

Guest Editorial:
Concepts and paradigms in spatial information:
Are current geographic information systems truly generic?¹

P.A. BURROUGH, Department of Physical Geography, University of Utrecht, Post box 80.115, 3508 TC Utrecht, the Netherlands.

A. U. FRANK, Institut für Landesvermessung und Ingenieurgeodäsie, Technische Universität Wien, Gusshausstrasse 27-29/127, A - 1040 Vienna, Austria.

Abstract. This article considers the philosophical and experiential foundations of human perception of geographical phenomena and their abstraction and coding in geographical information systems. It examines the role of culture and language in describing geographical reality and explores the ways geographical data models reflect how people view the world. Differences between those who see the world as made of exact entities and smooth continuous surfaces, and those who prefer to view reality as a dynamic and complex are explored in terms of five aspects of spatial data, namely (i) objects versus fields, (ii) single scale versus multiple scales, (iii) Boolean versus multivalued logic, (iv) static versus dynamic descriptions and (v) determinism versus uncertainty. These five aspects are further divided into nine factors of geographical data which indicate the differences in the way people perceive spatial data. Eight "typical" GIS applications and four generic methods of handling spatial data are examined in terms of these nine factors to define a GIS "hyperspace". The locations of the typical applications and the generic methods in this hyperspace show why no single generic approach to spatial data handling is sufficient for all possible applications. The analysis reinforces the authors' contention that spatial data analysis tools need to be chosen and developed to match the way users perceive their domains: these tools should not impose alien thought modes on users just because they are impressively "high tech". The implications of this conclusion for choosing or developing spatial information systems, for data standardization and generalisation and for the further development of "GIS" as a discipline in its own right are presented as topics for further discussion.

¹ Burrough, P.A., & Frank, A.U. 1995. "Concepts and paradigms in spatial information: are current geographical information systems truly generic?". In *IJGIS*, **9** (2), pp: 101-116.

1. Introduction

Geographical information systems (GIS) are no longer an emerging technology. They are used by a wide range of academic and technical disciplines, not because of the need for multi-disciplinary interaction, but because they provide supposedly universal tools for handling spatial data. The ideal concept of a tool, or a generic tool box, is that of an "enabler". Like a catalyst it permits changes to occur, providing the material it acts upon is in a certain form. During many discussions with national organisations and other users of GIS held during recent years as part of the COSIT, EUROGI, EGIS, NCGIA² and other initiatives, we have noted that there is a widespread acceptance of two ideas. The first is that GIS will be much more useful when all data are readily available in a standard format. The second is that the design and operation of the tool kits we call "GIS" are *discipline-independent*. Taken together, these ideas imply that the possession of data and skills with this technology automatically brings mastery in many fields. Examples of "technomastery" encountered have been computer scientists who used GIS to pronounce on epidemiology, managers of utility systems who claimed expertise in resource management and job advertisements that require "GIS specialists" who can operate a specific commercial system. We believe that these views are naïve. They reflect the limited and simplistic theoretical base of most spatial analysis carried out with current GIS and the general lack of an underlying theoretical basis for understanding spatial and temporal data within the context of a given discipline.

Conventional GIS are good at treating spatial data as exact objects (points, lines, areas) or discretised fields (pixels, voxels) but they have severe problems in dealing with certain forms of complexity, scale differences, generalisation and accuracy. We believe that people fail to appreciate that the inbuilt structures and assumptions in GIS technology may affect perception and the recording, storage, analysis and interpretation of data, thereby leading to distortions and misunderstanding. "*The logic of the information system cannot tolerate any vague or undefined notions*" (Circic 1975). This happens because the basic concepts of space and time, and the methods for handling both, used in current GIS are not necessarily congruent with the needs of the application area. The problem is particularly acute in areas where multiscale and multivariate data have to be handled.

It is our contention that the failure to explore alternative ways of describing the earth's surface and to develop appropriate ways of dealing with a wide range of different spatial data models and methods for handling spatial data lies behind much of the confusion and disappointment experienced by many new users of geographical information systems today. It should be a major aim of GIS research to lay the intellectual foundations for a general framework of spatial data modelling and analysis that extends far beyond the limited capabilities of current commercial systems.

This editorial aims to open a discussion on how users from different application areas perceive the phenomena they deal with, and how their spatial information should be handled. We address first the question of discipline-related perception, then consider the problems of spatial data models and GIS design that follow in terms of the following questions:

1. Is there only one possible set of geographical representations of space and time for modelling spatial phenomena? Do all users think like geographers, cartographers or computer scientists?
2. Is two-valued (Boolean) logic necessary and sufficient for understanding spatial and temporal patterns and processes?

²COSIT - European Conference on Spatial Information Theory; EUROGI - European Umbrella Organisation for Geographic Information; EGIS - European Geographical Information Systems Conferences; NCGIA - (US) National Center for Geographic Information and Analysis.

3. What is the role of culture and experientialism in determining which kinds of spatial data models and logic are most appropriate in any given situation?

This leads to the development of a "GIS-hyperspace" for displaying the relations between different classes of spatial data handling, conventional generic methods of spatial data handling and major groups of users and finally to the question whether "GIS" (the activity) should aspire to become a discipline in its own right.

An earlier version of this text served as a lead paper for the recent ESF GISDATA Specialist Meeting on Spatial Conceptual Models for Geographic Objects with Undetermined Boundaries, held at Baden, Austria, 8-12 June 1994, which will be reported separately (Burrough and Frank - in preparation).

