INT. J. GEOGRAPHICAL INFORMATION SCIENCE, 2000, VOL. 14, NO. 6, 501–520

Research Article

A conceptual framework for incorporating cognitive principles into geographical database representation

JEREMY L. MENNIS, DONNA J. PEUQUET

Department of Geography, The Pennsylvania State University, University Park, Pennsylvania, 16802, USA; e-mail: jmennis@gis.psu.edu and peuquet@geog.psu.edu

and LIUJIAN QIAN

Spatial Products, Oracle Corporation, One Oracle Drive, Nashua, New Hampshire 03062, USA; e-mail: lqian@us.oracle.com

(*Received 15 April 1999; accepted 27 December 1999*)

Abstract. The advancement of GIS data models to allow the effective utilization of very large heterogeneous geographic databases requires a new approach that incorporates models of human cognition. The ultimate goal is to provide a cooperative human-computer environment for spatial analysis. We describe the pyramid framework as an example of this new approach within the context of some important aspects of how humans conceptually store spatial information. The proposed framework provides the means to create multiple structural interpretations of observed geographic data and the ability to build knowledge hierarchies through the application of data mining and other statistical techniques.

1. Introduction

The representation of geographic phenomena in digital databases is one of the most central and fundamental issues in Geographic Information Science. Although robust and efficient techniques for representing a wide variety of modern business enterprise data have been developed, the availability of representational techniques that are up to the task of modern GIS requirements for very large, shared geographic databases is still problematic. The representation of geographic phenomena in a way that is appropriate for a wide range of application contexts has also proven to be particularly challenging in recent research (UCGIS 1996, Mark 1999). One reason for these deficiencies is that the conceptual models currently employed for digital geographic data representation do not incorporate any explicit consideration of how humans cognitively store and use geographic knowledge (Peuquet 1988, 1994, Couclelis 1993). The shortcomings of these conventional GIS data models' ability to represent information in a way that is more 'natural' to humans is now widely acknowledged (Peuquet 1988, Frank 1992, Burrough and Frank 1995).

Conventional geographic databases have only been concerned with the storage of observational data; higher-level knowledge derived from that data resided only in

International Journal of Geographical Information Science
ISSN 1365-8816 print/ISSN 1362-3087 online © 2000 Taylor & Francis Ltd
http://www.tandf.co.uk/journals

the head of the user, or in a separately defined model. This was effective when data sets were relatively small, and the analyst could rely on his or her own knowledge to guide searches of the data and interpret it. However, given the current amount, variety, and rate of accumulation of digital data, it is far beyond human capability alone to effectively do this.

Many explorations in improving methods for geographic representation have employed the standardized relational data model of Relational Database Management Systems (RDBMS) and object oriented data model design techniques (e.g. Herring 1992, Worboys 1994). Nevertheless, geographic data, whether they are called vector/raster or object/field representations, are still conceptually divided into the same point, line, area or grid cell units within the 'wrapper' of RDBMS or object oriented design. In essence, these approaches are generally utilized at the implementational level with little change to the underlying conceptual model.

There is a considerable tradition within Geography that explores how humans cognitively represent geographic-scale environments (Downs and Stea 1973, Golledge 1993). We believe it is this perspective, combined with the variety of cognitive evidence from Psychology, that needs to be integrated into GIS database design in order to advance geographic database representation. A fundamentally new approach to database modelling is needed that begins at the conceptual level. A GIS database should present a logical view of the data, as well as the derived higher-level knowledge that corresponds to people's own cognitive view. While a number of GIS researchers have explored characteristics of cognitive representation and conceptual modelling (e.g. Nyerges 1991, Usery 1993, Mark and Frank 1996), they have not extended their findings into a usable framework for database representation. Raper and Livingstone (1995) perhaps have come closest, having developed a specific representation for observational geomorphologic data within a cognitive context.

In this paper we describe a general framework for geographic representation that seeks to integrate the principles of cognition into geographic database representation. We do not attempt to imitate every known aspect of human cognition within a database representation; rather, we identify the primary organizational elements of cognition that can provide a generic framework within which to facilitate human-machine interaction and at the same time enhance computing efficiency. The proposed framework and the conceptual modelling approach built upon it allow for the derivation of an implementable geographical model capable of representing both observational data, as well as higher-level semantic abstractions that can be derived from that data.

The framework presented in this paper, which we call the pyramid framework, is intended to: 1) allow the non-exhaustive search of observational data, 2) facilitate the exploration of new data by quickly identifying patterns and anomalies, and 3) provide a more integrated modelling environment than is currently possible in conventional GIS. In the remainder of this paper, we first review key principles of how humans store and acquire geographic knowledge and how these principles have been integrated into geographic database representations thus far. We then describe the pyramid framework and how it incorporates these key elements. Finally, we present an example conceptual model for representing meteorological phenomena.

2. Cognitive representation

2.1. *Category theory*

Cognitively, we reduce the vast amount of knowledge we store to manageable proportions by grouping knowledge into categories. Categories are also key to

organizing our available synthesized spatial knowledge. They are the repositories of our beliefs about the world, including known entities, the known properties of those entities and relationships between them (Stevens and Coupe 1978, Tversky and Hemenway 1983, Eastman 1985). Besides *things*, we can categorize places, events, actions, spatial relationships, social relationships, and many other types of entities, both concrete and abstract, over an enormous range (Lakoff 1987). We also group categories into higher-level categories that are related in complex ways.

A category is a conceptual grouping of entities that are somehow considered similar, or are treated in a similar way. They are assumed to share common properties, so that these properties can also be considered characteristic of the category. We have categories for everything we think about, and we tend to attribute a real existence to them. The groundbreaking research of Eleanor Rosch (Rosch 1973, 1978, 1983, Rosch and Mervis 1975) and subsequent research has since demonstrated that categories as known in the mind are a function of properties of the elements within the environment, but as *interpreted* by the perceiver. In other words, we understand the world in terms of our own stored knowledge, which is organized by categories.

