LANGUAGE ISSUES FOR GIS

A U FRANK AND D M MARK

Interaction with computing systems can take place in formal or natural language. This chapter reviews the role of language in GIS, with particular emphasis on the design of user interfaces. Formal languages have their roots in the models used to conceptualize and represent geographical data, and here there are strong parallels between language, and its implications for patterns of human reasoning, and the models adopted by GIS. Natural language has subtle and mostly unrecognized effects on the structuring of space, and these become particularly important in a cross-linguistic context. Later sections of the chapter review the significance of natural language interfaces in GIS, particularly in such applications as vehicle navigation systems, and the problems presented by linguistic boundaries.

INTRODUCTION: DEFINITION OF TERMS

Users must be able to interact with GIS. For this reason, it is important to study how communication between users and GIS can be made most effective. In the past, for GIS, as for other information systems, most of the effort in this area has been centred on constructing appropriate query and command languages. The approaches proposed have been to use natural language or to construct formal languages. This chapter reviews and discusses why language issues in a broad sense, including cognitive issues, are crucial for the further development of GIS.

Natural languages comprise the 'everyday languages' that people use – English, French, German, Chinese, Spanish, and so on – with all their rules, exceptions and 'idiomatic' expressions. In principle, natural languages have three components: a vocabulary (lexicon), listing words (terms) that are used; a syntax and grammar, which describe how valid sentences can be formed from these words; and semantics, indicating what the sentences mean. None of these can be fully formalized (at least not today). The language that native speakers actually use is much richer and of more interest to GIS than the 'prescriptive' view of the language defined by dictionaries, academics and grammarians.

Formal languages, on the other hand, are artificially constructed languages, following formally defined rules. These too have vocabulary, syntax and semantics, and often are modelled on natural languages, but in these cases all of the components are fully described in a rigorous format.

The use of either natural or formal language to query a GIS may result in problems such as how to describe the information needs of a user. If natural language is used, the program must 'translate' the question into an unambiguous form which can be processed. This can be difficult, given the complex structure of natural languages. On the other hand, if users are made to express a query in a formal language format, they are forced to learn this language and to translate their information needs into this format. This has the potential to limit a user's ability to interact with systems, and also limit system access itself to specialized and trained individuals.

In the past, designers of formal query languages have not paid much direct attention to the linguistic aspects of the problem. Instead, they have selected formal structures following well-known models, probably inadvertently introducing structural elements of their own natural language. The tendency for the formal language to mimic the structure of the designer's natural language will make such systems easier to use for speakers of the same language, but perhaps more difficult for other users.

The terminology used formally has not always been based on concepts that people would use naturally; this applies to all information systems. In geographical systems, discussions on design have mostly been centred on the formulation of GIS queries, and have not considered the visualization and interaction necessary to inform the user of the result.

In this chapter a comprehensive approach is advocated. People use a finite set of concepts to organize their perception of space. These concepts should be respected when designing systems to communicate regarding spatial situations, both when users formulate queries and when responses are presented to them. On the other hand, computers require formal definition of terms in order to retrieve the necessary information. Thus the communication process can be seen as a translation between human spatial concepts and the formal spatial concepts in a computer program.

The discussion in this chapter is on a conceptual, logical level, and does not attempt to explain how these methods should be implemented. It seems very important that the GIS literature separates the concepts involved in a program from the mechanics of its implementation as a program. Such a separation has been advocated in a database design standard (Tsichritzis and Klug 1975), separating the conceptual database schema from the physical storage arrangement (internal schema). The standard also defines a third, external schema, namely 'user views', which describes subsets of the conceptual view, as appropriate for a specific task and which may be different from the 'corporate' view (Fig. 11.1).

The confusion in GIS literature between concepts and implementation is a severe impediment to progress in the field. Discussion is needed to determine what a system does independently from how it is achieved, especially in light of the extremely rapid development of the computer tools, where the 'how' might change overnight. The discussion of the 'what' is intrinsically linked to cognitive and linguistic issues,

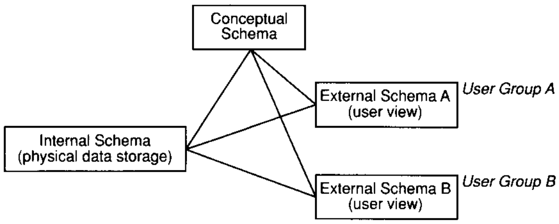

Fig. 11.1. The three-level schema.

and this chapter discusses the 'conceptual' and 'user view' (in terms of the standard mentioned above) and their relations.

MATHEMATICAL FORMALISMS FOR GEOGRAPHICAL INFORMATION

A computer is a machine whose chief function is to execute a set of instructions which manipulate symbols. These symbols are selected and structured to represent some situation in the real world, as it is perceived by people. The critical problem is that an individual's cognitive methods are informal and fluid, and allow ambiguities or contradictions to exist. The data processing in a computer is formal and follows strict rules of logic; even when computers are used to mimic 'fuzzy' human reasoning, a strict formalism is used to explain the fuzziness (e.g. Zadeh 1974).

In GIS a structure to represent spatial situations (as perceived by people) must be formally defined. Once defined, the appropriate operations and their outcome must also be defined. This is a language issue, as the objects and the operations applicable to them are defined in terms of the GIS user's chosen spatial language (Woodcock and Loomes 1989).

Despite the fact that space is a fundamental, everyday notion, its formalization is not simple. People sometimes look to the objects that fill space, thus treating space as an attribute of the objects (this follows a Kantian viewpoint). Or, it is possible to look at space and the properties of objects that are encountered at each point, and thus attributes and objects become properties of the location (following a more Descartian viewpoint). People use both these methods interchangeably, depending on what is more suitable for the task at hand.

The well-known concepts of Euclidean geometry (see Gatrell 1991 in this volume), which seem to be fully consistent with an appropriate concept of space, in fact help to obscure the problem. In fact, Euclidean geometry captures only limited and highly abstracted aspects of geometry and space. One aspect of the abstraction is the concept of a continuum of possible coordinates, positions and lengths. This may be fully consistent with the view held by scientists, engineers and many other GIS users, but does not apply strictly to all human experience and inference. Perhaps more critical, the real numbers and continuity they apply cannot formally and perfectly be implemented on finite computer systems, and this technical mis-specification may occasionally influence results. For example, in a Euclidean, real-numbered world, a precise polygon overlay procedure (no tolerances or 'snapping') could combine three coverages in a way that would not be order dependent; however, when all coordinates are represented on a computer, order dependency may result.

The conceptual implications of these two major geometrical data models (feature based or location based) are discussed in this chapter. Computer scientists generally use the term 'data model' for the tools or methods that are available to describe the conceptual structure of the data, that is, the language available to describe reality or our perception of reality; this is different from some uses in the GIS literature. The consequences of the internal representation and the data structures necessary to implement these two views, often referred to as 'vector' and 'raster', are treated elsewhere in this volume (Egenhofer and Herring 1991 in this volume). Here, the discussion concentrates only on how people interact with the data stored in the GIS. This is parallel to the discussion regarding data models, primarily the network and the relational model (Codd 1982), in computer science. The discussion of user interfaces here is thus, in principle, on a purely conceptual level and in theory, completely independent from implementation; in practical terms, however, it is not known how to translate between the two major geometrical data models without small but observable differences.