2. Perception

How do you see the world? When you look out across the landscape do you automatically see an assemblage of unique inter-locking objects or entities, that you might call land use parcels, soil mapping units, river terraces or ecotopes? Or do you see the world as a smooth, but complex continuum, your mind's eye transforming the patterns and processes of nature into reams of partial differential equations? Are you a westerner, steeped in atomistic ideas derived from the Ancient Greeks? Or are you from the east, where concepts and ideas seem naturally to blend in fuzzy, diaphanous ways, as does the mist on a Chinese painting? Do you recognize objects or process patterns?³ Or, are you a fractal freak, seeking self-similar worlds within worlds within worlds, nested images like Russian Dolls, or Jonathan Swift's fleas? How do you describe patchy phenomena such as baldness or erosion? Are you a determinist, seeking exact reasons for every action, even though the sum total of many simple deterministic actions may be more chaotic than predictable, and linear models may fall far short of describing reality? Are you a stochastician, content to describe your world in terms of probability distributions and probable interactions, and if so, do you look more to Bayes or to Fisher for inspiration?

In reality we all see the world in a mixture of ways. Based on experience and training, we build models, recognize patterns and analyze structures. Because we are what we are, humans are particularly skilled at distilling structure from complex images - this has been essential to our development and survival. Understanding the meaning of patterns can be crucial - the tiger in the grass is not just a pretty pattern of black and brown, but the face of Karl Marx perceived in the shadows in the limestone cave, though it may inspire some, is more whimsical than threatening. "Give us a sign O Lord", cry the would-be faithful in *The Life of Brian* (Monty Python 1979) - life is a process of recognizing signs, of creating and interpreting symbols and models and, depending on context, of taking or of not taking action.

The symbols we use and the models we build are not only influenced by our own past experience, but also by the accumulated experience of our culture, of mankind in general. Experiential realism as advanced by Lakoff (1987), Johnson (1987) and Lakoff and Johnson (1980) provides a philosophical basis to include cognitive aspects into science. In this theory, the meaning of words is related to experience; basic terms are related to direct bodily experience, abstract thought is connected to direct experience through metaphorical transformations [metaphors we live by]. As most bodily experience is spatial, spatial terms are fundamental to cognition and are used metaphorically to express many non-spatial, abstract relations.

³ "Pattern processing may be something we can do well. Kanji [the Chinese characters used to write Japanese] is much more a pattern than an alphabet, and the structure of our language isn't logically oriented - its a fuzzy language. So the way we communicate is fuzzy compared to the Western style" (Junichi Shamada, 1993).

Experiential models of the world are based on the interaction of humans with our environment. Many fundamental aspects of our interaction with the world are physiologically determined (i.e. the eye) and therefore, with exceptions caused by anomalies such as colour blindness, or tunnel vision, are common to all human beings. Experiential models of space can reveal themselves through spatial reference in natural language, through experiments with human subjects, through the observation of spatial behaviour, or through study of the artifacts of such behaviour (Mark and Frank, 1994). Experiential realism has been discussed in a geographic context by Couclelis (1988), Mark and Frank (1989) and Mark (1989).

In Science, pattern recognition and model building are formalised activities for discovering and describing patterns or processes that repeat in a predictable way.

Science ... is ..the search for algorithmic compressions. We list sequences of observed data. We try to formulate algorithms that compactly represent the information content of these sequences. Then we test the correctness of our hypothetical abbreviations by using them to predict the next terms in the string. These predictions can then be compared with the future direction of the data sequence. Without the development of algorithmic compressions of data all science would be replaced by mindless stamp collecting - the indiscriminant accumulation of every available fact. Science is predicated upon the belief that the Universe is algorithmically compressible... a belief that there is an abbreviated representation of the logic behind the Universe's properties that can be written down in finite form by human beings.

J.D. Barrow. (Theories of Everything, Vintage, 1991, p.11)

Models and symbols permit compressed descriptions of the world, and science is the search for such algorithmic compressions that at best should explain why the universe is as it is, and at least should be useful for predicting important aspects of our daily life.

Like many other kinds of signs and models used by humans, scientific concepts, models and methods may also be culturally influenced. Science, particularly Western Science, has developed powerful mathematical tools for algorithmic compression and mathematics has become the generally accepted universal language of science. It is indeed remarkable that mathematics forms such a useful and pervading basis for the formal ways in which "truths" are discovered, described, formulated and propagated (Barrow 1992).

As with all human theory and language, mathematics depends on people's perception of the world, their experiences and generalization of these. Geometry - as one of the earliest examples of abstract thinking - is the product of abstraction from measurements of observable figures on sand with imprecise physical instruments like a ruler and compass to the abstract ideas of a ideal point and a perfect circle. Mathematics in its pure form is a formal theory, a play with symbols without any intrinsic meaning - if meaning is construed as the relation of a symbol to particular objects in the world.

Experimental science has triumphantly demonstrated that for logic and mathematics, and the models built with them, to be useful, the symbols and objects they deal with must be capable of being related to observable phenomena.

*The isomorphism between these properties of the numeral series and certain empirical operations which we perform with objects permits the use of the series as a **model** to represent aspects of the empirical world.'* (Stevens, 1946).

Clearly, useful results can only be obtained if the level of abstraction in the models and symbolism is appropriate for the task in hand. It is also essential to be able to link both the model and the observations through conventional languages as well as mathematics. However, there is no golden guarantee that just "*because prolonged and accurate mathematical calculations have been made, the application of the result to some fact of nature is absolutely certain*" (A.N. Whitehead 1953).

The history of scientific thought shows that Western philosophers, mathematicians, lawyers and computer scientists have placed enormous emphasis on abstract ideas in their formulation of models of the world. For example, the binary division between "Truth" and "Falsehood". The logical laws of *identity*, *non-contradiction*, and *the excluded middle* have long been thought sufficient by the Western mind to usefully describe the workings of the world, even though this logic sometimes led to intuitively absurd situations, as is the case with many self-referencing statements or with complex problems. Not all societies have been troubled by strict logical strait jackets. Barrow (1992) remarks that the Jains in ancient India recognised seven grades of truth or falsehood. Western logic, like geometry, developed from approximation to abstraction, and became divorced from reality.

