comprehend the overall picture, reconcile the facets, and project the results as a set of two-dimensional images on the map. Readers of the map may possess enough of the same background training, experience and information to enable them to reconstruct the wider picture from the fragments. But the computer does not.

Some fragmentation of map information is not necessarily inherent in the survey process but arises from the limitations of pen, paper and printing press. In contrast to this older mapping technology, computer spatial models need not be restricted to two dimensions, to fixed scales, to predetermined map sheets, to limited aspects of geoscience, to single modes of visualisation, nor to separate archiving or publication of maps, images, quantitative data, sketches and explanatory text. In principle, an object store could be created during the survey procedures, and the information subsequently filtered, combined, visualised and communicated in response to specific user requirements.

Good practical reasons remain for partitioning some facets of surveying activities. In studying the depositional model by measuring a vertical section, the field geologist might correct for distortions and small faults due to later structural deformation. There is little point in recording the detail of the corrections, as they are complex, irrelevant to deposition, and can be taken for granted. Their outcome is not a set of incontrovertible data records, but the interpretation of a skilled surveyor. The structural information that is merely a nuisance in developing the depositional model may be of significant interest for the structural model, but is then observed and recorded in a different light. Our skilled surveyor, therefore, does not integrate all observations in a single database, but while developing an unrecorded holistic overview, records and interprets separately various aspects of the geology from diverse observations, often at a single exposure.

In the context of a computer-based knowledge system, IT offers procedures for object-oriented analysis and design (Coad and Yourdon, 1991). The surveyor might view the things of interest as objects, which could play a role in various models. The surveyor, in this context, is observing and interpreting actual occurrences or *instances* of objects and the relationships between them, such as an outcrop of the Exe sandstone, and its spatial relationships with adjacent beds and the topography. An object instance belongs to a more abstract object *class* (referring to the Exe sandstone in general), and the classes are part of a hierarchical classification (the Exe sandstone being part of the class of sandstones, being part of the class of sediments, and so on). The same object may belong to various other hierarchies. For example, the observations at the outcrop might refer to a specific part of the Exe sandstone; the Exe sandstone may be part of the Wye formation; and so on up the stratigraphic hierarchy. Objects are concerned with things, models with processes and concepts. But objects may contain processes and models refer to objects, and the distinction is not important for immediate purposes.

One aim of the object-oriented design is to record the characteristics (properties, composition, relationships, behaviour) of object classes within the hierarchies, with the ability to inherit characteristics from higher levels in the hierarchy. Knowledge that a particular exposure is of Exe sandstone thus places it in an object class that carries information on a range of expectations, which may be supplemented or refuted by additional information referring specifically to the object instance. The class could thus provide a template within which the instance could be described. A knowledge system is also concerned with the relationships between objects, enabling the geologist to study fragments of a more general model, and establish the links that can provide an integrated picture. Information on the links and the likely behaviour of the objects may consist of short text explanations or microdocuments for use by the human reader. It might also, possibly in a mark-up language, encode software instructions identifying constraints on the behaviour and interaction of the objects within specific computer models or standard classes of model.

Explicit records of relationships ensure that surveyed objects are linked appropriately. For example, they might indicate which orientation measurements refer to the depositional model, which to the structural model, and at what level of completeness the two models are reconciled. As another example, individual measurements from a gravity survey might have no direct links to those from a geochemical survey. Separate maps based on the measurements, however, might show spatial patterns of gravity and geochemical anomalies that could be compared with one another and reconciled through a shared explanation, say of a hidden granite body, based on the general geoscience model. The link is at the level of the completed maps, not at the level of individual data points, and is indirect, through a shared explanation external to both. The data items could be regarded as object instances, related to other items in the same dataset and to the relevant map object at a higher level in the object hierarchy. Such patterns of reasoning are relevant to understanding and evaluating the conclusions, but are difficult to record on a conventional map.

It is conceivable that Geological Surveys will build digital geoscience spatial models within a standardised framework based on the general geoscience model, initially transforming our legacy of existing maps and subsequently transforming the surveying process. I think it inconceivable that they should not, but have no view on how many generations the full transformation will take. Convenient and robust equipment to use in the field or office for studying and amending a full computer-supported knowledge system would help, but is inevitably some way off. However, the development and testing of prototype conceptual frameworks is feasible on a much shorter timescale. A new framework and ontology, discussed later, is needed in order to know where to put things, where to find them, and what to do with them, as surveying moves towards more flexible representations and repositories of information, and away from the familiar, rigid and stylised framework of the conventional geological map. For the geologists of today, existing fragments from digitised maps could be placed in such a framework, moving to a new learning curve at significant cost and with limited initial benefits. For the geologists of tomorrow, the outcome (suitably adjusted to match changing expectations) could open up a new world of field survey.

Transformations and processes

Through the fog of our ignorance, we glimpse fragments of geology, which we can assemble only by reference to our expectations. Our understanding of the fragments is tied together by reasoning based on incomplete knowledge of the underlying processes of the general geoscience model. "[A] candidate for a bridging principle between empirical observations and scientific theories is the so-called Principle of the Uniformity of Nature, which Hume (1777, section 32) summed up in the phrase 'the future will resemble the past'." (Howson and Urbach, 1993, page 3). Hume's acquaintance, James Hutton, with his interest in geology, concluded that the present is the key to the past. Gaylord Simpson (1963, page 24) emphasised that geology has the characteristics of a historical science. "The unchanging properties of matter and energy and the likewise unchanging processes and principles arising therefrom are *immanent* in the material universe. They are nonhistorical, even though they occur and act in the course of history. The actual state of the universe or of any part of it at a given time, its configuration, is not immanent and is constantly changing . . . History may be defined as configurational change through time, i.e. a sequence of real, individual but interrelated events. These distinctions between the immanent and the configurational and between the nonhistorical and the historical are essential to clear analysis and comprehension of history and of science."

The approach is made more precise by the concepts of transformations and invariants. The immanent properties and processes are invariant (do not change) when transformed from one point in time or space to another. However, the configuration of objects and events does change through time, and of course affects the outcome of the processes. The concept of transformations, in particular spatial transformations, has wider relevance to geology. The cartoons drawn by d'Arcy Thompson (1942)

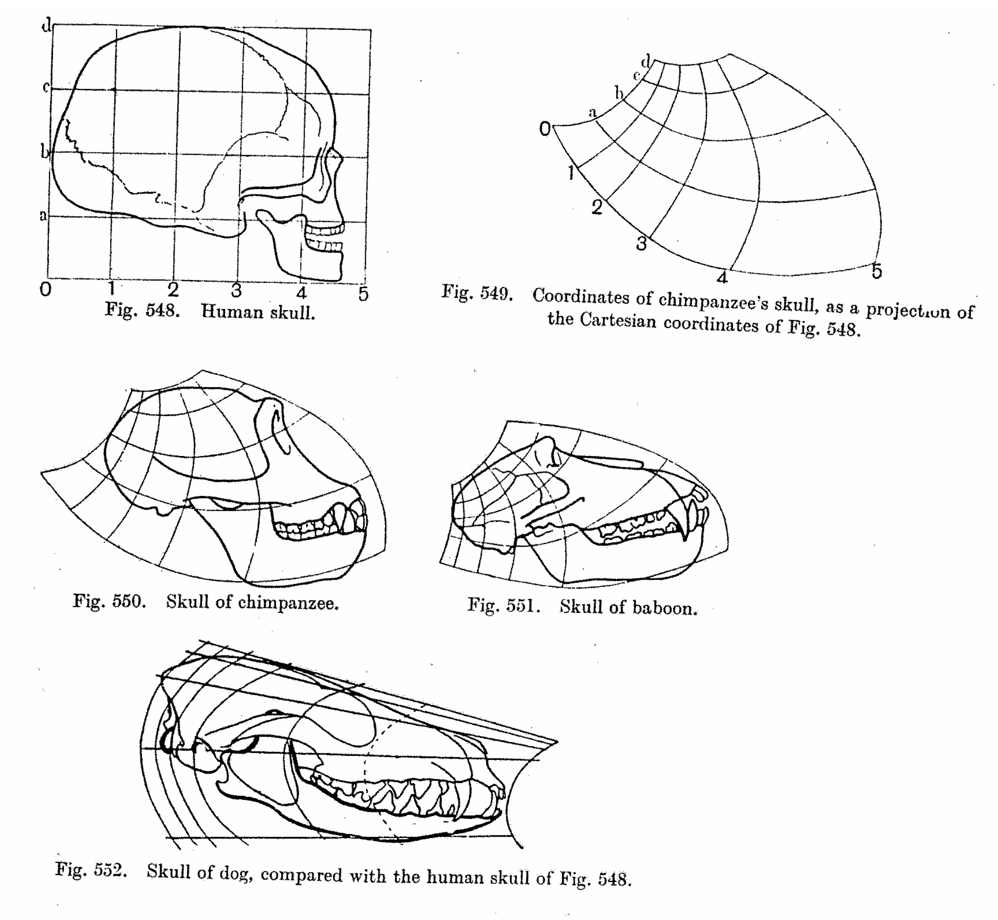

Figure 2: A flexible grid of coordinates illustrates the spatial transformations of anatomical features in related species. From d'Arcy Thompson, 'On Growth and Form' (1942), Cambridge University Press, © Cambridge University Press, reproduced with permission of Cambridge University Press.

illustrate the relationships between fossil diagrams by overlaying a grid that could be manipulated as though it were drawn on an elastic rubber sheet (figure 2). He avoided altering the spatial relationships of their anatomical structures and introduced no overlaps, gaps or rips in the grid. Thus he did not contravene what he knew of the immanent processes by which the skeletons had once operated. He did, however, move, stretch, shrink and turn the various parts of the rubber sheet to align the anatomical features of the different species. The transformed grid of squares clarifies the changes of form, accounting for the differences and revealing the similarities. Computer graphics can of course transform (morph) the specimen diagrams to match the same square grid. Stratigraphic or other surfaces that lack features for correlation can be transformed to normalise shape statistics.