The extension of these concepts to three-dimensional GIS problems will not be trivial; because 3-D GIS is a fairly new and as yet small subfield, such problems will not be discussed here (Raper 1989; Turner 1990; see also Raper and Kelk 1991 in this volume).

Regular tessellation models

A system with a square regular tessellation geometrical data model (a raster) is built on the notion of a subdivision of space into cells of regular size and shape (Fig. 11.2). Methods other than subdivision in squares have been studied (Diaz and Bell 1986) but are not widely used (Fig. 11.3). For implementation in a data structure, regular subdivisions of space, which allow a hierarchical structuring of areas of varying size are very convenient (Fig. 11.4). An example of this is the very popular quadtree structure (Samet 1984). This, however, does not alter the conceptual data model and the behaviour of the operations at the user interface. Everything a raster system can do a quadtree system can do, and vice versa – the only noticeable difference would be the execution speed and other computer resource requirements.

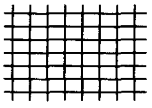

Fig. 11.2. Example raster.

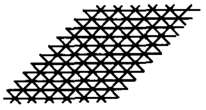

Fig. 11.3. An alternative tessellation.

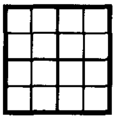

Fig. 11.4. A hierarchical square raster.

For each raster cell, values for all attributes can be made available to the user. The concept can be thought of as an array with integer indices, where each array cell contains a value for each attribute of interest. This can also be visualized as a superposition of similar arrays, which each contain the attribute values for one property.

$$map : property \times i_x \times i_y \text{ -->> } value$$

Based on this geometric data model, a 'map algebra' can be built (Tomlin 1983a, 1983b, 1989). Operations are defined such that they have one or two maps as input and produce a new map according to a specific rule. This output map has exactly the same spatial structure as the input map (but a different content) and can thus be used as input in further operations.

Irregular tessellation model

The alternative is an irregular subdivision of space, an 'irregular tessellation' geometric data model (Fig. 11.5). This model is based on the concept of cells as defined by algebraic topology (Alexandroff 1961; Giblin 1977; Spanier 1966). Each cell has a specific dimension; terminology varies:

0-cell = point, node, vertex;
1-cell = line, arc, edge;
2-cell = area, region, cell.

Each cell is bounded by cells of lower dimension, for example, a line is bounded by two points at each end and each cell is the boundary for some cells of higher dimension, for example, a line is the boundary between two areas. For all points, coordinate values are given that determine their location. Also in some systems, the lines between nodes can have arbitrary shape, in others they are restricted to straight lines. In order to simplify implementation it has been argued that only triangular cells – or generally simplices – should be allowed and other cells subdivided accordingly (Frank and Kuhn 1986).

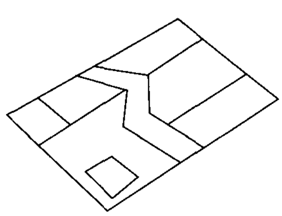

Fig. 11.5. Example of an irregular tessellation.

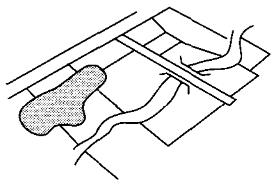

Fig. 11.6. The integrated data model.

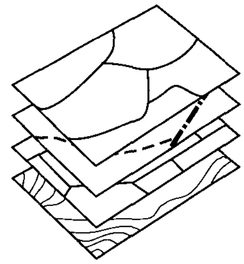

Fig. 11.7. The layered model.

Using an irregular tessellation as a geometrical foundation, three views are possible:

- An 'integrated geometry' concept, where all the layers are integrated at once and the cells are the largest areas for which all attribute values are homogeneous (Fig. 11.6). Nevertheless a 'layer view' can be established as a 'user view' utilizing the operations of the map algebra.
- A 'layer' concept (Fig. 11.7), where each layer represents a single property and the cells in this layer show the largest spatial units that have a common attribute value for this property (e.g. a layer for 'soil classification' and another one for 'ownership'). The user can then overlay, or combine, these layers and produce new layers. The operations are similar to the map algebra mentioned previously.
- An 'object' concept, where attention concentrates on the individual objects, each having geometrical and non geometrical properties. In this view, the fact that all objects fill the space (or even create the space) is not stressed.

Properly built software should allow these three viewpoints to coexist on top of a single data collection. Each of these viewpoints, however, is sufficiently different from the others to require its own language for a user to think about, to formulate queries with, and to understand the responses from the system.

Earlier, GIS were built as collections of points and lines – sometimes referred to as 'spaghetti' (Fig. 11.8). The lines were just lines and all intersections had to be deduced from coordinates. The points represented either point features or areas (centroids) whose boundaries had to be sought in the collection of lines using operations from analytical geometry. Again, this is essentially an 'irregular tessellation' data model as all equivalent properties can be deduced. These deductions are not only costly in terms of computer operations, but also they cannot be executed reliably (i.e. without leading to internal contradictions) on a finite computer. Thus users can observe artefacts which are due to the limitations of the specific implementation.

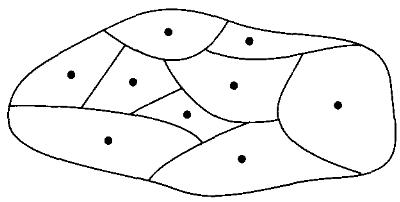

Fig. 11.8. An example of spaghetti: note the undershoots and overshoots at junctions.

In the irregular tessellation data model, topological relations are included (it is thus a 'full topology' model). The neighbours of each cell are known. The topology in the model is very often also used to verify that the actual data are a valid representation of a topologically correct situation (Corbett 1975, 1979). Of course, in a 'regular tessellation' model, topology is also included, although it is implicit in the tessellation pattern and tile numbering system.

The regular tessellation and earlier, line-based representations should in principle be equivalent. In practice, however, they are not – the limitations of computers as finite machines affect the exact behaviour of the operations. It is a reasonable assumption that the spatial resolution of a regular tessellation model (i.e. the fineness of the grid of cells) is much coarser than the spatial resolution of an irregular tessellation model (i.e. how many bits are used for point coordinates); typical examples are 100 m cells for the one and coordinate values with centimetre precision, which results in a 10^4 ratio of resolutions. This results in observable differences between corresponding operations executed in one or the other system.

Other geometric data models

The two models discussed present methods to deal with situations where space is subdivided in mutually exclusive and collectively exhaustive cells. There are phenomena for which these models are not appropriate. In this subsection two additional models are briefly mentioned.

Feature based models

If only isolated features are of interest, like roads, rivers, lakes or settlements, localized in space but without concern for the areas between them, using an irregular tessellation model may be inappropriate. Instead, a number of GIS have been built around a geometrical data model that includes:

- point features,
- linear features,
- areal features,

each defined by sets of coordinates. Details of individual models differ (Burton 1979; Cox, Rhind and Aldred 1980). A number of alternatives are discussed in the proposed US cartographic data exchange standard (Digital Cartographic Data Standards Task Force (DCDSTF) 1988).