During the last two hundred years mathematicians have progressed from a simple atomistic and logical view of the world to include smooth continuous fields as well as discrete objects. They have also learned to cope with complex variation in time and space (e.g. Stewart 1989, Mandelbrot 1982) and are beginning to explore the limitations of reductionist thought (Cohen and Stewart 1994). Mathematical methods for dealing with uncertainty are firmly established in the form of mathematical statistics (e.g. Porter 1986) and in the ideas of fuzzy logic (Zadeh 1965).

Though mathematics has progressed, the new ideas have not diffused completely to all disciplines that deal with the spatial analysis of the earth's surface, nor in the design and implementation of geographical information systems. There is a large gap between the richness of the ways in which people can perceive and model spatial and temporal phenomena and the conceptual foundations of most commercial geographic information systems. The information systems have been designed to meet a limited paradigm in which real world phenomena can be approximated by exact geographic data types which are termed "objects" or by smooth, continuous "fields" (see section 3.1). This reductionist approach is sensible from a cartographic viewpoint when we are dealing with anthropogenic entities such as land parcels, houses, cables or administrative units that can be modelled by simple, easy-to-combine abstract objects but it may be totally inappropriate for phenomena that a) vary over many scales simultaneously, b) are difficult to define exactly, c) are dynamic, and d) are probabilistic. This includes most of the natural world.

Professional training in science and technology includes extensive exposure to the concepts, models and classification systems that are generally perceived to be necessary and appropriate for working in any given area. If it is generally perceived that a reductionist, atomistic view of the world is a necessary framework for operating in, people modify their true perceptions of reality to meet accepted paradigms. For example, many efforts in pattern recognition used in the analysis of remotely sensed data (which are treated as continuous fields) implicitly assume that "exact objects" such as soil units, ecotopes, land use classes and so on, can be derived by filtering and classification. Certain patterns of attribute values in a continuous field may be perceived as "objects": for example, vortexes in atmospheric circulation patterns are sometimes explicitly named hurricanes. The survey and classification of geology, soil or vegetation into exact, sharply defined classes or areas (e.g. Soil Survey Staff 1976) that meet abstract ideas of nicely defined, properly circumscribed units has been an enormous exercise in forcing continuously varying phenomena into "exact" moulds. Although much scientific evidence has been assembled to demonstrate that soils are not exact, homogeneous entities, (Burrough 1993) the abstract entity paradigm is still used because the technologies used for handling spatial data, and the organisation of the definition and

adoption of "standards" actively require it. In areas where that paradigm has been discarded or is not used, such as geological exploration or hydrological modelling, other abstractions are used. It is a real problem to combine data that have been recorded according to different paradigms, not because of a lack of technological standards, but because the perceptions governing the different observations have been constrained in different ways.

To summarise, we distinguish several phases in the way humans organise spatial data. First there is the act of *perception*, which is closely followed by *representation*, which we define here as the choice and use of an appropriate conceptual model by the human brain. Once people have agreed on a system of representation, and have embodied it in *language*, it is necessary to *code* the representation in a such a way that computers (or other non-human information processing technology) can handle the data reproducibly. The creation of the information system is then a matter of determining an efficient *implementation of the codes*. In geographical information systems, implementation includes the technology used to portray spatial phenomena, such as is understood by the terms "raster" and "vector" displays, and the various means of implementing them in topological networks, runlength codes or quadtrees (Burrough 1992). In principle, perception and representation should always precede coding and implementation, but the establishment of accepted thought patterns and their technological implementation may impose strong constraints on the observer - a case of the tail wagging the dog.

3. Spatial data paradigms and GIS design

To better understand how to handle spatial phenomena, we must separate the areas in which current paradigms and methods are adequate from those in which they are either impotent or impose such strict constraints that the information handling seriously distorts perception (irrespective of the technical elegance of the graphic products!). We need to go beyond the already classic vector-raster dichotomy and consider the following five aspects:

- 3.1 objects versus fields
- 3.2 single scale versus multiscale phenomena
- 3.3 single criteria versus multicriteria
- 3.4 static versus dynamic systems
- 3.5 determinism versus uncertainty

3.1) objects versus fields

As noted, these embody the main features of conventional western thought by assuming the world is composed of two substantially different kinds of static representations of space which can be handled with conventional mathematics and Boolean logic. The first representation stresses individual objects with well determined boundaries and attributes, which is termed "*the entity view*". The second representation deals with continuous variation by modelling it in terms of a *continuous field* of data values (Coullelis, 1992).

In the entity view the world consists of exactly definable and delimitable individual objects (Nunes, 1991). Each object occupies space and has properties. Objects may consist of parts and sub-parts. Basic geographical objects are the point, the line (a set of linked points) and the polygon (a set of linked lines, or contiguous points). Their properties are described by tuples of attributes measured on nominal, ordinal, interval or ratio scales. Most analyses involve logical operations or compute new attributes from data stored in the attribute tuples. Logical operations follow the laws of Boolean logic which is reinforced by embodying it in standardized data retrieval languages such as SQL. The geometry of the entities does not usually change, except as a result of polygon overlay (intersection of spatial sets).

In the continuous field view (Goodchild 1992) every point in space can be characterized in terms of a set of attributes - terrain use, elevation, average annual rainfall, soil texture, etc. measured at a set of geometric coordinates in a Euclidean space. Each attribute is measured on an appropriate nominal, ordinal, interval or ratio scale. It is assumed that most properties vary smoothly and gradually, so that attribute values at unsampled locations can be determined by interpolation. There is no mention of an object or entity in the cognitive sense, though these terms may be used in the sense of object-oriented software engineering to describe the characteristics of particular continuous fields.