Some basic geometrical transformations that change the position or form of an object relative to the chosen origin, scale and axes, are:

- translation – bodily movement of defined distance and direction
- rotation – turning of the object about an axis through the origin
- magnification – enlargement or miniaturisation of the object
- scaling – different magnification along different axes
- projection – reduction of the number of dimensions
- perspective projection – diminution of size with distance from the viewpoint, mimicking the effect of perspective

Complex transformations can be defined by a sequence of basic operations of this kind. For example, rotating an object about its long axis might involve: translation, to centre the axis on the origin; rotation; and translation back to its original position. The order of the operations affects the result, as you can demonstrate by thinking through a few geometrical examples. Complex transformations can thus be analysed into their basic geometrical components. Corresponding operations in matrix algebra bring about the same result, establish a link between geometry and algebra, and provide a means of computation. The development of computer graphics has made available a wide range of software for spatial transformation.

Looking back at the functions for surface fitting discussed earlier, the functions can be seen in this light as patterns of vertical translations of points from a horizontal plane. In manual contouring, the geologist is more inventive. The structural geologist might visualise rotation and stretching and squeezing (scaling) of the original beds when drawing a contour map. The sedimentologist contouring

a sand bar might have a particular form or shape in mind and mentally rotate and magnify it to see whether a larger or smaller version would better fit the measurement points. Slopes and curvatures might be visualised and transformed as well as relative locations. One can readily but imprecisely imagine the results of these transformations. They are less consistent, but may lead to greater geological understanding than the surface fitting by computer methods described earlier. They raise the prospect of exploring the forms of objects resulting from geological processes, with the help of spatial transformations on the computer, and feeding back suitably invariant aspects into the procedures of interpolation and interpretation.

The work of Harbaugh and others illustrates how geological processes can be simulated numerically on the computer. The SEPM Special Publication (Harbaugh et al., 1999) contains many examples. The majority of the simulations address a forward model, that is, they start with a specification of objects and their properties and derive possible results of processes operating on them. This throws light on the likely forms of rock bodies created by a defined process. But geologists have to address the inverse problem: the outcome is known, but what were the processes that created it? Inverse models address this question and are effective tools in fields like seismic exploration and downhole logging, where the investigators can to some extent control the conditions and the physical processes can be expressed mathematically. Given well-defined forward models, inverse procedures may be able to select the most appropriate of them, even in fields such as stratigraphy (Cross and Lessenger, 1999). But, as Chamberlin (1897) noted, inverse models may have no unique solution and therefore require consideration of multiple hypotheses. One approach to narrowing the range of possible solutions is to bring to bear independent sources of information: geophysical, geochemical, experimental, seismic, bio-stratigraphic and so on. The object-oriented approach is well suited to such an approach. But, in attempting to link models of past processes to present-day observations, there is another issue, considered next.

Complex systems

Many geological processes belong in the realm of complex, self-organising systems (described by Heylighen, in press). A system is a set of interacting parts that function as a whole, and its organisation is the arrangement of parts to promote a specific function, such as maintaining a particular configuration despite disturbances that would otherwise disrupt it. In a self-organising system, patterns may appear to arise spontaneously through the interaction of adjacent parts according to simple rules without any central control. This is a consequence of feedback effects, where input and output are connected by a causality loop, that is, each stage in the operation of the causal processes of the system is affected (intensified or inhibited) by the outcome of a previous stage.

The properties that emerge from the self-organisation are known as emergent properties, and because of them, the behaviour of the system as a whole cannot be explained by building up a picture from the behaviour of the parts. No wonder that, as suggested earlier, the geological surveyor takes a holistic approach. The reductionist mode of explaining or modelling complex emergent phenomena by reducing them to simple parts controlled by mechanical processes governed by the deterministic laws of physical science is inadequate. The linear equations familiar in physics, where effect is proportional to cause, may not apply. A complex system must be viewed as a single, coherent and organised whole, with properties that cannot be reduced to those of its components, and, because small causal changes can have large effects, the outcome of processes in the system may be inherently unpredictable.

Since the 1960s, according to Nicolis and Prigogine (1989, pages ix, 2, 8), a revolution in both mathematical and physical sciences has imposed a new attitude in the description of nature. The gap between simple and complex, between order and disorder is much narrower than previously thought. “Two disciplines have dramatically modified our outlook on complexity – nonequilibrium physics and the modern theory of dynamical systems (and the prevalence of instability). The new methods developed in this context lead to a better understanding of the environment in which we live.” “Today, wherever we look, we find evolution, diversification and instabilities…we are forced to acknowledge the existence of stochastic processes, those whose dynamics is nondeterministic, probabilistic, even completely random and unpredictable.” Methods of studying complex behaviour appear to be more familiar to geographers and ecologists than to most geologists. Nevertheless, Baas (2002) introduces his study of a complex system in coastal geomorphology with a useful account of chaos theory, fractals, attractors, self-organisation and self-organised criticality.

The work at ExxonMobil of Van Wagoner et al (2003) must surely cause many geologists to adjust their expectations. Their investigation, of unprecedented extent and detail, looked at the shape of sand bodies ranging in length from a few centimetres to 1000 km, created by many processes in many depositional environments including the laboratory flume. “From the shape alone it is impossible to determine the size or depositional environment of these bodies. Thus, shape is independent of scale and place of deposition… empirical and statistical similarities in shapes indicate that these bodies were deposited by a common physics. The physics at the local instantaneous scale are the well-established laws of fluid and sediment dynamics. However, these dynamics do not explain the cause of the global organization of the bodies observed in nature. A deeper, more encompassing explanation is required. We believe that the explanation can be found in nonequilibrium thermodynamics and energy dissipation… All open systems (i.e. systems through which energy and matter are transmitted) evolve toward increasing complexity with time … as these systems form dissipative structures to minimize gradients…” They conclude that: “the sedimentary rock record is built of scale-invariant hierarchies of sedimentary bodies… similar in shape and property distribution… evolv[ing] along a well-defined pathway… independent of depositional environment.”

The results obtained by these investigators suggest that the complex systems approach, and the search for common physics, will lead to equally significant findings in other fields of geoscience. Their fractal model of deposition from jet-plume pairs raises issues of how to use knowledge of the resulting shape to interpolate or fit surfaces, and of how we identify and model the boundaries of sedimentary bodies. The mathematical process generating a fractal typically starts with a simple geometrical object (the initiator). Its form is replicated by a process (the generator), which incorporates smaller versions of itself at appropriate points on the initiator, then even-smaller versions on each smaller version, repeating at ever-diminishing size until stopped for practical reasons. In this case, a flow of water carrying sediment initially erodes a channel, but, as the flow spreads and slows, it preferentially deposits the coarser material. The deposited material builds up and tends to split the flow, creating two channels, which in turn split, branching out like a tree. The location and configuration of the feedback effects that result in channel splitting are in effect unpredictable, even if the immanent processes are known. Downstream from each successive branch, smaller and finer-grained sediment bodies are deposited, with a gradual overall transition from coarse sandstone to fine mud.

The fractal model (like any other model) is a greatly simplified view of reality, but one that extends our understanding, and leads to insights of scientific and economic value. It calls for observations at many scales and levels of detail, and mapping sharp lithological boundaries is not adequate for visualising its form. To represent the spatial objects resulting from such models, we need to go beyond map-bound concepts, such as: “The formation is the fundamental unit in lithostratigraphic classification… No formation is considered valid that cannot be delineated at the scale of geologic mapping practiced in the region when the formation is proposed.” (North American Commission on Stratigraphic Nomenclature, 1983, page 858). Fractal models have been applied to other aspects of geology, such as ore emplacement (Turcotte, 2002) and rock fracturing (Paredes and Elorza, 1999), and again the results do not lend themselves to visualisation on a fixed-scale map although they fit well into an object hierarchy.