Given Rosch's (1978) assertion that the form of cognitive category systems allows the retention and use of a maximum amount of knowledge with minimum effort, she further posited that categories are arranged in hierarchies. The vertical dimension of any particular hierarchical system of categories (e.g. animal, dog, collie) concerns the amount of inclusiveness of the category. Categories at the same level across a horizontal of the same hierarchical system (e.g. dog, cat, horse) would have the same amount of inclusiveness. This, then, forms a taxonomy, wherein categories are related to one another by means of class inclusion. The greater the level of abstraction, the greater the inclusion and the higher the location in the hierarchy. Attributes that are common to a class of objects are stored with that class, and all subordinate categories inherit those common attributes. This functionally allows information common to groups to be stored with minimum repetition.

The implication of this is that some levels of the hierarchy are more useful than others in that they relate more directly to how elements of the world are perceived. The most 'useful' level of category was termed by Rosch (1978) to be the *basic-level* category. It is at this level that properties attributed to the specific categories mirror those perceived in the real world. These are characterized by overall shape and motor interaction and are the most general level in the hierarchy at which a mental image can be formed. Thus, 'chair' is a basic-level category while 'furniture' is not. Basic-level categories are located in the middle of any given taxonomy, and are the first to be learned. Superordinate categories are added with increasing knowledge and are useful when referring to groups of objects (Murphy and Wisniewski 1989). Superordinate categories also provide spatial, taxonomic, and other types of linkages between entities as explicitly stored relationships.

Our stored knowledge is, generally speaking, organized by *kinds* vertically within the hierarchy (taxonomies), and by *parts* (partonomies) within categories. At the basic level, our knowledge is mostly concerned with *parts* (Rosch 1978, Tversky and Hemenway 1984). For example, chairs have seats. Parts determine shape, and hence the way that an object will be perceived and imaged. We also impose this part-whole structure on events, implying that our knowledge of event categories is structured very much like our knowledge of physical object categories. (Lakoff and Johnson 1980).

To increase the distinctiveness and flexibility of categories at all levels vertically in any hierarchy, categories tend to become defined in terms of *prototypes* that

contain the attributes most representative of items within that category. In other words, categories have a *graded* internal structure (Rosch 1978). Thus, beanbags and barstools would be less central members of the category 'chair' than the straight-backed, four-legged type usually found in kitchens and dining rooms.

2.2. *Category theory as applied to geographic knowledge*

How does category theory apply to geographic categories? For objects that have a physical manifestation in the real world (i.e. they are not totally abstract), shape is still among the most important criteria for identification. The shape identified with a particular type of geographic object can include an overall shape, as well as the shapes of key component parts. As an example of geographic-scale objects, we visually distinguish churches from other kinds of large structures primarily by their distinctive overall shape, which is provided mainly by the distinctive shape of steeples, a key element characteristic of churches.

Certainly, geographical entities also tend to be taxonomically arranged in nested hierarchies. For example, country, region, state, city, and neighbourhood are conceptually arranged in a hierarchical order. There is also a container relationship built into this hierarchy. In other words, for any set of specific examples (e.g. the US, New England, Massachusetts, Boston, and Beacon Hill) each entity is spatially contained within the entity next above it in the hierarchy. A nested hierarchical structure of cognitive geographic knowledge has been experimentally confirmed by a number of researchers (Stevens and Coupe 1978, Eastman 1985).

Tversky and Hemenway (1983) investigated whether or not Rosch's basic-level categories could be extended to environmental scenes, as well as geographic-scale objects. They found that, in human cognition, the preferred level of description of objects embedded in spatial contexts is at the basic level. This level allows the greatest amount of information about the scenes and objects involved. Thus, categories such as *beach* and *city* provide much more information than the superordinate category of *outdoors*, and the subordinate categories *industrial city* and *Midwestern city* provide only marginally more information. Most information stored about basic-level scenes and objects are observable objects and their properties. Indeed, the mention of the term *city* to most people evokes images of tall buildings, dense residential development and lots of traffic, among other attributes. Lloyd *et al.* (1996) assert that the notion of key information contained within basic-level categories being of parts also translates into a container relationship for spatial categories: countries contain regions, which contain states, which contain cities, which contain neighbourhoods.

The concept of *schema* addresses the cognitive use of a priori knowledge to interpret and categorize new observational data. A schema is essentially a set of information about a type of object (or other cognitively distinguishable 'thing', such as an event) that is used to discover new instances of this type of object (Thorndyke 1984). For example, my schema for the visual recognition of the object 'church' may include such information as: is a large building, has stained-glass windows, has a square or cross-like shape, made of brick or stone, etc. When I recognize an object that meets enough of the church schema criteria, I identify the object as an instance of the church category with specific values for the generic properties described by the schema; a specific church has a certain size, a certain shape, and is made out of a specific material. Schema are formed by induction from repeated experience with the same type of object and may be based on the prototype example of a category of objects. A schema is not an exact representation, but is more like a general pattern.

2.3. *Separation of different types of geographic knowledge*

There is also a body of evidence suggesting that we cognitively store *what*, *where*, and *when* knowledge in separate categorical hierarchies, or knowledge structures, that capture differing characteristics for differing purposes (Ungerleider and Mishkin 1982, Anderson 1987, Farah *et al.* 1988, Sergent 1991). For non-abstract objects, knowing *what* an object is and being able to identify a specific type of object when encountered involves detailed and precise geometric properties, particularly shape and the relative orientation of component parts, in addition to other attributes such as colour and size. On the other hand, knowing *where* something is seems to be embedded in a knowledge of environmental scenes. This knowledge includes relative locational properties of objects, but these properties, such as 'containment', 'distance', and 'direction', tend to be very generalized. For example, farms tend to be located *outside of* cities.