Continuous field

Other phenomena, like temperature, magnetic fields, and so on, are thought of as continuous, that is, for every point in space there is a value and the values change gradually. The phenomenon is organized, mathematically speaking, as a function of the location in space $f(x,y)$. If there is a single value (not a vector) associated with each point, the field may be visualized as a continuous surface. The mapping from a point in space to a value can be determined by a mathematical formula or interpolated from measured values. Current GIS are not designed to handle such phenomena directly.

COGNITIVE SCIENCE, NATURAL LANGUAGE AND GEOGRAPHICAL SPACE

What is cognitive science?

The conceptual basis for this treatment of models of geographical space comes primarily from the field of cognitive science. George Lakoff offered the following definition:

> 'Cognitive science is a new field that brings together what is known about the mind from many academic disciplines: psychology, linguistics, anthropology, philosophy, and computer science. It seeks answers to such questions as:
>
> - What is reason?
> - How do we make sense of our experiences?
> - What is a conceptual system and how is it organized?
> - Do all people use the same conceptual system?
> - If so, what is that system?
> - If not, exactly what is there that is common to the way all human beings think?
>
> The questions aren't new, but some recent answers are'. (Lakoff 1987: xi)

Why is such a field introduced here? A fundamental premise of this work is that a main objective of GIS is to allow the user of the system to interact vicariously with actual or possible phenomena of the world (Mark 1989). From this, it follows that models of the human mind, and in particular how our minds deal with concepts and objects of geographical space, are a strict prerequisite to designing effective GIS. This is a bold claim and the need for explicit cognitively based data models as a basis for GIS data structures can be questioned. However, there is little question that effective user interfaces must be based at least in part on such models.

The Rosch–Lakoff–Johnson model of cognitive categories

The mathematical concepts that are typically used to provide data models and structures for GIS include classical set theory and, in particular, the idea of geographical concepts, regions and areal objects as sets. In the classical model, every member of a set is an equally good example of it, and furthermore there is some necessary and sufficient set of properties for determining whether some object is or is not a member of a set. The first part of this model would predict, for example, that every familiar bird would be an equally good example of the class of all birds. However, it is known that, when people are asked to give an example of a bird, they tend to give 'robin' or 'sparrow' or 'canary' far more often than they give 'ostrich' or 'penguin' or 'duck'; this is the concept of prototypes.

The second part of the model (necessary and sufficient observable properties to define set membership) also breaks down when applied to cognitive concepts. The repeated and largely unsatisfying attempts by quantitative and statistical geographers to use techniques such as discriminant analysis to define classic geographical regions such as the American 'corn belt', or 'Appalachia' are an indication that such regions are not equivalent to classical sets.

Whereas such problems with set theory were noted early on, it was the research and writings of Eleanor Rosch (1973, 1978) that provided a clear statement of the problem and summary for the evidence. Several solutions have been proposed, including the 'fuzzy set theory' of Zadeh (1974). Smith and Medin (1981) also treat the problem in great detail. They propose two groups of solutions. One group is probabilistic and is related to fuzzy set theory. While many of the problems of classical set theory are solved, solutions in this group still contain fundamental flaws, especially in the way they treat conjunctions of classes. The other approach discussed by Smith and Medin involves the concept of exemplars. Classes are defined by exemplars and by rules for establishing similarity to these exemplars. The image-schema model of cognition (Johnson 1987) and the broader philosophical position termed experiential realism (Lakoff 1987), appear to provide an even more appropriate basis for concept modelling. Rather than using actual instances as exemplars, classes have idealized or generalized prototypes. These prototypes for classes can in turn be based largely on a small number of schemata that embody properties and transmit them, through prototypes, to class members:

> 'A schema is that portion of the entire perceptual cycle which is internal to the perceiver, modifiable by experience, and

> somehow specific to what is being perceived. The schema accepts information as it becomes available at sensory surfaces and is changed by that information; it directs movements and exploratory activities that make more information available, by which it is further modified.' (Neiser 1976: 54)

Johnson (1987) follows Neiser in claiming that mental activities such as perception and cognition are heavily influenced by what Johnson calls image-schemata, which he defines as follows: 'A schema consists of a small number of parts and relations, by virtue of which it can structure indefinitely many perceptions, images, and events. In sum, image-schemata operate at a level of mental organization that falls between abstract propositional structure, on the one side, and particular concrete images on the other' (Johnson 1987: 29). He goes on to add: '... much of the structure, value, and purposiveness we take for granted as built into our world consists chiefly of interwoven and superimposed schemata...' (Johnson 1987: 126). Recently, Mark (1989) has discussed how an image-schematic model of geographical categories and concepts might operate, and how it could relate to concepts of 'user views' of a geographical database.

How language structures space

In the introduction to an earlier section, the term 'data model' was used for the tools or methods that are available to describe the conceptual structure of the data, that is, the language available to describe reality or perception of it. If 'cognitive/linguistic model' is substituted for 'data model', this is almost exactly the thesis of Leonard Talmy in his seminal paper, 'How language structures space' (Talmy 1983). The basic idea that Talmy presents in that paper is that human natural languages provide individual speakers with a set of terms that are linked to cognitive concepts. Then, consistent with the ideas of schemata discussed above, these mental concepts constrain the way people think, reason, and talk about both perceptual spaces, which can be seen from one viewpoint, and geographical spaces, which must be integrated over repeated experiences of parts of space. The somewhat controversial Sapir–Wharf hypothesis about language and thought extends this logically to the idea that speakers of different languages may think differently, at least about some topics (Lakoff 1987: 304–37).

A central issue involves the concepts of 'language universals' and 'language primitives'. A universal would be a property that applies to all natural languages. Explicit universals may be non-existent, but it is possible to turn to look for primitives, building blocks that may themselves be universal, and from which linguistic expressions can be built. Obviously, identifying such items would be essential to the design of a multilingual, natural language understanding system. Some of the primitives for the language of geographical space seem to be exactly the topological relations used in graph theory and in GIS, as discussed above. The image-schemata of Johnson (1987) form another set of primitives (Mark 1989). The search for still more conceptual primitives for geographical space, and the study of how they combine in particular languages, is an open research question, and a critical part of the research agenda. There is little doubt that geographers and others in the GIS community will make substantial contributions to cognitive science in this area.

Fundamental spatial relations

Freeman (1975) produced an important and early review paper on formal representation of spatial relations. He proposed that the following form a complete set of primitive spatial relations for elements in a (2-D) picture, a view of everyday (3-D) object space:

1. left of;
2. right of;
3. beside (alongside, next to);
4. above (over, higher than, on to);
5. below (under, underneath, lower than);
6. behind (in back of);
7. in front of;
8. near (close to, next to);
9. far;
10. touching;

11. between;
12. inside (within); and
13. outside.

Note that this is not a minimal set of relations, since some can be defined as combinations of others.

Freeman's list is very similar to the list of terms presented by Abler (1987: 306) in his discussion of the research agenda for 'Geographical Information and Analysis'. The cardinal directions can be added to Freeman's list through the addition of one more axiom. If 'north' is associated with 'up', then by deduction, 'south = down', 'west = left', and 'east = right' can follow. Peuquet and Zhan (1987) extended Freeman's relation set in exactly this way, including the cardinal directions as spatial relations without comment, and substituting 'north' for 'above' and 'south' for 'below' in the example they drew from Freeman's paper (Peuquet and Zhan 1987: 66). Note that the 'north = up' axiom is quite arbitrary. Indeed, the etymology of the Indo-European root for the word 'north' is based on 'left' (Svorou 1988); this relation results from an earlier 'east = forward' convention, and world maps in Medieval times were presented with an east up orientation.