The local systems that a geologist is likely to study are set in the environment of a hierarchy of more general systems, including the general geoscience model. The tendency of the system to create organised processes and objects can be described in terms of attractors – preferred configurations or positions in state space, such that a system starting in another state tends to evolve towards an attractor. The pattern of attractors can be visualised in terms of a fitness function measuring the degree of suitability or desirability of that state. The system adapts to its environment by optimising its fitness function. Present-day spatial objects reflect the stage that the adaptation has reached. On a geological time-scale they are far from equilibrium and still in a dynamic process of change, even although they appear static on the human time-scale of geological survey. Indeed, we may recognise objects frozen in time at different stages along evolutionary paths towards attractors at various levels in the hierarchy, such as temporary lakes and lacustrine sediments in the evolution of a river valley. The distinct environments and separate scales on which major processes operate (such as sea-level changes and tectonic, climatological, volcanic and sedimentological processes) complicate the study of geological systems. Nevertheless, the concept that hierarchies of subsystems evolve towards attractors offers a good match to the hierarchical classifications by which geologists make their subject manageable.

Nicolis and Prigogine (1989, page 15) suggest that nonequilibrium enables a system to avoid disorder, and to transform part of the energy communicated from the environment into an ordered behaviour of a new type, the dissipative structure, a regime characterised by symmetry breaking (such as channel splitting), multiple choices, and correlations of a macroscopic range. They describe these correlations (page 13) as statistically reproducible relations between parts of the system. The semivariograms of geostatistics and kriging would seem to be an example of a simplified model of such a correlation. They also point out (page 14) that: "Far from equilibrium, that is, when a constraint is sufficiently strong, the system can adjust to its environment in several different ways… Chance alone will decide which of these solutions will be realized. The fact that only one among many possibilities occurred gives the system a historical dimension, some sort of 'memory' of a past event that took place at a critical moment and which will affect its future evolution." Complexity (pages 36-37) is concerned with systems in which evolution, and hence history, plays or has played an important role in their observed behaviour. An essential feature is the ability to perform transitions between different states. "Everywhere we look in our environment at large we find relics of the past, which, far from being fossils, are actively preparing the history yet to come." This calls to mind the distinction drawn by Simpson (1963), mentioned earlier, between the configurational and the immanent.

The study of complex systems seems, reassuringly, to track ideas that have long been part of geological understanding. It does so in a framework that recognises the limitations of reductionist methods and the uncertainties of prediction, that requires an integrated top-down view, clarifies concepts and carries new insights across many applications. It demonstrates how a hierarchy of processes can create recognisable patterns across a range of scales. The details of the overall process may be obscure and the outcome partly unpredictable. Nevertheless, patterns can be recognised, analysed and to some extent explained. Uncertainty has always been recognised in geoscience investigation, and predictions are accurate only in a statistical sense. Reductionist methods, and direct, inverse and many other types of model have proved their worth in understanding the Earth and predicting its properties. Complex systems can be seen as another concept that helps us to understand more fully the processes, the uncertainties and the occasional big surprise. Their products are unsuited to the cartographic model.

Model diversity

A Geological Survey map aims to provide a reasonable and realistic interpretation that weighs up views from many sources, and typically reconciles them to present one single depiction of the overall geology within the area. The word reconciled is used here in the sense of Kent (1978, page 202): "By reconciliation, I mean a state in which the parties involved have negligible differences in that portion of their world views which is relevant to the purpose in hand." While surveying, geologists must force a concoction of ambiguous, contradictory, subjective and vague ideas from many sources into the precise and rigid geometry of the map. The single image of constant scale is overloaded with diverse information. Consequently, the map user has problems in: determining the sources, evidence and reasoning; recognising uncertainty, scenarios and multiple hypotheses; identifying the salient points, and the sampling schemes or design of the investigation; reconciling overlapping maps of related topics; adjusting the map to a new or evolving interpretation; and reinterpreting the map to meet specific objectives.

There is a strong case for a Survey providing (as at present) default options of standard views matching consistent, widely accepted, expert opinions. But where it is practicable, a range of other options, including more specific information and alternative selection, interpolation and visualisation procedures, could help users to meet their own particular requirements. A knowledge-based system might offer a sequence of standard views, corresponding to stages of interpretation and generalisation. This implies that the system framework should contain a view of the reasoning procedures to provide a reference to which the stages could be related and linked. Identification and description are a first stage of interpretation. A second stage of interpretation is interpolation to fit lines and surfaces to fill the gaps between points that are known from direct evidence. Further stages of interpretation involve generalisation to look at larger features at a smaller scale and coarser resolution, and to feed back revisions into broader views of the regional geology. The stages might broadly correspond to conventional products: field notes, survey-scale maps, generalised smaller-scale maps. Indeed, such a match is necessary to incorporate legacy information at local, regional, national and global levels of detail. Different objectives call for access at different levels of detail.

Users, because they should know most about their own objectives, might also benefit from an option to work with information that has not already been fully reprocessed for a different purpose. For the same reason, a Survey might maintain a secure long-term digital archive of source material (see National Library of Australia, 2004), including material from outside sources even where it is not reconciled with the Survey's viewpoint.

In creating or using a map or spatial model, our understanding is based on fragmentary evidence, and the fragments themselves may stem from diverging lines of thought. For instance, the multifaceted survey discussed earlier is made more complex by the range of objectives that influence the way the evidence is interpreted and the results are visualised. Examples of objectives are:

1. Providing a generic view, as a basis for more specific interpretations to meet explicit objectives. Example: a Geological Survey map.
2. Visualising a surface that takes a neutral view of the evidence. Examples: to suggest and guide exploration of alternative interpretations; to resolve a dispute about, say, allocation of oil reserve estimates between adjacent leases.
3. Deciphering the sequence. Example: working out the relative stratigraphic positions of the exposed beds, and establishing the local stratigraphy.
4. Testing initial interpretations. Example: a model to demonstrate that identifications and interpretations being developed in the field conform to a plausible geological pattern.
5. Estimation and prediction. Examples, to assess: the depth at which a well is likely to reach a particular formation; the gold reserves in a deposit; the amount of folding in an area.
6. Testing predictions. Examples: matching simulations of geological processes against the observed geology; testing predictions from other topics, such as geophysics or geochemistry, against the known geology or vice versa; predicting values to compare with new or withheld data.
7. Exploring best- or worst-case scenarios. Examples: estimating the largest feasible oil reserves in a lease before relinquishing it; estimating, from test boreholes, the maximum risk of significant faulting before mining a coal seam; assessing the geological threats to a radioactive waste repository.
8. Generalisation. Example: to obtain a regional overview by simplifying and combining salient points from detailed local models.
9. Recognising or detecting pattern. Examples: visualising the spatial characteristics of a surface to throw light on its origin; testing whether a surface with specific characteristics, such as a dendritic pattern of river valleys, could fit the available data; looking for characteristic patterns to detect faults.
10. Modelling abrupt changes and systematic changes. Examples: detecting, locating and representing faults, unconformities, on-lap, and off-lap.
11. Separating patterns of different scale or type. Examples: examining deviations from the regional structure to identify local anomalies that might indicate, say, data errors or wrong identifications, geological hazards (such as sink holes), or economic opportunities (oil-bearing anticlines, ore deposits); separating the effects of faulting and folding; separating depositional and structural features.
12. Categorising types of surface. Examples: establishing classes of deposit with similar properties (cluster analysis); characterising surfaces from high-energy and low-energy depositional environments (discriminant function); extending the classification to new areas (discriminatory analysis).
13. Comparing, correlating and reconciling surfaces. Examples: comparing surfaces from seismic surveys and well picks and combining information from both sources for a more accurate view of the structure; reconciling gravity and geochemical surveys with a geological model; relating the spatial variation of a property, such as porosity, to the variation of other properties like grain size or position within a basin.
14. Explaining the origin of surfaces. Examples: relating the observed surfaces to a conceptual model of their formative processes; examining the pattern of folding to throw light on past stress patterns and their variation through geological time.

There are many business settings that give direction to geoscience surveying, many different objectives, and many models for interpreting the results. Methods of modelling surfaces or mapping the boundaries of land-surface geology must be suited to the requirements. The emphasis placed on different sources of information varies from one application to another. Even basic observations reflect a particular viewpoint. No map or model can meet all the conflicting objectives, each of which may call for a different approach. Diversity enables ideas to evolve, but forces difficult choices on an information community, such as a Geological Survey, attempting to offer a broadly relevant view based on widely accepted procedures understood by all users. The object-oriented approach might enable a Survey to offer greater flexibility by supporting reuse of objects within various models related to different objectives. Metadata for the object-class hierarchy could record the suitability of an object for various classes of model. For some models a completed map might be an appropriate input object, other models might require access to objects recording raw observational data, yet others would have to be reconciled by the user.