Kosslyn and Koenig (1992) note that determining temporal relationships is qualitatively different from information relating to identity or location. There is thus a third system of what they term a motion relations subsystem, corresponding to what can be called the *when* knowledge representation system. For instance, as we move through our environment, we experience landmarks and other elements in specific temporal sequences, depending upon the direction we are travelling in space. How quickly we encounter these sequences of elements also varies according to how quickly or slowly we are moving. In the dynamics of urban change, *when* a specific factory closed can be at least as important as *where* in explaining a subsequent urban arrangement of land use.

Time relates to grouping information in two ways: first, properties of the patterns of movement, per se, define how moving objects will be grouped. For example, if we see a lot of houses being built in a particular area, we see them as a single unit (i.e. a subdivision), not as individual houses. This is called the Gestalt Law of Common Fate (Kosslyn and Koenig 1992). Second, a specific pattern of movement that tends to happen repeatedly in sequence will also be viewed as a unit. This is how we can recognize the beginnings of 'urban sprawl' within our own community, even though the individual components in such a change in land use over time are complex. Kosslyn *et al.* (1992) argued that, because of the fundamental difference in how they are used, the *where*, *what* and *when* systems are qualitatively distinct from each other, and as such, they are encoded separately within the brain. Clearly, however, *what* and *where* knowledge cannot be completely separate and independent (Kosslyn and Koenig 1992, Landau and Jackendoff 1993) as it is the interaction between these systems that allows us to function in the world.

The interdependency of the cognitive *what* and *where* representational systems was described by Marr (1982) specifically with respect to vision. Marr was interested in building a computer model of vision, and viewed vision as an information processing task. He asserted that processing of visual information must begin with the perceived *image* of the real world. Note that image is meant here in the machine vision sense, i.e. a location-based array of colour values, and not in the sense of 'mental imagery'. Directly observed phenomena must first be selected and abstracted into key characteristics of the image (the 'primal sketch'). These characteristics are interpreted using pre-existing knowledge, such as that which is stored in a schema; objects are eventually associated with locations and groups of locations in the image. Objects, as concepts, are thus always higher-order information. This information can then be placed (or confirmed) within the *what* system and it also becomes knowledge that can be used to interpret subsequent images.

2.4. *Summary*

This section has highlighted the features of cognition that we believe are particularly pertinent to geographic database representation. First, there is significant evidence that our knowledge of objects, our knowledge of scenes, and our knowledge of motion and process comprise three separate cognitive subsystems within the mind. Furthermore, there are significant differences in the way that spatial information about each is encoded. The *what* system of knowledge operates by recognition, comparing observed evidence with a gradually accumulating store of known objects. The *where* system operates primarily by direct perception of scenes within the environment, picking up invariants from the rich flow of sensory information. The *when* system operates through the detection of change over time in both stored object and place knowledge, as well as sensory information.

In addition, categorization plays a key role in geographic cognition. Taxonomic and partonomic hierarchies facilitate the recognition of objects from sensory perception of the environment by providing schema for generic types of objects. These schema are used within a cycle of knowledge acquisition in which sensory observations (Marr's image) are used to identify objects in the environment (higher order conceptual entities) through abstraction, which may then be used to modify and refine the original schema (figure 1).

3. Cognition and geographic database representation

While the integration of semantics into database representation has been an ongoing project since the advent of the relational data model (Codd 1970, Chen 1976, Hull and King 1987, Peckham and Maryanski 1988, Borgida 1991, Booch 1994), there have been a number of recent studies that focus specifically on the integration of cognitive principles into geographic database representation. For example, Usery (1993) notes that category theory can provide a useful model for development of a semantic 'feature-based' GIS. Of particular relevance to this paper is the idea that database representation can be based on the decomposition of

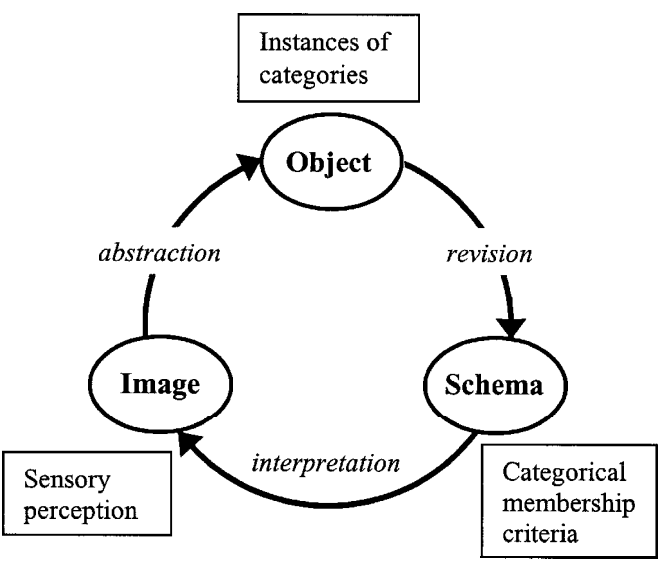

Figure 1. The cycle of cognitive geographic knowledge acquisition and refinement.

geographic information into space, time, and theme properties (Sinton 1978, Nyerges 1991), analogous to the where, when, and what components of cognitive representation described in section §2.3. Of equal importance is the idea that since cognitive representation (and acquisition) of geographic knowledge is divided between sensory information (image) and derived knowledge (schema), so too can geographic database representation be divided between spatiotemporal data and higher level geographic knowledge derived from that data (Rennison and Strausfeld 1995, Freksa and Barkowsky 1996).