Some cultural and linguistic groups, including the Hawaiians, use a radial coordinate system for referencing in geographical space (see Mark, Svorou and Zubin 1987). This uses the 'inside–outside' dichotomy of the Container image-schemata (see Mark 1989) for one spatial dimension, and 'toward some landmark' (spatial action, rather than relation) as the other. Other island peoples use similar spatial reference frames (see Haugen, 1957, for a discussion of this for Iceland).

Herskovits (1985, 1987) has discussed formal and computational models for locative expressions in English. In particular, she discusses about 30 'use types' for the English prepositions 'in', 'on' and 'at'. Recently, Mark (1989) has proposed a link between the image-schemata model of Johnson (1987) and the models of Herskovits. For example, conceptualizing something as a Container means that an English speaker is likely to use the preposition 'in'. The use of 'in' would, in turn, cause the listener or reader to use a similar Container schema in interpreting the meaning of the expression.

SPATIAL QUERY LANGUAGES AND DATABASE INTERFACES

Requirements for a query language

A query language is the tool a user needs in order to extract data from a database and present the result in a useful format. The query is expressed in a language understandable by the query execution program. Thus the query processor must fulfil two important functions:

- Select the subset of the data the user needs.
- Render the selected data in a format that is meaningful to the user.

This section concentrates on the expressive power, that is, what can be specified with a given language, and does not discuss the specific syntax or implementation issue. The discussion is restricted to properties in which spatial query languages differ from ordinary query languages, as they are commonly used for administrative data processing. Eight requirements for a spatial query language are listed, which deal with the selection of data and with its representation (Egenhofer 1989c).

Spatial selection criteria

The user will need to select data to be retrieved, not only based on predicates over attribute values (e.g. the standard question 'select all employees with salary > 50 000') but also based on spatial properties (e.g. 'select all parcels owned by Smith and on or within 100 metres of a lake or pond'). The query language must be extended with predicates to select data based on 'neighbour', 'connected to', 'inside', and so on. Such an extension should be systematic, that is, the set of predicates covers all cases and the predicates have meaningful relations to each other. The predicates must then be given a formal meaning in terms of one of the geometrical data models explained before. An arbitrary set of terms from, say, the English language is not a good starting point. For a subset of spatial relations, namely the topological ones, a systematic set has been established (Egenhofer 1989a). From these proposed base relationships, the user can construct more complex ones, with specific meaning appropriate to the application.

Selection based on pointing

Users of a GIS will naturally ask questions like 'what is this?', or 'who owns this building?'. The query syntax must, therefore, be extended to accept as values objects visible on the screen to which the user points. Integration of pointing gestures with queries is one of the central concepts of the CUBRICON interface (Neal and Shapiro 1990; Neal *et al.* 1989).

These requirements deal primarily with the selection of data for retrieval. The following group of requirements deal with rendering of the result on the screen. The standard query languages, as used in administrative data processing, assume automatically that the result of the query can be displayed in the form of a table; this is obviously not true for a GIS, where many outputs will be in map form.

Combination of query results

The visual integration of the result from more than one query is an important feature of a GIS. It must be possible to specify that the result of a new query is added to (superimposed on) the already displayed map, that it is removed from it, or that the objects selected are highlighted to make them easier to find.

Spatial context

The result of a spatial query cannot always be interpreted by itself, for example, the query 'show the town of Orono' would result in a point with a label on an otherwise empty screen as shown in Fig. 11.9 – this is hardly useful (Egenhofer and Frank 1988; Frank 1982). The query language must, therefore, include a means for the users to specify the necessary context needed to understand the result as in Fig. 11.10 (e.g. 'show AS TOPO SHEET ....', directing the system to use the standard content of a topographic map). It should also include default rules, such as how a query is expanded to include a minimal context, if the user opted not to specify one. This is necessary, because although context is ubiquitous in human interaction, people are not used to thinking about the context explicitly, and thus it is especially onerous for users to deal with context explicitly.

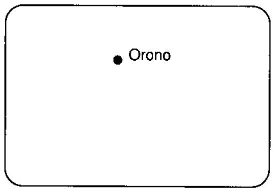

Fig. 11.9. The town of Orono, showing no context.

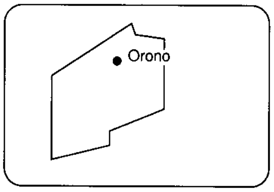

Fig. 11.10. As Fig, 11.9 but with some context (outline of the containing state).

Selection of query window

Access to data in a spatial database is usually explicitly or implicitly restricted by an 'area of interest', for example, 'show all school buildings' means 'show all school buildings in the Orono school administration district'. Again, the query language must include methods to describe the area of interest and contain rules for how a default is selected if the user does not specify one (e.g. use the value of the previous query).

Description of scale

In some instances a display must have a certain scale to be specified by the user.

Description of map legend

The map legend describes a mapping between the data objects in the database and their graphical rendering. The user must be able to select a map legend and change it when necessary.

Differentiate representation based on attribute values

Users often need to observe how an attribute value is spatially distributed (e.g. which houses in a city were built in which year). It is customary in manual cartography to build classes for objects with similar values and assign specific graphical values to them (e.g. colour, raster) (Bertin 1983). Thus the query language for a GIS must essentially include everything that is necessary to specify choropleth and similar maps.

SQL extensions

A number of research projects (Egenhofer 1984; Egenhofer 1989b; Egenhofer and Frank 1988; Frank 1982) and some GIS manufacturers (Herring, Larsen and Shivakumar 1988; Ingram and Phillips 1987) have addressed the problem of constructing a GIS query language by selecting a standard database query language and exploring what extensions would be necessary. Most often the SQL language, based on the relational data model, is used as a starting point (ANSI X3H2 1985; Chamberlin and Boyce 1974; Chamberlin *et al.* 1976). It is very obvious that a language that contains commands to deal with all these problems must be quite extensive and hence may become difficult to learn. Most of the extensions which have been studied fulfil only some of the requirements and are limited to the most important extensions using default methods for other parameters (for an overview see Egenhofer 1989c).

The conclusions drawn from these efforts to extend SQL are:

- Extending SQL with spatial relations and operators is straightforward once the semantics of the relations are formally defined (Pullar and Egenhofer 1988).
- SQL is not as easy to use for complex queries as is often claimed [it is much better than the previous proposals (Reisner, Boyce and Chamberlin 1975), but it was designed in the early 1970's and is considered dated]. GIS queries tend to include complex conditions, which require careful planning for translation into SQL.
- Extending SQL to include pointing as input is feasible, but no syntax or flow of actions that have been found are compatible with natural language/human cognitive patterns. A new keyword PICK (Frank 1982) or MOUSE (Ingram and Phillips 1987) has been created which can be used anywhere an object is required. During query execution the user is then asked to point to the appropriate object.
- It is advisable to build a separate SQL styled command language to deal with the graphical output issues, and not include these commands into the SQL select-from-where clause. This language will become quite complex.