A computer-based knowledge system should be able to model the content of surface and subsurface maps within the same conceptual and geometrical framework, giving a consistent view of the geological objects on and below the ground, with interpolation between outcrops matching interpolation between boreholes. It might attempt to identify, estimate and record the uncertainty in individual objects arising at various stages of the reasoning process. This raises general issues about the nature and causes of uncertainty that are hinted at on existing maps by broken lines and confidence intervals on contours. To resolve them, a conceptual framework is needed to identify and record the stages of geological reasoning, and the uncertainties associated with each.

Framework and ontology

The obvious starting point for the design of a knowledge base for geological survey is the analysis of present-day methods. Extensive studies relating to digital maps and mapping are indeed being undertaken, as well as studies of data management (POSC, 2004) and data transfer with mark-up languages (CSIRO, 2004). Development of the Canadian Geoscience Knowledge Network can be followed at GISWorld (2004). Numerous open-files reports on digital cartography are available in the U.S. Geological Survey catalogue. For example, a subject search for ‘digital cartography’ at http://usgs-georef.cos.com/ found 758 records, and ‘geologic map data model’ recovered 16 records. International links to geological map standards and databases can be found at USGS (2004). The papers imply, but do not resolve, the need for a more comprehensive design where map information can be tied to its wider context (Brodaric, 2000).

A specification of the underlying concepts – the *ontology* for the system – must obviously fit within the framework for science and knowledge systems in general. It must also take into account the ideas, practices and legacy of the conventional knowledge systems on which the study of geology is based. The conventional system depends on subtle understanding and complex interconnections (and now hyperlinks) that defy diagrammatic representation. At its heart is the general geoscience model mentioned earlier, organised from fragmentary evidence through reasoning and reconciliation, and constantly adjusted (by communal endeavour) to make sense of new ideas and observations. It contains shared ideas of the structure, concepts, procedures and vocabulary that enable geologists to communicate and develop a communal understanding of their subject. It is, more or less, the geoscience paradigm

(Loudon, 2000, part K, page 101), which is part of the background thinking of every geologist.

The knowledge of geoscience, within our present-day horizon, is entirely human knowledge, but the system reflects the technology that supports it. Its structure must therefore evolve to match the transformation of that technology from pen, paper and printing press to modern IT. In geological surveying, the computer can help to record, link, analyse, summarise, manipulate, store, browse, retrieve, transform, and communicate the information. The software can potentially support: desktop browsers and portable field survey stations with GIS, database and hypermedia capabilities for user selection and filtering of information; spatial modelling systems for simulation, analysis, integration, interpolation and generalisation; and visualisation systems. A framework is essential to enable these software components to work together, and for the human user to maintain a coherent view of what is going on. This framework is placed at the top level of figure 3. The diagram shows the geoscience knowledge system interfaced with the real world. The full interconnections within the general geoscience model, held in the human mind or as hyperlinks, are too complex to show in the diagram, but some extracts relevant to an IT implementation are shown as boxes.

The framework should be relevant, overall and in detail, whether or not the system is computer based and regardless of the information type (such as maps, sections, sketches, diagrams, images, text descriptions, text explanations, data, video clips, hypermedia) in which the information is represented. It must be designed to tap into, and make full use of, human knowledge and experience. Geologists must therefore get involved in determining the structure of their knowledge base; in defining their reasoning and its associated uncertainties; and in improving the representation of their spatial insights. But they must be wary of their lingering metaphors, which can inhibit progress by drawing attention away from the more radical approach needed to realise the potential benefits of more powerful technology. In designing the framework for a geoscience knowledge base, we need to look ahead, explore new possibilities and adjust our expectations. The framework must be flexible to handle existing legacy information as well as current and foreseen procedures, adaptable in order to adjust to its future evolution while carrying forward its customer and information base, and open-ended (extensible) to cope with unforeseen developments.

To meet future needs, the structure must support information for many topics, from many sources, at any level of detail. It must index this information and relate it to one uniform view of the underlying geology that aims to tie observations and predictions to spatial co-ordinates and thence to the real world. The framework is particularly concerned with those aspects of the paradigm where we can find or foresee technological solutions. We can think of the information containers generically as objects, rather than specifically as books, journal articles, maps, or datasets. The concepts and processes linking these objects can be regarded generically as conceptual models and process models. This object-oriented view matches the procedures of IT to patterns of human thought and offers a flexible and productive framework, at least until better approaches emerge.

The shared geoscience paradigm is a compatible part of the wider paradigm of science as a whole. It underlies the activities of geoscience in general, and thus of geological survey. A core surveying activity is collection of information in the field, by observation, measurement and recording. The aim of surveying is not normally to alter the paradigm, but rather to create or modify a local interpretation of the geology. It is based on the general geoscience model, which provides the underlying reasoning by placing observations in the context of processes and historical configurations. Surveying operates at the interface between the science and the real world, assigning values to geological properties at locations on or within the Earth's crust. Interpretative spatial models, traditionally in the form of maps and map explanations, are central to the process of geological survey. Surveyors work in detail on their local model, with a generalised view of the remainder of the general geoscience model.

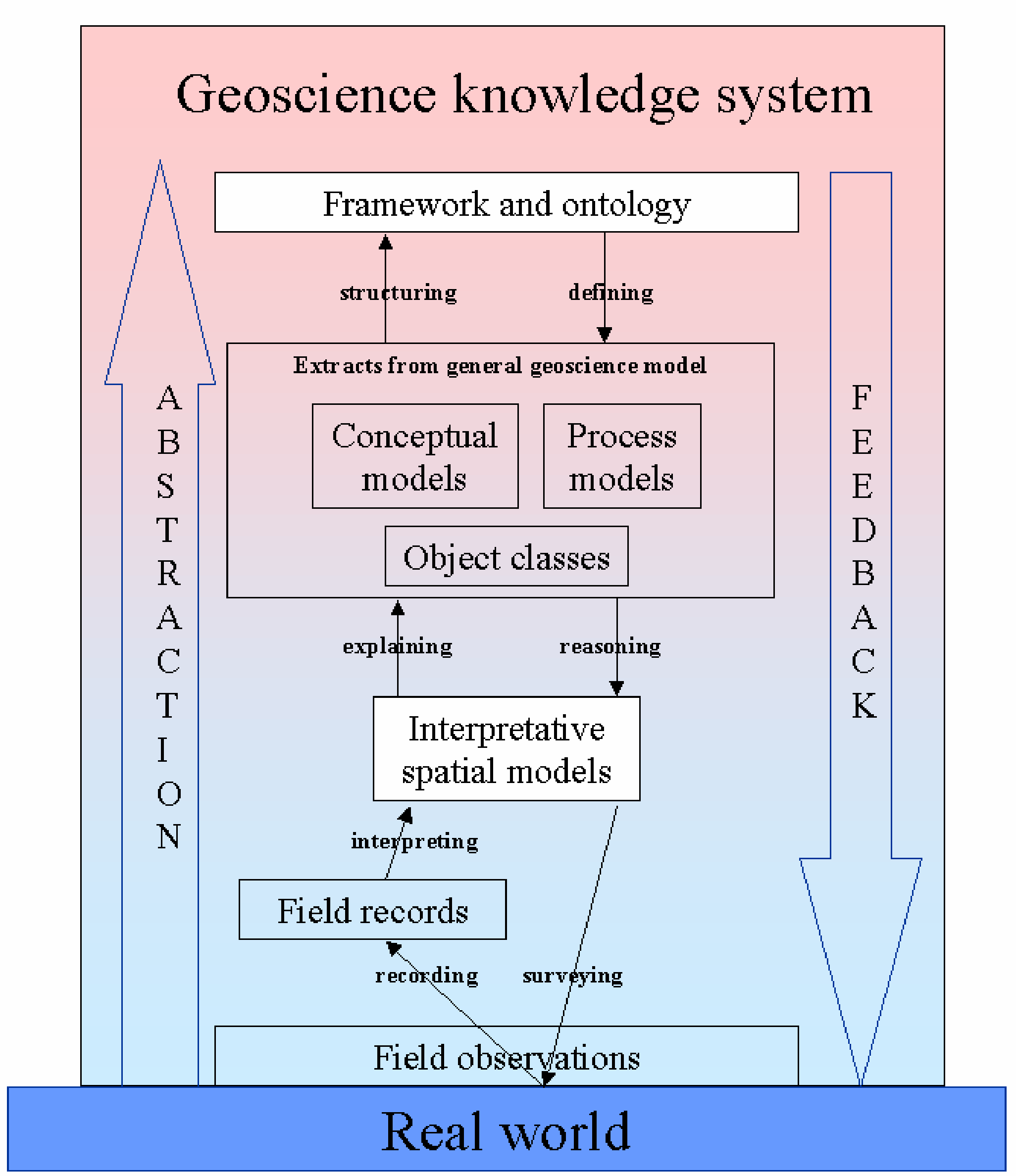

Figure 3: IT support for a knowledge-based system for geological survey.