In practice, continuous fields may be represented by mathematical functions, sets of discrete isolines, or approximated by a regular discretisation or tessellation for efficient storage and retrieval. Triangular irregular networks are sets of triangular polygons that approximate a continuous surface by an irregular tessellation of linked entities.

These two models are used by all current GIS, and, as noted in the previous section, the thinking they embody has pushed users to great lengths to force their view of the world in to one or both of them. Besides enforcing an particular abstract form of thinking on users, these models have several shortcomings with respect to modelling many real situations. First, in their simplest form they are not dynamic, but deal only with a static, invariable world: they do not describe *spatial* change. Second, the models do not allow for error or imprecision, neither in the values of the attributes that are recorded nor in the description nor identification of the entities involved (though there is no reason why extra tuples containing attribute standard deviations could not be added to the database, if known). Third, these simple models do not deal with complex objects which may a) consist of interacting parts, or b) display variation at many different levels of resolution.

3.2 Single scale versus multiscale phenomena

Complex entities and continuous fields with variation at a range of scales are commonplace in nature. In the entity model, complex objects can be defined in terms of sub-objects, but the relationships between objects and sub-objects are fully defined in special terms (belongs to, is part of, owns, inherits...) This is the approach of *object orientation* (Egenhofer & Frank 1988, Jacobson et al 1992), in which it is essential to know how complex objects can be the sum of their parts.

Continuous fields can be the result of spatial variation operating over several scales. Methods of *Fourier analysis* assume that complex continuous variation is composed of sets of continuous sine and cosine surfaces. Methods of wavelet analysis avoid the limitations of Fourier series when complex variation is continuous but varies in character between different parts of the domain (Daubechies 1990).

Geostatistical interpolation methods allow pattern analysis in continuous fields over several scales and directions (Journel and Huijbrechts 1978). Most simple methods of interpolation in GIS treat the point data as though only one scale of variation is present and they use one set of parameters defining search radii, interpolation model coefficients and directional factors. The methods of geostatistics analyse the spatial covariance structure of the data to determine a) whether data are sufficient to allow interpolation at all, b) the optimal weights for interpolation, and c) the possible presence of anisotropy, trends and multiple scales of variation. As with fourier analysis, the methods assume second order stationarity of at least the first differences in attribute values over a distance vector \mathbf{h} , but given sufficient data, the spatial analyses and interpolations can be tailored to different domains (Burrough 1993).

These extensions are beginning to be incorporated in some geographical information systems, either as new developments (van Oosterom 1993), add-on packages or separate specialist software modules linked by data exchange routines. Neither object-orientation nor geostatistical methods of interpolation deal directly with dynamic patterns, in the latter case because it is assumed that the spatial covariance structures are permanent features of the continuous field.

Geostatistical methods, however, are well provided for dealing with uncertainty because they can divide spatial variation into that which is spatially coherent at the level of observation, and that which is not. The latter appears as residual noise or error, and is known as *nugget* variation. Levels of nugget variation are not necessarily constant over a domain and depend on the complexity of the underlying (unknown) pattern, the density of observations and the interactions between these two. Besides predicting the values of quantitative attributes at unsampled sites, the latest geostatistical methods also allow the prediction of probabilities that a given value occurs, or will be exceeded, and can deal also with qualitative data (Bierkens and Burrough 1993).

3.3 Single valued versus multivalued criteria

Since the invention of formal methods for dealing with continuous logic by Zadeh (1965) there has been no need to force phenomena into the straightjacket of exact objects because the degree of matching an ideal concepts can be specified. Methods of fuzzy logic are straightforward extensions of binary logic in which statements are no longer TRUE or FALSE, but take a truth value somewhere between 0 - 1, depending on context. Experiments indicate that multiple criteria selection and classification using fuzzy logic are much more robust and less error prone than standard methods using Boolean logic (Heuvelink and Burrough 1993). Using only Boolean logic for data selection by sieve mapping may reject data unnecessarily and give results that are intuitively unsatisfactory or not sensible (Burrough 1989): multicriteria methods have demonstrated advantages (Carver 1991).

Fuzzy and multicriteria methods can be implemented on most geographical information systems that permit mathematical modelling. New technology is unnecessary.

3.4 Static versus dynamic systems.

The dynamic view admits the possibility of *process-driven change* in spatial patterns. In the exact object approach, dynamic processes are often linked to a topological network of lines for route analyses. Topological networks are good approximations of transport routes and are used in conventional GIS for route optimizing and analysis. The topological networks are static, but movement is modeled by the change of attributes along the connected parts of the network: the dynamic change of attributes and not geometry. For example, in this model the form of a lake is not dependent on the level of the water.

In the continuous field approach, dynamic fields are obtained by differentiating continuous surfaces that represent some form of potential. Vortex fields in fluids are sets of directed vectors (in the original physical sense of properties having magnitude and direction) derived from differences in energy potential. Drainage nets on digital elevation models are directed vectors indicating potential hydrological flow paths. By analogy, similar kinds of directed vectors can be derived for other kinds of potential surfaces such as cost, profit, attractiveness or population density. Many major research areas outside geography -hydrology, atmospheric science, fluid dynamics, population dynamics, etc. relate dynamic change to discretised, continuous fields using sets of partial differential equations.

Simple forms of dynamic analysis have been available in some raster GIS in the form of "Map Algebra" (Tomlin 1983), but it is only recently that hydrologists have been making serious contributions to solving the problems posed by requiring generic dynamic modelling functionality in a GIS (Moore et al 1993, Nyerges 1993, van Deursen & Kwadijk 1993). Until recently, most dynamic modelling required software external to the GIS, which was used only for data supply and display, but generic approaches to dynamic modelling within GIS are being developed (van Deursen & Kwadijk 1993).

Dynamic changes in continuous fields can be modelled in raster GIS by computing new attribute values for each cell for times t_1, t_2, t_3, \dots and displaying the results as a time sequence.