GIS query languages based on direct manipulation

The SQL language which is most often extended for GIS query languages is a typed language with traditional syntax and keywords. Since its design, a new paradigm for constructing user interfaces, based on direct manipulation, has been developed and applied successfully (e.g. the Apple Macintosh personal computer) (Shneiderman 1983, 1987; Smith *et al.* 1983). Some attempts to apply these concepts to query languages in general (Jackson 1990; Kuhn 1991) and to GIS in particular have been studied (Egenhofer and Frank 1988).

It has been found that the construction of a direct manipulation-based interface for a raster-based system is feasible and a number of implementations are known (Intergraph 1989; Pazner, Kirby and Thies 1989; Jackson 1990). The regular conceptual structure of the data model can be translated to visual objects and their manipulations. This could be extended to an 'irregular tessellation' data model, when the 'layer' structure is stressed. This type of interface is based on processes which combine or otherwise manipulate the 'layers'. It is procedural and the user is responsible for combining the processes in the correct order to achieve the desired result.

On the other hand, constructing non-procedural, purely descriptive spatial query languages for GIS with an object oriented user view is more difficult. Translating a keyword based language more or less literally to a screen-based input reduces some of the complexity of the interface – the user need not remember the keywords and is prompted for the parameters – and thus makes the interface more versatile and usable. However, it does not reduce the cognitive complexity of the object and operations.

More promising is the selection of appropriate metaphors and their correct visualization for a GIS query facility. The authors have explored metaphors for the 'pan' and 'zoom' operation which is very powerful in selecting the area of interest and will be expanding it to other suitable tasks, e.g. the selection of content (Jackson 1990; Kuhn 1991).

NATURAL LANGUAGE PROCESSING FOR GIS

Beyond being a basis for design of formal languages and associated data structures, natural language studies are themselves important issues for the designers of GIS for several distinct reasons. The most difficult problem in dealing with natural language in a computational sense is to understand its meaning. People understand each other based on both the formal and conventional structure of their language, and on a large collection of perceptual and cognitive experiences discussed. For a computer, the 'understanding' of natural language is taken to mean that the computer acts on commands, queries and other linguistic input with a response similar to the one that would be given by a person in a similar situation. Similarly, ideal natural language output by a computer will evoke in a human reader reactions that are similar to those which would be given to words and sentences generated by a person. Of course, this is more or less a restatement of the 'Turing test' of artificial intelligence fame, and a general solution is not expected immediately. However, for a limited domain such as GIS, or particularly for some GIS application area, success, or at least substantial progress, seems to be a reasonable expectation.

One important application of natural language processing in GIS involves the potential for the input of queries and commands in natural language; whereas natural language queries in typed form may be of only limited utility, natural language commands and queries will become a common form of system interaction when real-time interpretation of normal speech becomes practical. Next, increasing need for input of geographical data and information in text form is foreseen, ranging from text on biological specimen labels to interpretation of newspaper articles, explorers' journals, or tape-recorded field notes. Natural language production for limited domains, such as generation of verbal descriptions of routes for drivers is already possible, and currently is being extended to other, relatively simple cases and domains, such as the production of legal boundary descriptions for parcels of land. More general natural language text generation is further away, but production of grammatically correct descriptive paragraphs for direct inclusion in reports would be a desirable feature of future GIS.

Natural language input for queries and commands to GIS

In general terms, natural language understanding is still very difficult to implement. The problem of 'understanding' natural language in this sense can be restated as a problem of 'translating' between natural languages and formal languages within a very limited domain. For such a translation, among other things, a better, more formal understanding of spatial terms in natural language is needed, as well as more complete, formal geometrical data models, and, last but not least, translation methods between them. It is apparent that people may use several concepts of geometry and of geographical space, depending on the task, and a single, unified geometrical data model that satisfies all expectations has not yet been found. Therefore, translations between partial formal systems of geometrical reasoning may be necessary. Such translations are, mathematically speaking, mappings between algebraic systems (i.e. morphism) (see Egenhofer and Herring 1991 in this volume).

There are generally three advantages seen in natural languages query and command input:

- The system user needs less training ('everybody knows natural language').
- The user can better represent demands ('natural language is the best representation of human thought').
- Commands can be issued faster ('we speak faster than we type').

It is not clear whether the argument that natural language reduces training requirements applies to the GIS situation: it is based on the assumption that users are conversing with the system about objects of their daily experience for which they possess an adequate vocabulary. This is not necessarily the case. The current systems are still quite restricted in their representation of space and spatial situations. Thus users have to learn how to translate their concepts into the concepts of the system; therefore, training is required to convey to the user the concepts the system uses. It is possible to argue that training the user in the system's formal command language is an effective method to convey to the user the concepts that are utilized in the system. However, these arguments are not yet based on experience, given the lack of natural language systems. Thus the examples in training sessions show not only how commands are used, but also what they achieve and when they can be used – this is in the best of all possible worlds, where training is effectively organized and delivered.

It is further doubtful that the second argument, that natural language is the best expression for ideas, holds. It is not evident that there is a natural language expression for all spatial concepts that well-versed professionals use (e.g. experienced planners). Professional jargon is rife with artificially constructed words or ordinary words that are used with a different meaning. Again training in a formal language would establish a coherent vocabulary. Finally, understanding the natural language input is not sufficient for a system to be usable. If users have to detail every explanation in the most intricate manner and cannot rely on the 'common sense' and general understanding of goals, the circumstances of the task, they will not experience the natural language dialogue as natural.

Thus, natural language may become a very important input in the future for certain GIS applications. Natural language in itself does not solve many of the problems of GIS query languages and most of the problems discussed above in the section on formal query languages apply at least in part to natural language interfaces.

Natural language queries and commands

The process of natural language understanding should deal with the understanding of spoken language. Equipment is currently available to understand a limited vocabulary of spoken natural language, as issued by arbitrary speakers, or to understand a larger (but still limited) vocabulary by a speaker for whom the system has been trained. Typically, such equipment can only recognize words spoken with clean breaks between them (so-called discrete speech), and cannot cope with the continuous speech that people usually utter. Systems should also understand unrestricted vocabulary. Programs to analyse typed natural language sentences with a limited domain vocabulary and somewhat limited syntactical structure are available; users of these systems must either rephrase statements or be able to train the system to understand new terms.

For effective natural language interaction with GIS, systems will need to understand dialogue and not just isolated sentences. Research in dialogue understanding is underway, and has been applied to GIS situations. Such research should lead to usable systems within a few years. It also will be necessary to understand gestures and other non-verbal input, integrated with speech, since discussion of spatial situations between people typically involves substantial amounts of non-verbal gestures, sketches, and so on (Neal and Shapiro 1990). Despite some interesting results in advanced research the routine use of such tools is still a few years away.