Their spatial models are therefore highlighted in figure 3. Reasoning from the general model, modified by the surveyors' training and experience within the context of the local interpretation, guides the surveying procedures and observations that extend, test and modify the spatial models. In an IT context, the interpretation might be supported by reasoning processes based on feedback from models of geoscience concepts and processes, together with the accumulated abstraction of knowledge of the properties, relationships and behaviour of the object classes (figure 3).

The initial transient field observations are an abstraction from the real world and even the most basic field observations are thus selective. Outcrops give answers only to questions that are put to them (Harrison, 1963). The salient points are abstracted from the answers, recorded as field records, and further abstracted to extend or amend the interpretative model. The results from the local interpretation in turn extend and may correct the descriptions and explanations of the general geoscience model. Objects are containers that store information about sets, models and relationships and carry it across the interfaces from one part of the system to another. Figure 3 shows four related bodies of information within the geoscience knowledge system: framework and ontology; extracts from the general geoscience model; interpretative spatial models; and field observations and records at the real-world interface. The records in each part loosely correspond with areas of human memory (semantic, episodic, spatial, and short-term), and thus to geologists' thought processes (Loudon, 2000, part J). The diagram is specific to field survey, and other aspects would be pulled out from the general model to replace these boxes for, say, palaeontological or mineral exploration studies.

The framework can guide the orderly development of the system by describing a formal structure within which the objects, the models and the various processes and procedures of surveying can interact. Indexes can assist users to locate the information they require. The framework might be represented by diagrammatic depictions of logical relationships, such as entity-relationship diagrams. "But logic itself has no vocabulary for describing the things that exist. Ontology fills the gap: it is the study of existence, of all the kinds of entities – abstract and concrete – that make up the world. It supplies the predicates of predicate calculus and the labels that fill the boxes and circles of conceptual graphs. The two sources of ontological categories are observation and reasoning. Observation provides knowledge of the physical world, and reasoning makes sense of observation by generating a framework of abstractions called *metaphysics*.

"A choice of ontological categories is the first step in designing a database, a knowledge base, or an object-oriented system. In database theory the categories are usually called *domains*, in AI they are called *types*, in object-oriented systems they are called *classes*, and in logic they are called *types* or *sorts*. Whatever they are called, the selection of categories determines everything that can be represented in a computer application or an entire family of applications. Any incompleteness, distortions, or restrictions in the framework of categories must inevitably limit the generality of every program and database that uses these categories." (Sowa, 2000, page 51)

Sowa points out (page 52) that: "Philosophers usually build their ontologies from the top down. They start with grand conceptions about everything in heaven and earth. Programmers, however, tend to work from the bottom up. For their database and AI systems, they start with limited ontologies or *microworlds*, which have a small number of concepts that are tailored for a single application." Geological Surveys have long set their local findings in their regional and global context. Perhaps, therefore, it is timely for the philosophers and programmers to get together: for geoscientists – who have long developed classifications and nomenclatures for the subject as a whole and for details within the resulting specialised topics – to explore new information containers, adjust their ontological expectations, and start to bring the resulting framework into the global knowledge system that is being transformed by computer support. In outline, the overall scope of the geoscience knowledge system might refer to the process of abstraction: from the real world to geoscience information (by collecting and interpreting information through observation, measurement, experimentation, explanation, and the application of general scientific knowledge) to recording the results in a suitable form for communication. The aim of communication is to meet the needs of users. The information suppliers broaden the content of the knowledge base: the users narrow the focus to filter and extract knowledge to meet their own specific requirements. Their ability to do so rests on a clear framework and ontology, accessible from field and desktop browsers.

Reasoning, uncertainty and field recording

To fulfil its role as an interface between the real world and the geoscience knowledge base, geological surveying must connect to the underlying evidence and reasoning. This is also necessary for at least two other reasons. The first is that, in a scientific product, it should be possible to assess and verify conclusions by tracing them back to the evidence. Scientists should be able to explain their procedures, and their colleagues should be able to follow their reasoning and reproduce the results. The second reason is to clarify how far the knowledge base must be amended to respond to small changes in the interpretation or observations. For instance, erosional features might be discovered at a supposedly conformable junction, or a supposed outcrop might turn out to be a large boulder. The design must ensure that limited adjustments of this kind avoid needless disruption to the system as a whole, but that where necessary their knock-on effects can be propagated through appropriate levels of interpretation. Equally, two-way hyperlinks should help human experts to identify the consequences of a change to the metadata (such as a revised definition of a stratigraphic unit), or help to explore the outcome of alternative hypotheses (figure 1). Ideally, therefore, it should be possible to trace the process of abstraction from the final summary back to the details on which it was based, and vice versa.

In the earlier section on multifaceted survey, we noted some features of geoscience reasoning. It has a strong historical element, in that it is studying the outcome of events and processes that took place in the distant past and are not necessarily observable now. The reasoning has a strong spatial element, for much of its scientific value, and more of its practical value, is related to its ability to explain and predict the present location, spatial relationships and form of geological objects and their properties from explanations based on conceptual configurations of events in the past. The past processes can be determined only from their present-day outcome, and this inverse reasoning is essential to building and relating the fragments of the general geoscience model that give coherence to geological reasoning. It must deal with complex systems. It has a strong probabilistic component, for our procedures are imprecise, our knowledge of the configuration and outcome of historical events is incomplete, much of our knowledge is tacit or unrecorded, and the behaviour of complex systems is inherently unpredictable. Furthermore, the outcomes that we observe are the consequence of many interacting processes with effects that cannot readily be disentangled. These include deposition, compaction, phases of folding and faulting, metamorphism, erosion and weathering. For practical reasons we have to piece together separate fragments of knowledge relating to different aspects of the geology. Error and uncertainty arise not only from the geological processes but also in our concepts (Brodaric and Gahegan, 2000) and in our procedures, such as identification, sampling, measurement and reasoning (Bardosy and Fodor, 2003).

In these circumstances, objective statistical analysis and mathematical simulation are unlikely to provide an accurate assessment of our overall uncertainty. There has recently been a strong revival of interest (BIPS, 2004) in Bayesian statistics, which may have a significant future role in geological reasoning. “According to the Bayesian view, scientific and indeed much of everyday reasoning is conducted in probabilistic terms. In other words, when evaluating an uncertain claim, one does so by calculating the probability of the claim in the light of given information… inductive probability treats the probabilities of theories as a property of our attitude towards them; such probabilities are then interpreted, roughly speaking, as measuring degrees of belief. This is called the subjectivist or personalist interpretation. The scientific methodology based on this idea is usually referred to as the methodology of Bayesianism because of the prominent role it assigns to a famous result of the probability calculus known as Bayes’s theorem” (Howson and Urbach, 1993, pages 1, 11).

The strength of Bayesian methods lies in their probability calculus, in effect allowing geologists to deal with their uncertainties and calculate the overall probability of a particular outcome, as a gambler or a bookmaker might systematically calculate betting odds. After all, every time geologists decide where to place a line on a map they deal with uncertainty and may attempt to take into account all the relevant information and weigh up their beliefs in the various aspects that bear on their decision, perhaps sketching a tentative line to see how it would match their holistic view, and modifying it if need be. A practical problem in the Bayesian approach is providing numerical values for estimating the odds or degrees of belief. This is not necessarily insuperable, for numerical experiments (Harbaugh et al., 1999) provide probabilities of the outcome of experimental processes in the required quantitative form, which might help geologists to develop a feel for estimation. In an object-oriented environment, the behaviour of object classes (for example, during interpolation) can be recorded, and the application to specific instances could be largely automated. The probability calculus leads to quantitative conclusions about outcomes that could be tested statistically against the real world.

Bayesian methods may or may not provide an acceptable quantitative foundation for estimating, discussing, reconciling, combining, adjusting, and recording degrees of belief throughout the reasoning process. One can think of good reasons to suppose that they might improve on existing alternatives, and that further investigation of their potential in a geoscience knowledge base might therefore be worth serious consideration, initially as an academic exercise. Geological explanations are generally expressed in words, not in numbers, and it has been suggested that fuzzy sets and logic could help to make this accessible to computer analysis (Nordlund, 1999). Applying comprehensive quantitative analysis of uncertainty to geoscience reasoning as a whole may be beyond our short-term capabilities, but is a direction that could bear further exploration.

The system framework and process of reasoning does not depend on quantification, however, and the importance of the legacy from earlier workers implies that the framework should provide a long-term structure that can accommodate explanatory objects or microdocuments in many forms: such as written text, sketches, maps, quantitative explanations, video clips or multimedia. In the spirit of a mark-up language, the microdocuments may or may not include information directly usable by computer software. This in turn points to the benefits of a system that can accept composite objects of any extent. For example, an existing map may have to be considered as an indivisible object if the symbols and objects portrayed have no useful significance when separated from the map, or if the map is in, say, its original paper form with no individual access to its components. The general reasoning process can then involve the object as a whole, even although its components, because of the way the information was assembled, cannot be analysed individually in an external context. A knowledge system might thus develop as a structured component of a more general hypermedia repository, connected internally and externally by hypertext and a shared browsing system (BGS, 2004). It must be based on human knowledge and reasoning, complementing it with the storage and processing power of IT. It should support the most effective representations of the underlying knowledge available at the time, with the flexibility to cope with other representations, and must be extensible to cope with new ideas as they emerge. The way we organise and represent our knowledge (Sowa, 2000) is central to its development.