Building on these themes, Peuquet (1994) describes the Triad framework for geographic representation which decomposes geographic entities into interrelated location-based (where), time-based (when), and object-based (what) information. This approach seeks to capture different perspectives on a given geographic phenomenon, but does not distinguish between observational data and derived knowledge. Consequently, the Triad framework cannot represent categorization and other important features of cognitive geographic representation and knowledge acquisition as described in figure 1. Raper and Livingstone (1995) take a somewhat similar approach to develop a spatiotemporal geomorphologic database model. Their data model represents a geomorphic phenomenon as a semantic decomposition into attributes derived from the basic-level categories form, process, and material. These attributes are referenced to a four-dimensional spatiotemporal framework with axes of location (in three dimensions) and time.

Other researchers have sought to improve geographic database representation using object oriented data modelling. The object oriented abstraction techniques generalization and aggregation are analogous (but clearly not identical) to taxonomic and partonomic cognitive categorization hierarchies, respectively. In geographic database research, these techniques have primarily been applied to the representation of complex geometry of geographic phenomena (e.g. Egenhofer and Frank 1989, Milne *et al.* 1993, Worboys 1994, Tang *et al.* 1996). For a more detailed review of object-oriented conceptual modelling the reader is referred to Rumbaugh *et al.* (1991) and Booch (1994).

4. Description of the pyramid framework

4.1. *Overview*

The pyramid framework is founded on the principles and techniques described in the above mentioned body of research, including object oriented modelling and other approaches, in order to develop a rigorous, yet comprehensive, framework for geographic database representation. The conceptual framework proposed here is intended to be non-domain specific (i.e. inclusive of a variety of geographic phenomena) and is based on reducing knowledge about a geographic phenomenon into perspectives based on the cognitive separation of what, when, and where knowledge. The pyramid framework incorporates key cognitive structures including categorization, part-whole relations, and behavioural rules, in order to create a high-level semantic representation of an interrelated web, or system, of geographic phenomena. We also employ theories of geographic knowledge acquisition, from sensory perception to cognitive conceptualization, in order to separate the representation of geographic data from the representation of geographic knowledge that is derived from that data.

Thus far we have used terms such as data, information, and knowledge relatively loosely, but it is helpful at this point to define them more formally. *Data* are

observational measurements that have been recorded in some way, whereas *information* is data that is generalized, ordered and contextualized in ways that give them meaning. Information thus is selective toward data, separating the important from the relatively unimportant. *Knowledge* is a cumulative understanding of information, i.e. an overall representative structure and a set of generalized rules of the relevant phenomenon. Here, we use the term knowledge for all levels of derived description, including information, in order to distinguish it from observational data.

Based on the semantic difference between data and knowledge, the pyramid framework is composed of two distinct, yet interrelated, components: the Data component and the Knowledge component (figure 2). The Data component represents uninterpreted observational data, analogous to the sensory *image* as described by Marr (1982). These observational data are represented by three distinct data 'perspectives': location, time, and theme.

The Knowledge component represents derived knowledge that is analogous to the recognition of objects embedded within the visual image (i.e. the observational data) and the classification of those objects into categories. It is in the Knowledge component that the three perspectives of observational data come together to form a semantic object (conceptual entity) that is of interest to the user, hence the graphic pyramid structure and the name of the framework. Here, we identify two types of classification that mirror the principles of cognitive categorization: taxonomy (superordinate-subordinate relationships) and partonomy (part-whole relations). However, cognitive research demonstrates that this process of knowledge derivation from observational data is not unidirectional; prior knowledge is also often used to interpret the observational data. Therefore, the Knowledge component also represents a rule-base for identifying objects and their categorical membership from the observational data, a function that is analogous to the *schema* as described by Thorndyke (1984).

As such, the Data and Knowledge components represent levels of abstraction and a 'cycle' of knowledge development analogous to how humans use sensory

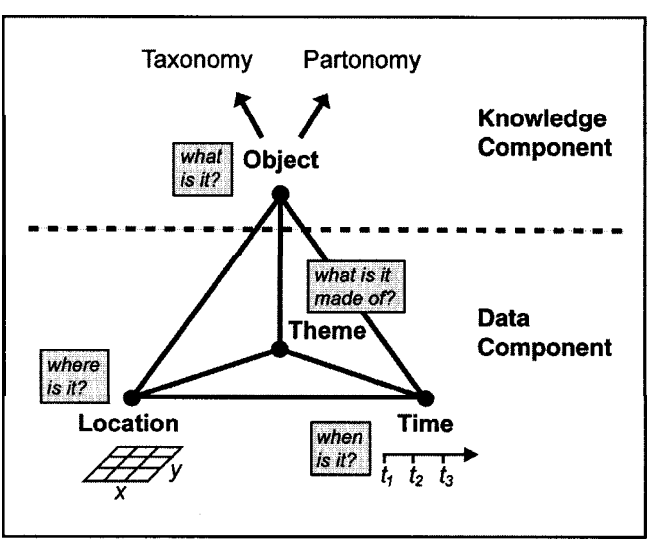

Figure 2. Overview of the pyramid framework: data component and knowledge component.

perception to learn about the world, as is illustrated in figure 1. Observational data about the world is stored in the Data component, then classified and interpreted to derive meaningful information that is stored in the Knowledge component. This knowledge may then be used to further interpret the observational data. The pyramid framework can therefore be thought of as a knowledge development *process* in addition to a static representational structure.