If enough images can be computed and stored, true motion models can be obtained, though they are usually not interactive. Most implementations of dynamic analysis still ignore uncertainty and errors as important parts of the information.

New technological developments are beginning to include both dynamic and complex analysis in GIS toolkits. In the exact entity view, object orientation permits geometry and location to be treated both separately and together, allowing objects to move through scenes and over and around other objects. These developments have been particularly successful in providing dynamic CAD-CAM and "virtual reality" systems, which have been used in architecture and landscape design. In these systems behaviour is almost always completely deterministic and uncertainty is neglected.

3.5 *Determinism versus uncertainty*

The choice of exact objects or smooth continuous fields as models or reality, and the use of Boolean logic admits no room for error or uncertainty. Many models of physical processes, likewise, assume implicitly that uncertainties are small, or can be quantified or reduced by better measurements, so that process models operate deterministically.

Studies of errors in conventional spatial data handling have concentrated on obvious aspects of geometry and attribute measurements (RMS errors) and correct topological linkages. Most vector and raster-based GIS can handle errors on attributes by including metadata and data on standard errors, but many users prefer to avoid doubling their databases to include this information if it is not strictly necessary. When deriving new data from attributes, the quality of the derived data can be obtained by statistical methods or error propagation analysis if the errors associated with the model inputs and coefficients are known (Heuvelink 1993), though these methods are not currently available in standard GIS.

Practical research is beginning to demonstrate that in application areas such as resource appraisal and environmental modelling, errors are not simple, and may be both spatial and related to different scales. As indicated above, geostatistical methods and methods for dealing with vague phenomena (fuzzy logic) are proving useful. With many process models it is important to pay attention to model calibration, validation and the propagation of errors (see de Roo 1993), and few GIS have any pretensions in this area (but see Goodchild and Gopal 1989, Heuvelink 1993).

4. Different classes of users of spatial information.

Objects versus fields, multiscale phenomena versus simple phenomena, boolean logic versus multi-valued logics to describe complexity or vagueness, dynamic versus static view of the world, error and uncertainty: these five main issues confront all users of spatial information and they should know how these issues relate to their own world view. For the sake of discussion, we have grouped users as follows:

4.1 Users in areas such as cadastral mapping, utility management and automated mapping usually deal with exact entities and continuous fields. Many commercial GIS match the requirements of these practitioners because the market for these kinds of systems is large, the application areas are reasonably well understood and are supported by legal frameworks. We call this group *managers of defined objects*.

4.2 *Planners and resource managers* are increasingly using GIS for spatial planning. This group has infused its tradition and methods of conceptualization of the spatial world into the GIS. Some commercially available systems cannot disguise their roots in a planning tradition which used transparent maps overlaid one on the other to find areas with certain properties. The U.S. vendors of these kinds of GIS sell not only GIS tools, but also a planning philosophy, which may be

suitable for planning new suburban areas in flat landscapes but not particularly attractive for planning tasks in old European cities that are strongly influenced by topography and that may be subject to other norms (Campari and Frank 1993).

4.3 Environmental scientists charged with inventory, modelling and monitoring physical processes in the earth's atmosphere, biosphere, hydrosphere, pedosphere and lithosphere use spatial information in many ways (Goodchild et al 1993). Environmental scientists deal with complex dynamic systems that may not be easily broken down into simple components. They may neither be able to define, nor to record the attributes of the phenomena they work with in their entirety and must resort to sampling. These users are more likely to want to model processes of distribution and change than they are to answer simple data retrieval queries. Instead of creating databases that are complete models of the world they have to make do with samples, with approximations. They deal with interpolation and extrapolation: they must trade certainty for probability. We call this group *space-time modellers*.

4.4 Decisions regarding space are made by *politicians and entrepreneurs*. They decide where a new shopping mall is to be built, where oil wells are drilled and how forest stands are exploited. They use advice from the groups listed above and they make use of output from GIS and other spatial analysis tools when it has been duly condensed and prepared such that the information can be understood and combined with other information they have to consider (mostly legal and economical) and included in formalized and legalized decision processes.

4.5 Last, there is the *public at large*, which lives in the spatial world. Humans navigate on foot and by car in this space in their quest for survival (i.e. to find the best 'conditorei' in Vienna). Enormous amounts of spatial information is constantly being processed, mostly in intuitive and informal ways before being stored. Decisions are often made in "real time".

The last two groups, despite being more influential or more numerous than the others, rarely use GIS directly, so we will not consider them further. However, we argue that (1) acceptability of GIS to any group depends on the appropriateness of the modelling tools, and (2) the modelling tools required differ for the five groups.

5. Conceptual hyperspaces in geographical information: who lives where?

Let us now consider how the spatial paradigms of the groups differ. At first sight differences are small: the GI tools used by the managers of defined objects appear to serve the needs of the planners and resource managers and the space-time modellers who both make maps using points, lines, areas and pixels. A contour map of the surface of an aircraft or car body looks similar to a contour map of a hillside which looks similar to a contour map of the variation of a chemical pollutant in the subsoil. A 3-dimensional model of geological layers in the underground resembles the drawing of a machine part. But independent tests, however, often demonstrate that the predictions made in the soil or on the hillside are far less accurate than they should be according to the tool. Why then, with so many similarities, do tools designed to work with one kind of paradigm serve one group well and not another?

The answer is though *graphic representation* may be similar, the phenomena are intrinsically different. The imposition of a common methodology over context is procrustean. Managers of defined objects protest and tell the planners and resource managers and the space-time modellers to define their problems more accurately, and to break the problems down into small, discrete units that can be handled mathematically or logically. When there is sufficient computing power and data the problems will be solved, they say..