Within a computer-supported knowledge system, the records are elements in a hypermedia knowledge repository, and spatial models (see next section) take the place of the map. Unlike the framework and ontology considered in the previous section, the spatial models deal with actual occurrences (instances) of objects, rather than the general characteristics of object classes. They are thus at a lower level of abstraction and closer to the real world, being linked to it through spatial coordinates. But they are set in the context of the general model, which guides the observations. This chain of processes of abstraction and feedback inevitably introduce errors and uncertainty. For example, the reasoning processes bring uncertainties to the interpretative spatial models. The surveying procedures can be designed to clarify the uncertainty, but are also prone to error. A hypermedia system might aim to clarify the dependencies and associated uncertainties throughout the system.

Conventional field recording and interpretation of geology rely on documents, such as field slips and survey-scale maps, of similar form to the finished geological map. Field observation may follow predetermined rules, as in many geophysical and geochemical studies, with the geological reasoning preceding and following the fieldwork. In other cases, such as detailed geological mapping, the interpretation and the procedures develop together as the study proceeds and more is learned. It is appropriate to record the chain of reasoning as it develops, in the field if appropriate. When developing technology enables surveying to be more fully supported by portable IT linked to an extensive knowledge base, field recording might generate a hypermedia stream of observational records, presumably in a mark-up language, such as XML.

The field records could make continual reference to the framework and ontology on the one hand and to relevant aspects of the interpretative spatial models on the other. The framework and ontology could provide templates listing the expected properties and behaviour of object classes in order to guide the observational and interpretational procedures and to standardise the vocabulary, and give the current view of the structure of the general geoscience model that relates the fragmentary interpretations. The interpretative spatial models could provide an up-to-date indication of the implications of new observations for the existing views of processes, events and configurations, and thus of the requirements for further investigation. Surveying is based on existing ideas about the interpretation, and the hypermedia stream of observational records could record and explain how the new observations extend and modify previous interpretations. Mark-up of the records could identify salient points for editing and generalisation of the interpretation and reasoning.

Spatial modelling and interpretation

Earlier comments pointed to the tendency to imitate manual products in the initial developments of computer methods of cartography and surface fitting. Because of the need to carry forward the customer base, novelty and the lingering metaphor go hand in hand. But comparing geological surfaces to powers of east and north coordinates, or even kriging on the basis of semi-variograms, does not seem to quite match patterns of geological thought. Recent developments provide mathematical models (and potential computer support) that more readily embody geological insights while opening up unfamiliar views. A realistic aim is therefore to build computer graphics systems that allow geologists to express geological insights based of their understanding of geological processes and spatial relationships. Geologists might, for example, think of spatial objects in terms of position, slope, form and shape, and might visualise geometrical transformations like moving, scaling, rotating, or stretching geological objects. Even notions of fractal processes creating intertwined objects of various sizes from microfold to nappe, or cross-bed to delta, must seem familiar to geologists. Expressing these ideas mathematically (as opposed to expressing mathematical ideas in a geological setting) might bring greater rigour to geological reasoning as well as making computer applications possible.

Jean-Laurent Mallet leads the gOcad project (Gocad, 2004), a consortium that develops software for spatial modelling. The project started in 1989 and was based on a strategy of discrete modelling of natural objects. Points in 3d space define the geometry of any object, and the links between the points model the topological relationships, that is, those invariant under rubber-sheet transformations. The links might, for example, join adjacent points as triangular facets on a surface. Each triangle can be fitted with a curved surface that can be blended into the adjoining triangles. This is a piecewise approach, in the sense that local patches of the surface, triangles in this case, are each fitted with their own functions, modified to merge smoothly into adjacent patches. The approach is familiar from computer-aided design, finite element methods, and GIS. It has been adopted in various geological projects, generally using Delauney triangles (Bonham-Carter, 1994). The same approach can be extended to lines, surfaces and volumes.

Mallet (2002) provides an advanced mathematical exposition of the methods developed in the gOcad Project and examples of many applications that are relevant to geoscience. He shows how the parametric methods of computer-aided design can provide functions that relate the location of a point in three dimensions (x,y,z) to two arbitrary parameters (s,t) that might be thought of as the coordinates of a plane. Instead of one function $z=f(x,y)$, there are separate functions, f, g and h. Thus: $x=f(s,t)$, $y=g(s,t)$ and $z=h(s,t)$. This makes it possible to handle multi-valued surfaces that are penetrated more than once by the same vertical borehole because of overfolding or reverse faulting. The piecewise approach makes it easier to break surfaces along fault planes or unconformities, and to match, adjust or break slopes on either side. It allows local recalculation of the surface to include new information, such as another borehole, without disrupting the remainder of the surface. Methods from differential geometry can map directions of maximum and minimum curvature (principal curvatures) analogous to fold axes and axial planes. Surfaces and volumes can be interpolated to conserve the original areas or volumes as in balanced cross-sections.

Piecewise methods allow us to think of spatial models in other ways that correspond better to geological observation. Consider three adjacent values, of a property such as elevation, arranged as a triangle on a surface such as the top of a subsurface horizon. It would be possible to fit a fifth-order polynomial (quintic) surface to the three points, defining a curved triangular segment of the surface. This could be interpreted as a function like the fitted surface mentioned in an earlier section. But it could also be interpreted geometrically as in finite element methods, described for example by Strang and Fix (1973). If the elevation, slope and curvature could be defined at each of the three points, they would provide all the information required to define a quintic function that would precisely fit the values. If the same procedure is applied to the adjacent triangles, using the same values at the shared nodes, a smooth surface is generated with no sudden breaks in elevation, slope or curvature. The lines lying on the surface and marking the shared boundaries between the triangles are similarly smooth, being represented by third-order polynomial (cubic) curves. Lines, such as the intersection of the horizon with the ground surface, and the surface itself can thus be represented in a consistent and compatible manner.

It is feasible, though computationally extravagant and not necessarily desirable, to build a picture of smooth surfaces and boundary lines from information about elevations, slopes and curvatures that are known or estimated at points. For present purposes, however, the relevance is that the method allows us to look at our model in a more natural way. Location, slope and curvature can readily be sketched or visualised and can be measured or estimated by eye in the field. Knowledge about shape and form, perhaps based on invariant geological characteristics of the surface known from other occurrences and experiments, could be used in filling the gaps (interpolating) between observation points. Even as a thought experiment, that raises geological issues about filling gaps. They are no doubt obvious, but do not seem to have been resolved satisfactorily, perhaps because the traditional map offers no solution, so here are a few more speculations.

In filling the gaps between known points, with a pencil or by computer interpolation, we tend to draw smooth lines or surfaces. We know the surfaces are in fact rough, but as always the selected model differs from reality. The model is smooth, perhaps to avoid the visual impact of a discontinuity that might suggest a feature of geological significance, such as a hinge-line or fault, for which there was no evidence. In the model, we do not attempt to include the small irregularities we can see in the field, but smooth them over to view the bigger picture, pointing to larger features, such as major folds, faults or steady gradients, which could have been concealed by the noise of small irregularities. But our observations are unevenly distributed in space, and we are therefore looking at the spatial object (such as a line, surface or solid body) with varying resolution.

Location, slope and curvature are features we can recognise and measure in the real world. We can view features at various distances, and so are accustomed to the idea of moving away in order to smooth irregularities by eye, and see a broader pattern. We look at slides under a microscope, look at hand specimens, stand back and look at the outcrop, climb the hill to view the landscape as a whole, and perhaps examine air photographs and satellite imagery. Evolution and a lifetime's learning help us to visualise and interpret patterns in two or three dimensions, and to relate patterns at different scales. It makes sense to design our models and computer visualisation procedures to take full advantage of our inherited abilities. Even variable resolution is something that our eyes and brain can handle, as we know from looking at a landscape through patchy mist or examining a scene in dappled light. Our brain can extrapolate from the detail in highly visible areas of the image and lead us to expect similar pattern in the obscured areas, possibly confirming it from limited visible detail.

The geological map, however, gives only inadequate clues, such as broken lines, to the variation in resolution. Furthermore, the results are ambiguous and the complex reasoning behind the map increases this ambiguity. The field evidence for the position of a formation boundary at the ground surface may rest on a combination of observation of outcrops, inference from topographic features, analogies with the behaviour of the formation elsewhere, extrapolation from the subsurface, and knowledge of the geological setting and history. A smoothly drawn curve might indicate that the boundary was seen to follow that line in the field, or that the feature was thought from other evidence to be gently folded, or it might show a likely position where detail was uncertain. An intricately convoluted line, on the other hand, might mean that the pattern was clearly visible in the field. Or it might mean that it was drawn parallel to a nearby tightly folded surface, or that it corresponded to a pattern of folding believed to be present throughout the area, or that it was a smooth surface intersecting complex topography. How should one depict the intersection of a smoothed surface known only from sparse boreholes and a land surface known in full detail? The shapes of lines on the map give mixed messages.