4.2. *The data component*

Observed geographic data are formally represented within the Data component by the *location*, *time*, and *theme* perspectives. The location and time perspectives can be conceptualized using a geometric 'container' metaphor, in which space and time can be thought of as an absolute and enclosed 'space' within which geographic phenomena exist and processes and events take place. This is relatively intuitive—real world geographic entities extend across space and over time. The general notion of the Data component can be understood by using the familiar 3-D 'space-time cube' (figure 3) with location extending across the x and y axes and time extending along the t axis. This cube defines a set of spatiotemporal subspaces, or voxels, each with an assigned location and time.

A theme is defined here as a property that can be sensed, measured, and assigned a qualitative or quantitative value. A theme behaves as a spatiotemporal field of measurement—it is measurable throughout all space and time in the model universe, even if the measurement values are null or zero at some locations and times. Temperature is a good example of a theme, as it used here, because its measurement yields a numeric value for any point on the earth's surface.

Although less intuitive than location and time, a 'theme space' can be conceptualized in a way similar to the space-time cube, such as in a multidimensional statistical space or in a *feature space*, as is commonly referred to in the remote sensing literature, where each sensed band of electromagnetic energy is defined as one dimension in the feature space (Landgrebe, 1998). The integration of the theme space with the space-time cube creates a multidimensional hypercube, such as in the spatiotemporal observational data framework described by Raper and Livingstone

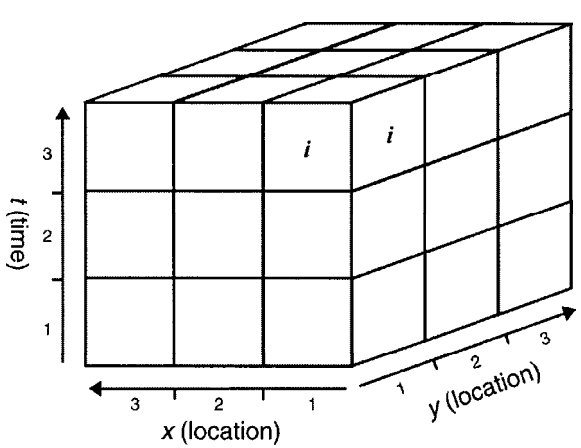

Figure 3. The space-time cube with two axes that describe location (x and y) and one axis that describes a time line (t). Voxel I defines the intersection of a particular location and time: $\{x, y, t\} = \{1, 1, 3\}$.

(1995). This multidimensional space with all themes included is termed the *data space*. Each multidimensional 'voxel' within the data space is defined by the intersection of its location, time, and multiple theme values.

4.3. *The knowledge component*

Whereas the Data component of the pyramid framework concerns 'raw' observational data with stored values for locations, times, and themes, the Knowledge component concerns the representation of derived geographical objects, their classification, and their inter-relationships. In this sense, the Knowledge component is representative of the conceptual entities that are recognized through the cognitive process of abstraction of sensory input as well as the cognitive categorical knowledge used in the abstraction process. The term *object* is used here to refer to a geographic conceptual entity that has a unique and cohesive identity, and is related to a specific combination of observational data stored in the location, time, and theme perspectives. An object here is analogous to an instance in object oriented modelling. We use the term *class* to refer to a type of object or a category of object, identical to the usage of the term class in object oriented modelling. In the pyramid framework, as in object oriented modelling, every object is an instance of, or a member of, a particular class.

For a specific object, the location and time perspectives of the Data component describe where, and at what times, the object was interpreted as existing within the data space throughout its lifetime. The theme perspective describes the range of thematic measurements that 'constitute' the object, or, in a sense, the composition of the object. The location, time, and theme 'history' of the object can therefore be retrieved by accessing the data space voxels to which it is referenced. However, only certain spatiotemporal properties of an object may be derived directly from the object's reference to the data space. An object may also have many semantic properties that are independent of its data space instantiation.

In order to represent semantic properties associated with an object, and to link the object with its observed properties stored within the data space, an object is represented within the pyramid framework through the use of an *object template*. To allow maximum flexibility in the types of information that can be recorded, as well as for the sake of uniformity throughout the Knowledge component, we use a frame type of structure. Originally proposed by Minsky (1975), a *frame* is a conceptual structure for representing a situation, a specific instance of an object, or for defining an object class. The unique aspect of a frame structure is that it can contain different kinds of information beyond simple attribute values and explicit relationships with other objects, including default characteristics, important exceptions, and behavioural rules. Frames can also contain procedures to retrieve attribute values on the fly or that may act as demons to simulate sensors or monitors that act on newly acquired observational data.

Figure 4 shows an example of a 'blank' object template where the attributes are defined as follows:

- Object ID—a unique identifier for the object
- Birth—when the object comes into existence
- Death—when the object ends existence
- Lifespan—duration of the object's existence
- Minimum size—smallest size of the object during its lifespan

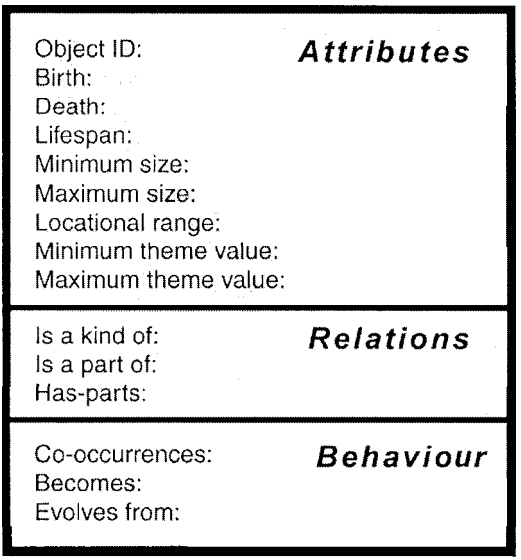

Figure 4. A 'blank' object template used to formally represent an object in the pyramid framework.