Input of textual geographical data to GIS

Considerable amounts of geographical data are collected not in the form of maps and diagrams but in textual form. In perhaps the most prominent current example, there are many millions of biological specimens in museums and herbaria; current efforts to computerize such collections are similar to the production of computerized catalogues for libraries, but also include a desire to geocode the data to allow for mapping and for entry into GIS (McGranaghan and Wester 1988). But, although these specimens have labels on which the locality data and collecting date are indicated, the location is hardly ever in the form of coordinates or a map – rather, it is in the form of natural language (McGranaghan 1989). For the small labels on birds and insects, the location may be just a place name. For plant specimens, which usually are mounted on paper, providing more room for description, the place name often is supplemented by a verbal description, roughly equivalent to instructions for relocating the site (McGranaghan 1989). If such data can be automatically analysed and translated to spatial locations in a GIS, they could be subject to both mapping and advanced forms of spatial analysis, and thus become far more valuable, especially for endangered species, which are often more common as museum specimens than as living examples.

Another current practical problem is understanding boundary descriptions of properties in deeds, and the translations among coordinate, graphic and verbal representations. In most jurisdictions, the contract for selling (or otherwise conveying) a parcel of land must include a description of the land, that is, its boundary. This description is often in verbal form and in many parts of the world the verbal description is the chief legal document. Most counties in the United States collect large amounts of such data, which are legally relevant, in their registry of deeds, but such data are not easily accessible for GIS. These data are needed in map form, such as the maps used by tax assessors. Descriptions of the boundaries of the ranges of biological populations are similar and often are found in reference volumes and checklists.

The input of other forms of geographical information in verbal form becomes more speculative and futuristic. It is conceivable that systems of the future might be able to assimilate and analyse explorers' journals, such as Columbus's logs or the journals of Lewis and Clark. They could be checked for consistency and perhaps new inferences could be made about the itineraries of their travels. Field workers of the future might be able to speak their notes into tape-recorders and later have the tapes not only transcribed but also analysed and integrated directly into GIS. A number of agencies and companies would be interested in devices which would accept spoken descriptions of locations from their field personnel and integrate their observations with map data. As appealing as these ideas may be, the development of such applications may be many years away.

Natural language production for GIS

Generation of verbal descriptions of routes for drivers already is possible, and indeed is an option of at least one current commercial GIS: the ARC/INFO 'Directions' command produces a verbal description of a route through a street network (ESRI 1989: Chapter 4: 9). Also, there are commercially available real-time, computer-based navigation aid systems for vehicles (ETAK 1988; Zavoli 1989) which produce in-car maps for use in route planning. There is considerable evidence, however, suggesting that driving instructions in verbal form may be more effective than maps (see McGranaghan, Mark and Gould, 1987, for a review). Davis (Davis 1986; Davis and Schmandt 1989) has described a system to provide driving directions over cellular phone systems. Experimental work also has investigated the idea of using complexity of verbal description as a cost heuristic for route selection itself, finding 'simplest-to-describe' paths rather than 'shortest' paths (Ma 1987; Mark 1985). This is a fertile area for further research.

Again drawing on speculation, the GIS of the future might produce grammatically correct paragraphs for direct inclusion in reports. One example is to go from a discrete set of observation points for some species of animal or plant, to a polygon representing the range of that organism, to a clear verbal description of where the organism can (or could) be found. Verbal descriptions of patterns, shapes and spatial relations are important parts of environmental impact reports, and again their generation by GIS would probably be desirable.

CROSS-LINGUISTIC ISSUES FOR GIS

In designing the GIS user interface, the GIS community should also pay attention to differences among various natural languages as to how they represent and express concepts, relations and objects of geographical space. It is most likely that currently favoured query languages are influenced by the natural language of their designers, commonly English or occasionally German. Without conscious choice, designers tend to select terminology and concepts based not only on their own everyday use of natural language, but also on word order and other major structural elements of that language.

Unless underlying concepts are explicitly and conscientiously used, compared and translated, the use of GIS in non-English speaking areas may be severely impeded. Any user–computer interface bridges the user's cognitive structure and the representation in the computer system. Obviously, constructing an interface must take into account both components. If a GIS is moved from one linguistic culture (say North American English) to another (say South American Spanish), some form of 'translation' of the interface is highly desirable, if not essential. In fact, many professionals and technicians in the non-English-speaking world have some working knowledge of English, and thus are able to use English-language software; but this is far from desirable, and restricts access to the technology to a very small subset of the population.

The 'translation' of an interface would include the actual translation of the words used in the interface, including commands, menus and help texts, but such translations are not always straightforward, since the concepts underlying the interface may not match those of the target language. Cross linguistic transfer of computer technology certainly calls for more than the translation of the manuals. Customarily that is the most that is done, but the GIS industry has not yet touched on the deeper issues. Computer science and the computer industry have studied systems supporting languages which are not based on the roman alphabet, and how such differences affect database query language (see King 1989 and other papers in the same volume).

Observations of adaptations of cultures to new languages often show a facility to adopt a new vocabulary, but to persist in using elements of the underlying structures of previous language. By analogy, it is suggested that the translation of a user interface's 'surface' vocabulary addresses the less urgent part of the problem, since using new words for old concepts may be fairly easy for users to adapt to in any case.

The first reaction to issues such as these is, almost invariably, to build more 'flexibility' into the GIS, and to make the interface more adaptable by the user. If this flexibility is well designed, it can indeed be used to adapt to individual differences. Ideally, the base structure of the programs – the geometrical data model and its operations – is available, and furthermore completely devoid of artefacts of the cognitive structure and linguistic traditions of the designer(s). However, in practice it is doubtful that this is ever the case. Otherwise, the construction of the interface has to translate the data model into a structure akin to the cognitive structure of the class of users. This is clearly no simple task and it will be desirable to identify parts which are useful everywhere, and others which depend on certain categories and which vary between target languages. The exploration of 'language universals', discussed briefly above, thus becomes very important, as it would allow the separation of what is generally applicable from what needs to be adapted to a local language, culture, or subculture. These cross linguistic issues for GIS have been addressed in more detail by Mark, Gould and Nunes (1989).

CONCLUSION

Observing natural language usage to describe spatial situations is very important for the designers of GIS to achieve a system which is compatible with the way users conceptualize their problem domains. Since current GIS structures are based largely on maps and mathematics, and since maps and mathematics represent previously formalized representations of naive concepts of geographical space, adoption of a cognitively based concept structure for GIS will not necessarily involve radical changes.

The most important issue is to understand the separation between the conceptual view of GIS, which explains how the system operates to the user community, and the implementation, which should be of interest only to programmers and systems maintenance personnel. The GIS literature continues to mix the two sets of issues, beginning even before the famous 'First International Study Symposium on Topological Data Structures for Geographical Information Systems' (Dutton 1979) hosted by Harvard University in 1977, and continuing to the present. The conceptual design, the geometrical and attribute data models, the user interface style – all of these have fundamental ties to cognitive and linguistic research, whereas the implementation of the GIS software connects to computer science.

To the extent that they have been recognized at all, linguistic issues in GIS have in the past been seen primarily as issues for user interface design, especially in spatial query languages. However, recent research in these areas has shown that a GIS needs an approach based on a 'dialogue' model, and that the standard 'question and answer' model is insufficient (Mark *et al.* 1989a; Neal *et al.* 1989). The debate concerning whether natural language interfaces are more appropriate than formal system interaction 'languages' is ongoing (Shneiderman 1981) and definite results cannot be expected before some of the technical limitations of understanding naturally spoken language (continuous speech) have been removed.