These ambiguities cannot be resolved from a single representation such as the conventional geological map. They can be resolved only through an ability to trace the reasoning process through a succession of images. For example, the significance of the intersection between a topographical and a geological surface would be easier to decipher if it were possible to look at separate models of the two surfaces. Equally, it would be helpful on occasion to visualise and follow the path from observation through reasoning to interpretation and to watch their relationships develop in an image on screen. This could become possible if the relationships were recorded during survey.

Returning to the finite element method that enables us to describe a surface by elevations, slopes, and curvatures, there is another problem with data distribution. The slope and curvature depend on the area over which it is measured. Mathematically, it refers to a point of zero area, and has no meaning on a rough surface without tangents. In geology, we replace the real feature by an artefact. We can model the surface by a triangular grid. If each triangle is a flat plane going through the three corner points, the slope of the planar facet can readily be calculated. The slope at the node required for the quintic polynomial could be taken as the average slope of these triangles over the area of interpolation. Deciding on the area of interpolation is not straightforward, for the triangles vary greatly in area. One possibility is the Thiessen polygon (Bonham-Carter, 1994), which is the area closer to the subject point than to any other data point. It is not possible, however, to avoid the implications of the data distribution.

The information on which the model is based stems from known points on the surface but is not necessarily limited to them. Background knowledge of the invariant aspects of the shape or form of the surface may be available from studies of other surfaces that developed in similar conditions or from consideration of the processes of their formation, perhaps supplemented by experimental work. The form of a surface may thus be known without knowledge of its location. One might know, for example, that a pattern of tight folding exists in an unexposed area, without any information on the location of the folds. Features of smaller wavelength than the sampling pattern cannot be accurately placed. Mallet (2002) draws a useful distinction between interpolation to determine the expected value of a surface at any point, and simulation of a specific scenario. The latter could give a more realistic view of the form of the surface and its geological characteristics at the expense of less accurate estimation of elevations. The range of possible scenarios or shape characteristics can help to determine the probability envelope within which the surface is likely to lie. And again, a single image cannot contain all this information.

Filling the gaps between observations on a geological map, say between elevations of a horizon estimated from wells on a subsurface map, might be seen as a process of adjusting a surface to fit the observations by means of spatial transformations. In the primitive computer methods considered earlier, this involved vertically scaling and adding the smooth curves of basis functions given by geographical coordinates raised to various powers. Alternatively, in manual contouring, similar procedures can be used to sketch in smooth surfaces passing through the point values, or shapes thought to be typical of the form of surfaces created by geological processes might be scaled and adjusted to pass through the points. The result of any of these procedures is a spatial model that is thought in some respects to resemble or to throw light on the real surface. The manual methods might be regarded as a form of pattern recognition, where the influences of known elements of pattern are explored by adjusting them with geometrical transformations to fit the available information. The computer equivalent is the technique of wavelet analysis (Graps, 2003). This has had considerable success in data compression for two-dimensional images, but the difficulties of finding suitable functions to represent geological shapes and the laborious computing needed to implement them have so far limited geological applications to well-defined forms such as faults. Surely this is fertile territory for future exploration, where well-established manual methods can be carried through to a computer environment, and mathematical procedures can bring greater rigour to geological insights.

Looking ahead

It is no doubt foolish to predict, but I do so in the confident belief that the procedures of science will quickly consign my follies to oblivion and in the hope that some ideas, suitably adjusted, will make their ephemeral contribution. The sightings outlined in this paper suggest to me that Geological Surveys will continue to change from their past emphasis on producing The Geological Map, through digital cartography, spatial modelling and electronic delivery, to a future emphasis on contributing to the geoscience knowledge system. Powerful and flexible spatial models dependent on IT can overcome important limitations of the cartographic model. Electronic field notebooks will increase in computing power and in the bandwidth of their wireless interconnection to the geoscience knowledge base. This will eventually make it possible to record a hypermedia stream of observational records, linked by a mark-up language to an ontological framework. This could supply templates of object classes and their expected behaviour, thus supporting description of the object instances. It could link to spatial models and point to the need for new observations to confirm or adjust the tentative interpretation.

The reasoning process could thus be documented, uncertainties and degrees of belief quantified, and expectations adjusted against the changing backdrop of a shared fragmented record of a holistic view. The abilities of IT to record, store, compute, interconnect, manipulate, communicate, transform and present information will be more effectively interfaced to the background knowledge and human insights of geoscientists. In particular, computer graphics should tie the construction of spatial models to human visual skills for more effective monitoring of interpolation, correlation and generalisation, and for integrating the complex consequences of events in many facets of geology. Concepts such as emergent and complex systems will continue to slide in and out of fashion while taking their place as important new perspectives on geoscience.

Geologists can learn much from brain scientists to guide their IT applications. Nevertheless, the human brain works in mysterious ways. Experienced geologists have learned subtle responses by mulling over numerous observations, filtered from a vast amount of sensory input, training, discussion and browsing in earlier records. They can, at least partially, understand the thought processes of their colleagues. We may not fully understand the process but it works, and uses knowledge far beyond the reach of IT systems. A geoscience knowledge base is a repository of human knowledge. But geological investigation and communication of that knowledge are technology-based, and modern IT can support human activities with more powerful tools and more flexible information containers. Conceptual models, process models, and object-oriented methods within a defined framework and ontology can offer greater rigour than conventional methods. The containers we choose for knowledge-base fragments have a feedback effect on the science. They will alter the channels by which we communicate and the lines along which we think, reason and explain. The exploration and transformation of the interface between human thought and representations of the outcome have scarcely begun. Nevertheless, computers have stimulated work in many areas that throw new light on geology, can better support the skills of the geologist, and are already transforming the task, the products, and the means of communication in geological survey.

The resulting opportunities to link surveying to the general geoscience model within the context of a comprehensive computer-assisted knowledge-based system and the need for long-term digital preservation and access to records within a repository (National Library of Australia, 2004) may require some redefinition of institutional boundaries and responsibilities, including such areas as surveying; digital storage and preservation; framework and ontology design; standards and quality assessment; and dissemination of geoscience knowledge. Geological Surveys, as a long-established international network on information communities, are well placed to play a leading role in transforming the structure of the geoscience knowledge base. But economic geology, particularly in the oil and gas industry, with greater resources, more narrowly directed commercial pressure and a willingness to outsource, is in fact leading the development (POSC, 2004). Academic research, development and training are also essential contributors to a knowledge system. Responsibilities for aspects of the knowledge system and procedures to implement them may well be adjusted and transformed as a result of new technology.

As many a failed IT firm can testify, good technology is not enough. Commercial and scientific success depends on IT supporting a service perceived by its customers as an improvement. Surely IT can facilitate better science, provided geologists understand and drive the developments to meet their needs. The scope for transforming unsatisfactory features of a Geological Survey's products is considerable, but the customer appeal of new products is notoriously unpredictable. The complexity of the transformation, with eggs dependent on the chickens that depend on the eggs, could lead not to change but to gridlock. The system must therefore evolve step by step. But to guide the steps, a forward-looking business plan is needed along with careful design of a framework. This must look ahead to a knowledge-based system that can handle existing methods and legacy information, but also remain open to predictable future developments and be capable of evolving to meet unforeseen new ideas. Only geoscientists, initially the explorers and transformers in academic research, can guide this development towards a satisfactory outcome.

There is a case for pathfinder standards for knowledge-based geoscience systems based on existing best practice. Even at this early stage, it could tentatively define the mainstream and encourage collaboration among the communities around the world that are working in these areas. The aim of such standards is to provide a consistent design framework to focus effort and reduce arbitrary decisions that cause pointless incompatibilities, while not inhibiting the diversity needed for successful evolution. In due course, they might help to provide the critical mass of system users needed for commercial development of a robust, industry compatible, IT-supported system. The result should enable a knowledge base of geological information, from survey to end use, to be shared globally, in both geographical and disciplinary senses. The actual and potential, direct and indirect, value to the users is a measure of the benefit of the work and the effectiveness of the system. The relative costs of surveying within the framework are a measure of the efficiency of the system. An analysis and forward projection of cost and value could help to guide the pace and direction of development.

The outcome will be unexpected, and will take time. As an exemplar, this paper began with Christopher Columbus, for whom containers were also all-important. Even going eastwards, he could never have reached Cathay in a horseless carriage, owing to its unreliability, a lack of roads, and to not having one. For convenient access from Iberia to Asia, Boeing had to develop the 747. For that to happen, America had to be discovered. Adjusting expectations and taking a long-term view, Columbus can retrospectively be congratulated on the perspicacity of his business plan. As for mathematical geologists, huge progress has been made in the few years since that meeting amidst the transformational events in Prague. Our splendid colleagues, to whom IAMG owes its existence and success, must surely have done something right and, as always, the really exciting transformations lie just ahead.