- Maximum size—largest size of the object during its lifespan
- Locational range—bounding polygon where the object exists during its lifespan
- Minimum theme value—smallest theme value of the object during its lifespan
- Maximum theme value—largest theme value of the object during its lifespan

As can be seen, the template for an object stores only those values that are critical in the history of the object, as well as those attributes and attribute values that describe an object as a whole. The function of the object template is not to exhaustively describe the object. Rather, the object template serves two interrelated functions: first, to provide a guide for finding the location, time, and theme history of the object within the Data component, when desired, and, second, to represent enough information to define the object's unique identity, including its relationships with other objects and its behavioural characteristics. These relationships and how they function are defined below in relation to the storage of classes.

Classes are represented using a *class template*. This is a frame structure similar to an object template, but the attributes of a class record ranges of values that define the *criteria* for membership in the given class rather than values that record the observed character of an object. In this way, class templates play a role analogous to that of schema in the cognitive process of object recognition. The ranges of attribute values described in the class template serve as a rule-base for the interpretation of the data stored in the Data component. These are used in discovering specific instances of that class, i.e. objects. For example, while an object has the attribute 'lifespan', a class has the attributes 'maximum lifespan' and 'minimum lifespan'. All objects that are members of that class are noted to have an actual lifespan within that maximum and minimum range to be a class member. Additional information may be included to note whether the stored attribute value ranges are required or simply expected for members of that particular class.

In addition to the location, time and theme information that describes the

attributes of a class, the class template contains slots for attributes that are not directly related to the Data component but concern the semantics of the object category. This includes information that describes the behaviour of an entire class. These include the behaviour *becomes*—what other type of object objects of this class can change into, thereby undergoing a change of identity. An example of this would be that the class 'tropical storm' *becomes* 'hurricane'. Of course, not all tropical storms become hurricanes, but this establishes a potential behavioural path for that class. The *co-occurrences* relation defines what other types of objects are typically spatially and/or temporally coincident with objects in this class.

The class template also includes information on the interrelationships of the various classes that mimics cognitive categorization structures; each class may have taxonomic *is-a-kind-of* and *has-kinds* relationships. This allows each class to be categorized into a higher-level superclass, which in-turn may be categorized into another higher-level class, and so on, forming a generalization, or taxonomic, hierarchy. In addition to generalization relationships, classes can have aggregation relationships. Some real-world geographic entities are part of other, larger and/or longer-lasting, geographic entities. To capture this, classes can have partonomic *is-a-part-of* and *has-parts* relationships with other classes. This relationship property indicates that the objects that are members of that class can have aggregation relationships with objects in other classes.

Generalization and aggregation relationships, and behavioural properties, exist not only for classes but also for specific objects and they are therefore represented in the object template (figure 4). In the object template, the *is-a-part-of* and *has-parts* relationships indicate composition between specific objects (as opposed to the general and potential aggregation relationship indicated for classes). An object also has an *is-a-kind-of* relationship with the class of which it is a member (by definition an object cannot have a *has-kinds* relationship with any other object). The behavioural attributes listed in the object template, such as *becomes*, *evolves-from*, *and co-occurrences*, also refer to other specific objects.

Note that some objects can be defined entirely from other objects and not defined directly from the data space. Objects that are derived directly from their observation in the data space are termed *atomic*; those objects that are composed solely of other objects are termed *composite*. Note that the terms 'atomic' and 'composite' are not used in the conventional sense to refer to primitive and complex data types, such as integers and arrays, respectively. Rather, the terms refer to an object as the term 'object' is used here—a conceptual entity derived from the interpretation of observational data. Like the generalization hierarchy, the aggregation relationship can be extended from the simple atomic-composite object relationship to a hierarchy of composite objects. Many composite objects may have *part-of* relationships with another composite object, forming an aggregation hierarchy, or partonomy.

So far we have described two separate and parallel hierarchies built upon generalization and aggregation. These two hierarchies are interrelated—each class in the generalization hierarchy may have one or more superclasses and/or subclasses, and also may have a set of objects (instances). Each object, whether atomic or composite, is a member of one or more particular classes in the generalization hierarchy. Class behaviour and relationship attributes are passed down to subclasses and member objects through inheritance. This approach ensures that information that is common to an entire class does not have to be redundantly stored in each object template for every member object.

5. Discussion of the pyramid framework

In the pyramid framework, we have described two interrelated components: first, the Data component that encompasses observed location, time, and theme data and, second, the Knowledge component. The Data component provides the observational data with which objects are 'constructed' and is always maintained in its original, uninterpreted state. The Knowledge component stores categorical criteria that allows for the recognition of conceptual objects from the Data component as well as properties of the actual objects themselves. This model represents real-world geographic entities in a manner that is better informed by human cognition than current conventional geographic data modelling approaches. This is also analogous to how humans modify knowledge based upon comparison of new observational information to recollection of previous observations.

It is up to the domain expert to determine the classification criteria in the Knowledge component by either a 'top-down' or 'bottom-up' approach, or a combination of both. In the top-down approach, the domain expert creates classes with predetermined membership criteria before defining the member objects within each class. In this classification approach, a priori expert knowledge of the modelled domain is used to create predefined object classes.

The bottom-up approach is more exploratory and is appropriate when the nature of the actual objects that populate the modelled universe is unknown. Here, some statistical clustering technique or inductive classification method can be used in order to develop classes and objects from the data space. Instead of integrating expert knowledge into the pyramid framework, this approach seeks to facilitate knowledge discovery by minimizing a priori assumptions about the nature of the data and the conceptual entities they describe and maximizing the 'objective' recognition of patterns embedded in the data.

The templates as given in this discussion are not meant to be exhaustive, but rather illustrative and to provide a general basis for geographic knowledge representation within a digital database. There are many classes where the storage of additional information would be advantageous, or even essential. For example, if the recurrence interval of a type of object is particularly important in its recognition, as is the case in classifying flood events, than an attribute describing that recurrence may be included to define that class.