The cognitive and linguistic issues in GIS span a much wider set of issues, and touch on the general problem of how to translate users' concepts into executable operations. The current model, which is to train users in translating their task into a procedural, formal geometrical problem (which is then submitted to the GIS) severely limits GIS technology to trained users. The relative merits of the tendency of technology designed in this way to promote and maintain a specialized class of 'gurus' to act as 'gate-keepers' to the technology is beyond the scope of this chapter, and has been discussed in detail by Winograd and Flores (1986). To go beyond this approach needs a more profound understanding of how people in general think and reason about space and things spatial.

Last, but not least, the structure of GIS interfaces, as available today, is, essentially, the product of an Anglo Saxon (or at least, Germanic) cognitive and linguistic culture. In order to make GIS useful in other language groups and cultures, it may not be sufficient to translate the 'surface structure' of the systems, such as the command language, or menu contents, or manuals. Instead, attention must be given to the deeper syntactic and cognitive structures that underlie other languages. It is clearly desirable to build alternatives to the 'verb-oriented' languages currently used for commands and queries. A metaphor-based, direct manipulation interface is clearly an attractive alternative, but such interfaces have their own problems. Unfortunately, the most appropriate metaphors and associated visualizable image-schemata are not known for geographical information for any natural language, let alone across many languages. Also, building such a visual interface will not resolve the cross-linguistic problems of GIS technology transfer and use, since the use of visual symbols often is as much culturally and linguistically determined as are the languages themselves.

It is to be expected that these issues will increase greatly in importance and recognition within the GIS community world wide.

REFERENCES

Abler R (1987) The National Science Foundation National Center for Geographic Information and Analysis. *International Journal of Geographical Information Systems* **1** (4): 303–26

Alexandroff P (1961) *Elementary Concepts of Topology*. Dover, New York

ANSI X3H2 (1985) *American National Standard Database Language SQL*. American National Standards Institute, Washington DC

Bertin J (1983) *Semiology of Graphics*. University of Wisconsin Press, Madison, Wisconsin

Burton W (1979) Logical and physical data types in geographic information systems. *Geo-Processing* **1** (4): 167–81

Chamberlin D D, Boyce R F (1974) Sequel: a structured English query language. In: Rustin R (ed.) *Workshop on Data Description, Access and Control*. ACM SIGMOD, Ann Arbor, Michigan, 249–64

Chamberlin D D, Astrahan M M, Eswaran K P, Lorie R A, Mehl J W, Reisner P, Wade B W (1976) SEQUEL 2: a unified approach to data definition, manipulation, and control. *IBM Journal of Research and Development* **20**: 560–75

Codd E F (1982) Relational database: a practical foundation for productivity. *Communications of the ACM* **25** (2): 109–17

Corbett J P (1975) Topological principles in cartography. *Proceedings, International Symposium on Computer-Assisted Cartography, AUTOCARTO2*. Reston, Virginia, US Department of Commerce

Corbett J P (1979) Topological principles of cartography. *Technical Report 48*. Bureau of the Census, US Department of Commerce, Washington DC

Cox N J, Rhind D W, Aldred B K (1980) A relational database system and a proposal for a geographic data type. *Geo-Processing* **1**: 217

Davis J R (1986) Giving directions: a voice interface to a direction giving program. *Proceedings, 1986 Conference, American Voice I/O Society*. September, pp. 77–84

Davis J R, Schmandt C M (1989) The back seat driver: real time spoken driving directions. *Proceedings, First Vehicle Navigation & Information Systems Conference (VNIS '89)*. IEEE, New York, pp. 146–50

Diaz B M, Bell S B M (eds.) (1986) *Spatial Data Processing using Tesseral Methods (collected papers from Tesseral Workshops 1 and 2)*. NERC Unit for Thematic Information Systems, Natural Environment Research Council, Swindon

Digital Cartographic Data Standards Task Force (1988) The proposed standard for digital cartographic data. *The American Cartographer* **15** (1): 9–140

Dutton G (ed.) (1979) *First International Study Symposium on Topological Data Structures for Geographic Information Systems*. Addison-Wesley, Reading, Massachusetts

Egenhofer M J (1984) Implementation of MAPQUERY, a query language for land information systems (in German).

Report 79. Institute for Geodesy and Photogrammetry, Swiss Federal Institute of Technology (ETH), Zurich

Egenhofer M J (1989a) A formal definition of binary topological relationships. In: Schek W L, Schek H-J (eds.) *Proceedings, Third International Conference on Foundations of Data Organization and Algorithms (FODO), Paris*. Springer-Verlag, New York, pp. 457–72

Egenhofer M J (1989b) *Spatial Query Languages*. Unpublished PhD dissertation, University of Maine

Egenhofer M J (1989c) Spatial SQL: a spatial query language. *Report 103*. Department of Surveying Engineering, Orono Maine

Egenhofer M J, Frank A U (1988) Towards a spatial query language: user interface considerations. *Proceedings, 14th International Conference on Very Large Data Bases, Los Angeles*. Morgan Kaufmann, Los Altos, California, pp. 124–33

Egenhofer M J, Herring J R (1991) High-level spatial data structures for GIS. In Maguire D J, Goodchild M F, Rhind D W (eds.) *Geographical Information Systems: principles and applications*. Longman, London, pp. 227–37, Vol 1

ESRI (1989) *Network Users Guide*. Environmental Systems Research Institute, Redlands California

ETAK (1988) *ETAK MapEngine, Programmers Guide*. ETAK, Menlo Park California

Frank A U (1982) MAPQUERY – database query language for retrieval of geometric data and its graphical representation. *ACM SIGGRAPH* **16** (3): 199–207

Frank A U, Kuhn W (1986) Cell graph: a provable correct method for the storage of geometry. *Proceedings of the 2nd International Symposium on Spatial Data Handling, Seattle*. International Geographical Union, Williamsville New York, pp. 411–36

Freeman J (1975) The modelling of spatial relations. *Computer Graphics and Image Processing* **4**: 156–71

Gatrell A C (1991) Concepts of space and geographical data. In: Maguire D J, Goodchild M F, Rhind D W (eds.) *Geographical Information Systems: principles and applications*. Longman, London, pp. 119–34, Vol 1

Giblin P (1977) *Graphs, Surfaces and Homology*. Chapman and Hall, London

Haugen E (1957) The semantics of Icelandic orientation. *Word* **13**: 447–60

Herring J R, Larsen R, Shivakumar J (1988) Extensions to the SQL language to support spatial analysis in a topological data base. *Proceedings of GIS/LIS '88*. ASPRS/ACSM, Falls Church, pp. 551–60

Herskovits A (1985) Semantics and pragmatics of locative expressions. *Cognitive Science* **9**: 341–78

Herskovits A (1987) *Spatial Prepositions in English*. Cambridge University Press, Cambridge Massachusetts

Ingram K J, Phillips W W (1987) Geographic information processing using a SQL-based query language. *Proceedings of the Eighth International Symposium on Computer-Assisted Cartography, Baltimore*. ASPRS/ACSM, Falls Church, pp. 326–35

Intergraph Corp (1989) *Tigris Imager Reference Manual*. Intergraph Corporation, Huntsville

Jackson J (1990) Developing an effective human interface for geographical information systems using metaphors. *ACSM/ASPRS Annual Convention* **3** (1): 117–25