However, space-time modellers question the basic assumptions of the managers of discrete objects, particularly when they have to deal with complex, inexact phenomena. There are important philosophical differences. Managers of discrete objects do not tolerate uncertainty as being inevitable - it is just a question of defining the system properly and fully, collecting sufficient data and measuring accurately, plus being able to master the necessary mathematics to relate all objects and their attributes logically. Space-time modellers are aware that errors are inevitable, though they can be estimated and reduced, but increasing the resolution of observation only reveals more detail. Therefore they do not accept a fully deterministic view of the world and must use concepts of probability and possibility. Indeed, they are supported by theoretical studies that demonstrate that many important mathematically posed problems in science, economics and the environment possess such complexities that they are computationally intractable. Only by accepting approximations can answers to some, but not all, problems be obtained (Traub and Wozniakowski 1994).

Space-time modellers also realise that complete knowledge is impossible (if only for the trivial reason that for some phenomena complete sampling and measuring would completely destroy what is being measured). They accept prediction and interpolation as rational tools, but they know that variations may occur at many different scales. Moreover, they know that it is not possible to measure all phenomena on interval or ratio scales but that often one must make do with nominal or ordinal data.

The differences between the groups can be illustrated by constructing a kind of "GIS hyperspace" to locate the different groups and their various perceptual models. The main independent axes of this hyperspace might illuminate the differences between them and, perhaps remove some of the confusion that surrounds the popular use of the term "GIS".

Based on the discussion of the five aspects of spatial information (section 3) we have chosen nine not necessarily independent factors that seem to be of fundamental importance. We have used multivariate analysis to see which three dimensions seem to most completely explain the whole range of GI activities that we have discussed. The factors are:

1. Are spatial phenomena modelled as objects or fields? (1/3)
- 2.1 Are the phenomena essentially *single scale* or *multiple scale*? (1/4)
 - 2.2 Are the representations of the phenomena in the database *generalised* or *detailed*? (1/3)
 - 2.3 Are the phenomena capable of *exact definition* or are they somewhat *vague*? (1/3)
3. Do the data retrieval and logical analysis require *single valued logic (Boolean)* or *multivalued logic (fuzzy)*? (1/2)
4. Are the spatial phenomena *static* or *dynamic*? (1/4)
 - 5.1 Are the phenomena measured on *nominal, ordinal, interval or ratio scales*? (1/4)
 - 5.2 Can the phenomena be measured by *complete enumeration* or must they be *sampled*? (1/2)
 - 5.3 Is the approach to the spatial analysis *deterministic* or *stochastic (probabilistic)*? (1/3)

Numbers in brackets give the minimum and maximum scores used to indicate the subjectively assessed location of a group of users or an application with respect to each question. For example for question 1 the scoring is: only objects - 1, both - 2, only fields 3.

Note that in these nine questions we understand that detailed entities attempt a complete description of a phenomenon; generalized entities only attempt a broad description of the most salient features. Also, a simple real-world phenomenon is one that can be described and visualized in its entirety. This includes all man-made objects, including land parcels, buildings and utility networks. Simple phenomena are made up of easy-to-define basic units and have a single spatial scale or a clear hierarchy of spatial scales. The relations between the various components are well understood and can be defined. In contrast, complex phenomena such as soil patterns, sub-surface ground water flow, the behaviour of large groups of people, animals or plants, do not have a single scale or well-defined set of scales; their behaviour can only be approximated by idealized basic units. Further, we note that in object-oriented systems, the term "complex" is used for objects that are composed of other objects. Object-oriented complexity is much simpler than the much more demanding, situation in naturally varying systems, but again there is a problem of words.

The three main user classes have been divided into 8 "representative" groups of application areas that cover a wide range of approaches to handling spatial data. Some concern static, exact approaches, others dynamic, deterministic (routing), others sampled and probabilistic). Four standard GI technologies for handling spatial data have been added. All have been given scores for the nine questions (Table 1). The results were analysed by correspondence analysis and the resulting scores were plotted in three dimensions.

Table 1. Scores used for correspondence analysis of GIS factors.

Question	1	2.1	2.2	2.3	3	4	5.1	5.2	5.3
<hr/>									
Application:									
Utility	1	2	3	1	1	1	2	1	1
Cadastre		1	1	3	1	1	1	1	1
Network routing	1	1	3	1	1	3	2	1	1
Land evaluation	2	3	2	3	2	1	2	1	1
Regional Planning	2	4	1	3	2	1	2	1	1
Hydrology	3	4	2	2	1	4	4	2	2
Crop modelling	2	3	2	2	1	3	4	1	1
Pollution mapping	2	3	3	2	2	2	4	2	2
<hr/>									
GIS approach:									
Vector GIS		1	1	3	1	1	1	2	1
Object-oriented	1	2	3	1	1	2	3	1	1
Raster GIS		3	3	1	2	2	2	3	1
Geostatistics		3	3	3	2	2	2	4	2
<hr/>									

The use of correspondence analysis to extract the main features of the data is in a sense pseudo science, because any analysis only reflects the subjective data entered. Nevertheless, the results help summarize a complex situation.

Figure 1 presents a generalized summary of the results of the first three axes, which between them expressed ninety percent of all the variation in the data set. It is perhaps symbolic that in this analysis the east-west axis (1) is dominant (52% variation): it expresses the tension between the complexity of the real world phenomenon under examination (tending to the east) and the increasing degree of exactness with which such a real world phenomenon can be described (tending to the west). Obviously, the more complex the real world phenomenon, the less easy it is to describe it exactly and vice-versa.

<Figure 1 near here>

The north-south axis (2) - 27 % variation - describes the relations between static (north) and dynamic (south) spatial data handling. Binary (Boolean) logic also contributes to the northerly direction, with multivalued logic contributing to the south.

The vertical axis (3) - 11% - represents the tensions between deterministic, exact (above the plane) and stochastic (probabilistic) descriptions of reality (below the plane).