Conclusions

Information technology enables us to do things differently. In thinking of its significance for geological survey, we therefore focus on the underlying objectives, not on the conventional products. Object-oriented analysis provides more flexible containers for information, regarding the things of interest as objects, and the processes and concepts as models, with attached records of their mutual relationships. It distinguishes between object instances, such as the actual occurrences of features seen in the field, and the more abstract notion of object classes, such as the stratigraphic formation or fossil species to which the object instances belong. Object classes can be hierarchical, each inheriting features from those above. The approach matches aspects of geological thinking, but separates visualisation from recording, and breaks free of the constraints of the static, two-dimensional geological map, with its fixed scale and predetermined sheet boundaries and content. The geological objects are multimedia representations, and can connect through mark-up languages to computer software. They can represent the many facets of surveyed geology, the numerous sources of information, and the multiplicity of hypotheses. They can relate these to the processes of abstraction, reasoning and explanation in the multidimensional general geoscience model, and connect to procedures for simulation, selection, generalisation and visualisation. They can handle the consequences of complex systems and their fractal models, emergent properties, generators and attractors. For a consistent view of this system of IT support for human knowledge of geology, a conceptual framework and ontology are required – enabling software components to work together, and helping users to know where to put things, where to find them, and what to call them. Bayesian statistics could give a coherent philosophy for handling the uncertainties that build up during the abstraction and reasoning processes. Discrete spatial modelling enables geologists to go beyond the overloaded single image of a map, and to make greater use of their inherited and trained visual abilities in relating and interpreting spatial information. Pathfinder standards could help the diverse geoscientific communities involved to collaborate in designing the emerging system, which may lead to new divisions of responsibilities.

References

Baas, A.C.W., 2002. Chaos, fractals and self-organisation in coastal morphology: simulating dune landscapes in vegetated environments. – Geomorphology, 48: 309-328; Amsterdam.

Bardosy, G. and Fodor, J. 2003. Geological reasoning and the problem of uncertainty. – http://www.jiscmail.ac.uk/files/GEO-REASONING/bardossy-w8.pdf

BGS, 2004. The BGS Geoscience Data Index (GDI). – http://www.bgs.ac.uk/geoindex/home.html

BIPS, 2004. BIPS: Bayesian Inference for the Physical Sciences. – http://astrosun.tn.cornell.edu/staff/loredo/bayes/

Bonham-Carter, G.F., 1994. Geographic Information Systems for Geoscientists: Modelling with GIS. – Pergamon; Oxford, 398 pp.

Brodaric, B., 2000. Digital geological knowledge: from the field to the map to the Internet. – U.S. Geological Survey Open-file Report 00-325; Reston, VA. http://pubs.usgs.gov/of/of00-325/brodaric.html

Brodaric, B. and Gahegan, M. 2000. Geoscience map data models, open systems GIS and semantics. – http://www.gisworld.org/Conference/2000/1147.pdf

Chamberlin, T.C., 1897. The method of multiple working hypotheses. – Journal of Geology. Reprinted in 1995. - Journal of Geology, 103: 349-354; Chicago.

Coad, P. and Yourdon, E., 1991. Object-oriented design. – Yourdon Press; Englewood Cliffs, NJ. 197 pp.

Cross, T.A., and Lessenger, M.A., 1999. Construction and application of a stratigraphic inverse model. – [In:] Harbaugh, J.W. et al (eds.). Numerical experiments in stratigraphy: recent advances in stratigraphic and sedimentologic computer simulations. SEPM (Society for Sedimentary Geology) Special Publication No. 62; Tulsa: pp. 69-84.

CSIRO Australia, 2004. XMML: online data transfer for the exploration and mining industry. – http://xmml.arrc.csiro.au/
GISWorld, 2004. GIS Division of the Geological Association of Canada. – http://www.gisworld.org/
Gocad, 2004. The Gocad Web Site Home. – http://gocad.ensg.inpl-nancy.fr/
Graps, A., 2003. Amara's wavelet page. – http://www.amara.com/current/wavelet.html
Harbaugh, J.W., Watney, W.L., Rankey, E.C., Slingerland, R., Goldstein, R.H., and Franseen, E.K., (eds.) 1999. Numerical experiments in stratigraphy: recent advances in stratigraphic and sedimentologic computer simulations. – SEPM (Society for Sedimentary Geology) Special Publication No. 62; Tulsa. 362pp.
Harrison, J.M., 1963. Nature and significance of geological maps. – [In:] Albritton, C.C. (ed.), 1963. The fabric of geology. Freeman, Cooper & Co; Stanford, pp. 225-232.
Heylighen, F., in press. The science of self-organization and adaptivity. – [In:] The Encyclopedia of Life Support Systems. EOLSS Publishers Co. Ltd. http://pespmc1.vub.ac.be/Papers/EOLSS-Self-Organiz.pdf
Hinze, C., Sobisch, H-G., and Voss, H-H., 1999. Spatial modelling in geology and its practical use. – Mathematische Geologie, 4: 51-60; Dresden.
Howson, C. and Urbach, P, 1993 (2nd edn) Scientific reasoning: the Bayesian approach. – Open Court Publishing Company; Chicago. 470 pp.
Hume, D., 1777. An enquiry concerning the human understanding (edited by L.A. Selby-Bigge, 1902). – The Clarendon Press; Oxford. 371 pp.
Kent, W., 1978. Data and reality. – North-Holland Publishing Company; Amsterdam, 211pp.
Loudon, T.V., 2000. Geoscience after IT: a view of the present and future impact of information technology on geoscience. – Pergamon; Oxford, 142 pp. Also available as Computers & Geosciences, Special Issue, 26(3A): A1-A142 [online in subscribing libraries]; Pergamon, Oxford.
Mallet, J-L., 2002. Geomodeling. – Oxford University Press; New York, 599 pp.
National Library of Australia, 2004. PADI: Preserving access to digital information. – http://www.nla.gov.au/padi/index.html
Nicolis, G., and Prigogine, I., 1989. Exploring uncertainty. – W.H. Freeman; New York, 313pp.
Nordlund, U., 1999. Stratigraphic modelling using common-sense rules. – [In:] Harbaugh, J.W., Watney, W.L., Rankey, E.C., Slingerland, R., Goldstein, R.H., and Franseen, E.K., (eds) 1999. Numerical experiments in stratigraphy: recent advances in stratigraphic and sedimentologic computer simulations. SEPM (Society for Sedimentary Geology) Special Publication No. 62: 245-251; Tulsa.
North American Commission on Stratigraphic Nomenclature, 1983. North American Stratigraphic Code. – The American Association of Petroleum Geologists Bulletin, 67 (5): 841-875; Tulsa.
Paredes, C. and Elorza, F.J., 1999. Fractal and multifractal analysis of fractured geological media: surface-subsurface correlation. – Computers & Geosciences, 25 (9): 1081-1096; Oxford.
Pshenichny, C.A., in press. A draft for complex formal approach in geoscience. – http://www.jiscmail.ac.uk/lists/GEO-REASONING.html
POSC, 2004. POSC specifications – Epicentre 2.2. – Petrotechnical Open Software Corporation; Houston, Texas. http://www.posc.org

Simpson, G.G., 1963. Historical science. - [In:] Albritton, C.C. (editor), 1963. The fabric of geology, pp 24-48. Freeman, Cooper & Co; Stanford. 372 pp.
Sowa, J.F., 2000. Knowledge representation: logical, philosophical and computational foundations. – Brooks/Cole; Pacific Grove. 594 pp.
Sowa, J.F., 2001. Processes and causality. – http://www.bestweb.net/~sowa/ontology/causal.htm
Strang, G. and Fix,G.J., 1973. An analysis of the finite element method. – Prentice-Hall, Englewood Cliffs, NJ. 306 pp.
Thompson d'A. W., 1917. On growth and form. – Cambridge University Press, Cambridge. 1116 pp. [revised and reprinted 1942]
Turcotte, D.L., 2002. Fractals in petrology. – Lithos, 65 (3): 261-271; Amsterdam.
USGS, 2004. International Clearinghouse for Geologic Map Databases and Standards. – http://ncgmp.usgs.gov/intdb/
Van Wagoner, J.C., Hoyal, D.C.J.D., Adair, N.L., Sun, T., Beaubouef, R.T., Deffenbaugh, M., Dunn, P.A., Huh, C., and Li, D, 2003. Energy dissipation and the fundamental shape of siliciclastic sedimentary bodies. – Search and Discovery, Article #40080. – http://www.searchanddiscovery.com/
Webster, R., 2000. Is soil variation random? – Geoderma, 97 (3-4): 149-163; Amsterdam.
Winchester, S., 2001. The map that changed the world. – Penguin Books; London. 338 pp.