Johnson M (1987) *The Body in the Mind: the bodily basis of meaning, imagination and reason*. University of Chicago Press, Chicago

King R (1989) Introduction to the special issue on non-English interfaces to databases. *IEEE Transactions on Database Engineering* **12** (4): 1–7

Kuhn W (1991) Are displays maps or views? *Proceedings of AUTOCARTO10*. ACSM/ASPRS, Bethesda Maryland

Lakoff G (1987) *Women, Fire, and Dangerous Things: what categories reveal about the mind*. University of Chicago Press, Chicago

Ma P (1987) An algorithm to generate verbal instructions for vehicle navigation using a geographic database. *East Lakes Geographer* **22**: 44–60

Mark D M (1985) Finding simple routes: 'ease of description' as an objective function in automated route selection. *Proceedings, Second Symposium on Artificial Intelligence Applications, Miami Beach*

Mark D M (1989) Cognitive image-schemata for geographic information: relations to user views and GIS interfaces. *Proceedings of GIS/LIS '89*, Vol. 2. ASPRS/ACSM, Falls Church, pp. 551–60

Mark D M, Gould M D, Nunes J (1989) Spatial language and geographic information systems: cross-linguistic issues. *Proceedings, II Conferencia Latinoamericana sobre el (Technologia de los Sistemas de Informacion Geograficos (SIG)*. Universidad de Los Andes, Merida, Venezuela, pp. 105–30

Mark D M, Svorou S, Zubin D (1987) Spatial terms and spatial concepts: geographic, cognitive, and linguistic perspectives. *Proceedings, International Symposium on Geographic Information Systems: The Research Agenda*, Vol. II. National Aeronautics and Space Administration, Washington DC, pp 101–12

Mark D M, Frank A U, Egenhofer M J, Freundschuh S M, McGranaghan M, White R M (1989a) Languages of spatial relations: Initiative Two specialist meeting report. *Technical Report 89–2*. National Center for Geographic Information and Analysis, Santa Barbara California

McGranaghan M (1989) Context-free recursive-descent parsing of location-description text. *Proceedings, Ninth International Symposium on Computer-Assisted Cartography*. ACSM/ASPRS, Falls Church, pp. 580–7

McGranaghan M, Mark D M, Gould M D (1987) Automated provision of navigation assistance to drivers. *The American Cartographer* **14**: 121–38

McGranaghan M, Wester L (1988) Prototyping an herbarium collection mapping system. *Proceedings 1988 ACSM-ASPRS Annual Convention*. ACSM/ASPRS, Falls Church, pp. 232–8

Neal J G, Shapiro S C (1990) Intelligent multi-media

interface technology. In: Sullivan J W, Tyler S W (eds.) *Architectures for Intelligent Interfaces: elements and prototypes*. Addison-Wesley, Reading Massachusetts
Neal J G, Thielman C Y, Dobes Z, Haller S M, Shapiro S C (1989) Natural language with integrated deictic and graphic gestures. *Proceedings, DARPA Speech and Natural Language Workshop*. Morgan Kaufmann, Los Altos CA
Neiser U (1976) *Cognition and Reality: principles and implications of cognitive psychology*. Freeman, San Francisco

Pazner M, Kirby K C, Thies N (1989) *MAP II Map Processor*. Wiley, New York
Peuquet D, Zhan C-X (1987) An algorithm to determine the directional relationship between arbitrarily-shaped polygons in a plane. *Pattern Recognition* **20**: 65–74
Pullar D, Egenhofer M J (1988) Towards formal definitions of topological relations among spatial objects. *Proceedings of the 3rd International Symposium on Spatial Data Handling, Sydney*. International Geographical Union, Columbus OH, pp. 225–41

Raper J F (1989) *Three Dimensional Applications in GIS*. Taylor and Francis, London
Raper J F, Kelk B (1991) Three-dimensional GIS. In: Maguire D J, Goodchild M F, Rhind D W (eds.) *Geographical Information Systems: principles and applications*. Longman, London, pp. 299–317, Vol 1
Reisner P, Boyce R F, Chamberlin D D (1975) Human factors evaluation of two database query languages – Square and Sequel. *Proceedings, National Computer Conference (AFIPS)*, pp. 447–52
Rosch E (1973) On the internal structure of perceptual and semantic categories. In: Moore T E (ed.) *Cognitive Development and the Acquisition of Language*. Academic Press, New York, pp. 111–44
Rosch E (1978) Principles of categorization. In: Rosch E, Lloyd B B (eds.) *Cognition and Categorization*. Erlbaum, Hillsdale New Jersey, 27–48

Samet H (1984) The Quadtree and related hierarchical data structures. *ACM Computing Surveys* **16**: 187–260
Shneiderman B (1981) A note on human factors issues of natural language interaction with database systems. *Information Systems* **6** (2): 125–9
Shneiderman B (1983) Direct manipulation: a step beyond programming languages. *Computer* **16**: 57–69
Shneiderman B (1987) *Designing the User Interface: strategies for effective human-computer interaction*. Addison Wesley, Reading Massachusetts
Smith D C, Harslem E, Irby C, Kimball R, Verplank W (1983) Designing the Star user interface. *Proceedings, European Conference on Integrated Interactive Computing Systems: Stresa, Italy*. North-Holland, Amsterdam
Smith E E, Medin D L (1981) *Categories and Concepts*. Harvard University Press, Cambridge Massachusetts
Spanier E (1966) *Algebraic Topology*. McGraw-Hill, New York
Svorou S (1988) *The Experiential Basis of the Grammar of Space: evidence from the languages of the world*. Unpublished PhD dissertation, Department of Linguistics, State University of New York at Buffalo

Talmy L (1983) How language structures space. In: Pick H, Acredolo L (eds.) *Spatial Orientation: theory, research, and application*. Plenum, New York, pp. 225–82
Tomlin C D (1983a) *Digital Cartographic Modeling Techniques in Environmental Planning*. Unpublished PhD dissertation, Yale University
Tomlin C D (1983b) A map algebra. *Proceedings, Harvard Computer Graphics Conference*. Cambridge, Massachusetts
Tomlin C D (1989) *Geographic Information Systems and Cartographic Modeling*. Prentice Hall, Englewood Cliffs New Jersey
Tsichritzis D, Klug A (eds.) (1975) *The ANSI/X3/SPARC DBMS Framework Report of the Study Group on Database Management Systems*. AFIPS Press, Montvale, New Jersey
Turner A K (1990) *Three-Dimensional Modeling with Geoscientific Information Systems*. NATO Advanced Research Workshop

Winograd T, Flores F (1986) *Understanding Computers and Cognition: a new foundation for design*. Addison-Wesley, Reading Massachusetts
Woodcock J, Loomes M (1989) *Software Engineering Mathematics*. Addison-Wesley, Reading Massachusetts

Zadeh L A (1974) *Fuzzy Logic and its Application to Approximate Reasoning, Information Processing*. North-Holland, Amsterdam
Zavoli W B (1989) Navigation and digital maps interface for fleet management and driver information systems. *Proceedings, First Vehicle Navigation & Information Systems Conference (VNIS '89)*. IEEE, New York, pp. A9–A14