These domain-specified location-, time-, or theme-specific object attributes can be accessed in the pyramid framework using a query process that evaluates the *dimensional dominance* of the desired set of attributes. The dimensional dominance is defined as the most constrained dimension (location, time, or theme) of the query (Langran 1993, Peuquet and Qian 1996). For example, if one inquires where an object is located at a specific time, then time is the dominant dimension of the query. Whereas if one inquires where an object has a specific theme value, then theme is the dominant dimension.

Slightly more complex dimensionally constrained queries can reveal other location-, time-, and theme-dependent information such as:

- the size of an object at a specific time
- the duration that an object exists (or remains) at a specific location
- the rate of movement of an object over a particular duration

- how often an object recurs at a particular location
- the range in theme values of an object over a particular duration

While some of these attributes may be stored in the object template for the purpose of object recognition, it would be inefficient to explicitly store the result of every dimensionally constrained query. Humans do not store everything that they know about any given object class, but rather store the general rules that allow the required specific information to be retrieved and organized appropriately when needed. They also distill-out the appropriate components of a larger knowledge structure and reorganize it to suit the particular situation at-hand (Portugali 1996). This makes sense from a database storage efficiency perspective, as well.

The Knowledge component of the pyramid framework is similarly intended as a flexible structure that can be manipulated in order to accommodate changing views of the modelled domain. The creation of the Knowledge component can be described as an iterative procedure: the generalization and aggregation hierarchies may be constructed, analysed, subsequently revised, re-analysed, and so on, as well as abstracted for specific purposes. This iterative approach supports data exploration and knowledge development because it allows the modeller to create a variety of structural representations from the same underlying observational data stored in the Data component.

6. An example: the representation of a storm system

We represent a particular type of storm system, the mesoscale convective complex (MCC) in order to demonstrate the representational power of the pyramid framework. MCCs are regional scale storm systems (as opposed to larger, synoptic scale phenomena) that are composed of many individual storm cells that can act in concert to create particularly violent storms (Maddox *et al.* 1986). These typically 'egg' shaped storms often occur at night in the Midwestern US during the summer or early fall. They can produce copious amounts of rain and are often associated with flooding and other climate hazards. The representation of MCCs and their derivation from observational data are particularly appropriate for this conceptual modelling task because their identification is based on a set of specific parameters derived from analysis of remote sensing data, analogous to the extraction of objects from a data space described here.

The first step in developing the conceptual representation of an MCC is to define the Data component. For this purpose, consider a data space with four dimensions. The two location dimensions define the Midwestern region of the US, regularly partitioned into locational cells 1000 m on a side. The time dimension defines the year-long period beginning 1 January, 1998 and ending 31 December, 1998, regularly partitioned into hour long intervals. For simplicity, we define only one theme, cloud top temperature. Cloud top temperature is derived from thermal infrared (TIR) satellite imagery and indicates the altitude of the cloud top (the higher the altitude, the lower the temperature), which in turn can indicate the stage of development or severity of a storm (Carleton 1991).

Development of the Knowledge component in this particular case is a top-down procedure, as opposed to bottom-up, because we are defining MCCs using pre-defined expert knowledge. As such, we begin by defining the MCC class template and its attributes which then later act as criteria for identifying actual MCC objects within the data space. The MCC class is described by its class template (figure 5)

Class ID: *MCC* Minimum lifespan: *6 (hours)* Maximum lifespan: *none* Minumum size: *50,000 (sq km)* Maximum size *350,000 (sq km)* Minimum shape eccentricity: *0.7* Maximum shape eccentricity: *none* Minimum cloud top temp: *none* Maximum cloud top temp: *-52.0 (degrees C)*	**Attributes**
Is a kind of: *MCS* Is a part of: *large-scale front* Has parts: *none*	**Relations**
Co-occurences: *night* *low level jet stream* *Evolves from: Squall line*	**Behaviour**

Figure 5. The MCC class template that describes the attributes of the MCC class, including the criteria for class membership, relationships with other classes, and typical co-occurrences.

which shows the MCC class attributes and their respective values. The criteria listed here are adapted from Carleton (1991).

- a cloud top temperature that is less than − 52°C
- a contiguous area of between 50 000 and 350 000 square kilometres
- a duration longer than six hours
- a shape eccentricity (minor/major axis of the storm area) greater than or equal to 0.7

Note that while the minimum and maximum shape eccentricity attributes may not necessarily be included in every class template (as the attributes minimum and maximum lifespan are), they are included here because eccentricity is essential for the identification of MCC objects. The class template also records the generalization and aggregation relationships of the MCC class: an MCC *is-a-kind-of* larger category of storm, mesoscale convective system (MCS) (Chappell 1986), and may be *part-of* an object that represents a specific type of synoptic setting associated with a large-scale front (Maddox 1983). In addition, the class template lists other attributes of MCCs: they can evolve from squall line storms (Maddox *et al.* 1986), they typically occur at night, and they are often associated with a low level jet stream (Carleton 1991).

The criteria captured within the MCC class template acts as a rule-base for a query on the data space to identify individual MCC objects. The query procedure to identify MCC objects steps through each of the class attributes described in figure 5 to select the portion of the data space that satisfies the class membership criteria. The following query procedure (described simply in plain English, as opposed to a formal query language) serves as an example of how MCC objects may be

identified from the data space. Note that each operation described below is applied only to the objects that were selected in the previous selection operation.