Armed with the description of GIS-space in Figure 1, we can now see what goes where. Utilities, cadastral mapping, occupy the north-west block above the horizontal plane, in the same general space where vector GIS and Object oriented GIS occur. Deterministic Network route modelling occurs nearby in the western part of the plane. Regional planning and land evaluation occur to the far north east because they deal with complex phenomena that are incapable of exact description, but which are generalized and analyzed with descriptive and simple logical models. They are located in a different part of the hyperspace from "standard" spatial information systems, which perhaps explains why planners continue to express their disappointment with GIS (Masser and Onsrud 1993). As is well known, exact, quantitative mathematical planning was rejected as impractical years ago.

Applications using mathematical modelling of the physical environment are located in the south-east of the plane: deterministic approaches hover above the plane, while stochastic, sampled methods lie below it. Fully geostatistical approaches for mapping (interpolation) lie below the plane in the centre - they do not seem to have great affinity with anything else. Raster GIS also occur in the east of the plane - they can deal with generalized complex phenomena, but do not attempt a too exact description of the real world phenomena. Their structure allows the use of logical, empirical and physical models, with an accent on the latter two, and they are able to handle stochastic surfaces and error propagation.

6. Discussion.

The above analysis reinforces our contention that it is essential to consider the relevant spatial paradigms and criteria before developing or adopting technological support systems for spatial information. Importing generic methods without consideration is easy for system developers, including researchers, and for vendors, but is costly for users, for three main reasons. First, they must invest a great deal of time and energy trying to understand how to reformulate their paradigms in terms of the views of the imposed or imported system. This can have considerable knock-on effects in terms of reduced work flow and productivity. Second, the imposition of inappropriate paradigms distorts perception and reduces understanding. Third, adopting technology for the wrong reasons usually leads to rejection and disappointment, not only of the imported technology, but of all similar ilk.

The analysis also suggests to us that problems not only occur with the choice of methods of spatial data handling, but must also be rife in issues of data collection, standardization and generalisation. How can one sensibly integrate different kinds of spatial data if each has been observed, recorded, modelled and stored according to its own particular set of paradigms? How much information is lost by insisting on universal methods of data collection, description and modelling that do not match the relevant application paradigm?

The main conclusion must be that methods of handling spatial information must be linked to the paradigms of the users' disciplines and that inter-disciplinary research to determine more accommodating paradigms than the "object-field" models is essential. In spite of this, there must be some common unity, however, otherwise "GIS" would not have become such an all-enveloping noun in its own right. But has "GIS" the capacity to aspire to the status of a separate discipline, like statistics or artificial intelligence? Should "GIS-training" (at *any* level) be given without reference to an application discipline? Should "GIS experts" be recognised for their skills in technology, or

for their knowledge of how to conduct spatial analysis of data in a particular field? We believe that these issues arise naturally from the material presented in this editorial. They are central to the development of the activities supported by this journal and its readers and we recommend you think about them carefully.

Acknowledgements.

We wish to thank Peter Fisher and three anonymous reviewers for their constructive criticism of an earlier draft of this Editorial and Max Craglia (ESF GISDATA programme) for encouragement.

References.

- BARROW, J. D. 1991, *Theories of Everything*, Vintage, London 223pp.
- BARROW, J.D., 1992, *Pi in the Sky*, Penguin, London, 317pp.
- BIERKENS, M. and BURROUGH, P.A. 1993, Stochastic Indicator Simulation. Parts I and II. *J. Soil Science* **44**: 361-368 & 369-381.
- BURROUGH, P.A., 1986, *Principles of Geographical Information Systems for Land Resources Assessment*. Oxford University Press. 194pp.
- BURROUGH, P.A., 1989. Fuzzy mathematical methods for soil survey and land evaluation. *J. Soil Science* **40**: 477-492, 1989.
- BURROUGH, P.A., 1992. Are GIS Data Structures too simple minded? *Computers and Geosciences* **18**: 395-400.
- BURROUGH, P.A., 1993, Soil Variability: a late 20th century view. *Soils & Fertilizers* 56:(5), 531-562.
- BURROUGH, P.A. and FRANK, A.U., (in preparation). *Spatial Conceptual Models for Geographic Objects with Undetermined Boundaries*. Taylor and Francis.
- CARVER, S.J., 1991. Integrating multi-criteria evaluation with geographical information systems. *International Journal of Geographical Information Systems* 5: 321-339.
- CAMPARI, I and FRANK, A.U., 1993. Cultural differences in GIS. A basic approach. *Proceedings of EGIS'93 Conference Genova*, EGIS Foundation, Utrecht, 1-10.
- CIRIC, in: Bie, S.W. 1975. *Soil Information Systems*. Proceedings of the meeting of the ISSS Working Group on Soil Information Systems, Wageningen, the Netherlands, Sept. 1-4, 1975. PUDOC Wageningen, 87pp.
- COHEN, J. and STEWART, I., 1994. *The Collapse of Chaos: Discovering Simplicity in a Complex World*. Viking Press, New York, 495pp.
- COUCLELIS, H., 1988. The truth seekers: Geographers in search of the human world, In: Gollidge, R., Couclelis H., Gould P. (Eds), *A ground for common search*, Santa Barbara: The Santa Barbara Geographical Press: 148-155.
- COUCLELIS, H., 1992, People manipulate objects (but cultivate fields): Beyond the raster-vector debate, Frank. U., Campari I., Formentini U. (Eds), *Theories and methods of spatio-temporal reasoning in geographic space*, Lecture Notes in Computer Science, Vol.. 639, Berlin-Heidelberg: Springer Verlag: 65-77.
- DAUBECHIES, I., 1990. The Wavelet Transform, Time Frequency Localization and Signal Analysis. *IEEE Transactions on Information Theory*, 36: 961-1005.
- DEURSEN, W.P.A. van and KWADIJK, J.C.J., 1993. RHINEFLOW: an integrated GIS water balance model for the River Rhine. In: K.Kovar & H.P. Nachnebel (eds.). *Application of Geographic Information Systems in Hydrology and Water Resources Management. Proc. Hydrogis*. IHAS Publication 211, Wallingford. 507-519.
- EGENHOFER, M.J. and FRANK, A.U., 1988. Object-oriented database: database requirements for GIS. Proc. Intern. Geographic Information Systems Conference: The Research Agenda, NASA, Washington, vII, 189-211.
- FRANK, A.U., 1992, Data models and data structures in GIS, *Computers and*

- Geosciences* 18: 409-418.
- GOODCHILD, M.F., 1992. Geographical data modelling. *Computers and Geosciences* 18: 400-408.
- GOODCHILD, M.F. and GOPAL, S., 1989. *Accuracy of Spatial Databases*. Taylor & Francis, London.
- GOODCHILD, M.F., PARKS, B.O. and Steyaert, L.T. (eds) 1993. *Environmental Modeling with GIS*. Oxford University Press, New York, 488pp.
- HEUVELINK, G.B.M., 1993. *Error propagation in Quantitative Spatial Modelling: Applications in Geographical Information Systems*. NGS Publication 163, Faculty of Geographical Sciences, University of Utrecht, 160pp.
- HEUVELINK, G.B.M. and BURROUGH, P.A., 1993.. Error propagation in Cartographic Modelling using Boolean logic and continuous classification. *Int. J. Geographical Information Systems* 7: 231-246. 1993
- JACOBSON, I, CHRISTERSON, M., JONSSON, P. and ÖVERGAARD, G. 1992. *Object-Oriented Software Engineering*. Addison-Wesley.
- JOHNSON, M., 1987, *The body in the mind*. University of Chicago Press, Chicago.
- JOURNAL, A.J. and HUIJBREGTS, Ch. J. 1978, *Mining Geostatistics*. Academic Press, 400pp.
- LAKOFF, G. and JOHNSON, M., 1980, *Metaphors we live by*, Chicago: University of Chicago Press.
- LAKOFF, G., 1987. *Women, fire and dangerous things: What categories reveal about the mind*. Chicago: University of Chicago Press.
- MAGUIRE, D.J., 1991, An overview and definition of GIS, in: D.J. Maguire, M.F. Goodchild & D.W. Rhind (Eds), *Geographical Information Systems*, Longmans, Harlow, Vol 1:9-20..
- MANDELBROT, B., 1982. *The Fractal Geometry of nature*. Freeman, San Francisco.
- MARK, D., 1989, Cognitive image-schemata for geographic information: Relations to user views and GIS interfaces, *Proceedings of GIS/LIS'89*, Orlando, Florida, vol. 2: 551-560.
- MARK, D. and FRANK A.U., 1989, Concepts of space and spatial language, *Proceedings of the Ninth International Symposium on Computer Assisted Cartography*, Baltimore, Maryland: 538-556.
- MASSER, I. and ONSRUD, H.J., 1993. *.Diffusion and Use of Geographic Information Technologies*, NATO Advanced Science Institutes Series, Kluwer, Dordrecht, Vol. 70, Series D. 349pp. 1993.
- MATHERON, G., 1971. *The theory of regionalised variables & its applications*. Les Cahiers du centre de morphologie mathematique de Fontainbleu. Ecole Nationale Superieur des Mines de Paris.
- MONTY PYTHON, 1979, *The Life of Brian* - the words. Monty Python Productions.
- MOORE, I.D., TURNER, A.K., WILSON, J.P., JENSON, S.K. and BAND, L.E., 1993. GIS and Land-Surface-Subsurface Modelling. In: M.F. Goodchild, B.O. Parks & L.T. Steyart, *Environmental Modelling with GIS*. Oxford University Press, New York, 196-230.
- NUNES, J., 1991, Geographic space as a set of concrete geographical entities, Mark, D., Frank A.U., Cognitive and linguistic aspects of geographic space, Dordrecht: Kluwer: 9-33.
- NYERGES, T.L., 1993. Understanding the scope of GIS: Its relationship to Environmental Modelling. In: M.F. Goodchild, B.O. Parks & L.T. Steyart, *Environmental Modelling with GIS*. Oxford University Press, New York, 75-93.
- OOSTEROM, P.J.M. van, 1993. *Reactive Data Structures for Geographic Information Systems*. Oxford University Press, Oxford. 198pp.
- PENROSE, R., 1989. *The Emperor's New Mind*. Oxford University Press.
- PORTER, T.M., 1986. *The Rise of Statistical Thinking*. Princeton University Press, Princeton.
- ROO, A.P.J. de, 1993. Modelling surface runoff and soil erosion in catchments using Geographical Information Systems. NGS Publication 157, Faculty of Geographical Sciences, University of Utrecht, 300pp.

- SHAMADA, J., cited in *New Scientist*, 2 October 1993, p. 47
- SOIL SURVEY STAFF, 1976. *Soil Taxonomy*. US Government Printing Office, Washington, D.C.
- STEVENS, S.S., 1946. On the Theory of Scales of Measurement, *Science*, June 7, 103, no 2684, p. 677.
- STEWART, I. 1989. *Does God Play Dice?*. Blackwell, Oxford. 317pp.
- TOMLIN, C.D., 1992. *Geographic Information Systems and Cartographic Modeling*. Prentice Hall.
- TRAUB, J.F. and WOZNIAKOWSKI, H., 1994. Breaking Intractability. *Scientific American*, 270:(1) 90B-93. (January 1994).
- A.N. WHITEHEAD, 1953. *Science and the Modern World*, CUP.
- ZADEH, L.A., 1965. Fuzzy sets. *Information and Control* 8: 338-353.

. The first three axes of the GIS hyperspace. Vertical lines indicate positions above or below the plane.