- Theme value criteria:
  Select only those voxels with a cloud top temperature $< -52°C$
- Size criteria:
  Group all contiguous selected voxels into individual (temporary) potential MCC objects
  Calculate the area of each potential MCC object
  Select only those potential MCC objects with an area between 50 000 and 350 000 square kilometres
- Shape criteria:
  Calculate the major and minor axes for each potential MCC object
  Calculate the eccentricity (minor/major axis) of each potential MCC object
  Select only those potential MCC objects with an eccentricity $> = 0.7$

Querying based on the duration criteria can be complex. In order to calculate the duration of each potential MCC object, it is necessary to maintain the identity of each potential MCC object over time. If an object is stationary, location can be used to identify the same object over time; however, MCCs are not stationary but often move. For the purpose of this example, it is possible to maintain the identity of each potential MCC object through time by simply stating that if the centroid of a potential MCC object is located within 200 km of the centroid of another potential MCC object at an adjacent time interval (i.e. either the prior or following time interval), it is recognized as the same potential MCC object that has moved over time. This method assumes that MCCs do not move (planimetrically) at a rate that is greater than 200 km per hour and that two or more MCCs do not occur within 200 km of each other within the span of two consecutive time intervals (two hours). Research on MCC storm movement suggests that this is typically the case (Maddox *et al.* 1986). Using this approach to maintaining identity over time, the duration criteria may be satisfied by selecting all potential MCC objects that exist (maintain their identity) over a duration of greater than six hours.

Each of the potential MCC objects that meet all the criteria listed above (theme, size, shape, and duration) are then defined (instantiated) as actual MCC objects. All MCCs that occurred over the United States during 1998 are represented as individual MCC objects with associated location, time, and theme data captured as a portion of the data space. Each MCC object also has an object template that captures the important attributes of that particular MCC: its minimum and maximum cloud top temperature, its minimum and maximum size, its duration, etc. As an example, figure 6 shows the object template for a particular MCC object that occurred the early morning of 23 August, 1998. Other attributes included in the object template are the MCC object's *is-a-kind-of* relationship with the MCC class and its *is-a-part-of* relationship with a particular large-scale front object, as well as its relevant co-occurrences, such as the fact that it occurred during the night. Note the difference between the class template (figure 5) and this object template (figure 6). The two templates are similar, however the class template stores attribute values that are common to the entire class (including ranges of values for certain attributes) while the object template stores attribute values that describe an actual geographic entity.

This example is relatively simple and straightforward. MCCs are modelled here as atomic objects, using only one theme, and with only one higher level of generalization and aggregation. One can imagine, however, that these same principles can be

Attributes

Object ID: *MCC [16]*
Birth: *August 23, 1:00 am*
Death: *August 23, 9:00 am*
Lifespan: 8 *(hours)*
Minimum size: *50,000 (sq km)*
Maximum size: *129,000 (sq km)*
Locational range: bounding rectangle
Minimum cloud top temp: *-68.0 (degrees C)*
Maximum *cloud top temp: -52.0 (degrees C)*

Relations

Is a kind of: *MCC*
Is a part of: *large-scale front [11]*
Has-parts: *none*

Behaviour

Co-occurences: *night*
low level jet stream [5]
Evolves from: none

Figure 6. The object template for a particular MCC that occurred during the early morning of 23 August, 1998 (numbers in brackets indicate the identification of a particular object).

used to develop a much more complex storm system representation if other data, such as precipitation rate and air pressure, and a richer hierarchy of storm types, such as single-cell storms and squall lines, were integrated into the representation. Further, note that the structure of the Knowledge Component is relatively flexible. This representation of MCCs in the Midwestern US during 1998 may be refined from both additional observational data and from declarative knowledge generated by domain experts. By altering the attributes (criteria) of the MCC class, new sets of MCC objects can be generated, providing a variety of different structural interpretations of the climate data for different application contexts.

7. Conclusion

Advances in geographic database representation are dependent upon a conceptual level approach that incorporates models of human cognition of geographic phenomena. In addition, the conventional view of database modelling as the development of a static structure that exists for the lifetime of the database inherently limits exploration by providing only one view of the data. As a representational device, the pyramid framework brings geographic database representation more in line with human geographic cognition and also allows for the development of flexible and modifiable geographic databases that can facilitate knowledge acquisition (i.e. learning).

The pyramid framework as described in this paper is conceptual in nature. It is intended as a framework that can be used for departing from traditional GIS representation techniques that have their roots in cartography, toward an approach to representation that has the capability of utilizing the full potential of computers as a representational medium, with much closer links to how humans represent and learn about geographic space. This represents a dramatic departure from the long-standard assumption that representation of geographic phenomena begins with

points, lines, polygons and pixels. Granted, the framework proposed here is more complex than previous approaches, but this is necessary in order to better reflect the complexity of human spatial cognition and learning and to thereby have a much more powerful and flexible approach for digital database representation.

The pyramid framework is not intended to be comprehensive in its treatment of cognitive geographic representation. Besides the futility in trying to imitate rather than selectively using aspects that are particularly advantageous, there is flexibility for adding aspects that may be particularly useful or appropriate within a particular application context. There is also opportunity to integrate more advanced knowledge representation schemes once they are developed. For example, there is ample evidence that cognitive categorization differs significantly from the set-based model of categories implicit in the object-oriented techniques used to represent categorical knowledge in the pyramid framework as presented (Rosch 1978, Lakoff 1987). Perhaps this particular issue may be addressed in future versions of the framework by drawing on recent research on the representation of 'fuzzy' geographic entities (Burrough and Frank, 1996).

There are other theoretical issues that warrant additional research concerning the interaction between the Data and Knowledge components of the framework. For instance, it is not specified in the framework exactly how objects and classes in the Knowledge component may be generated from observational data in the Data component. Further efforts in this area need to address a variety of issues that include the application of inductive and deductive logics, the use of statistical and rule-based knowledge discovery techniques, and the role of graphic representation at the human-machine interface.

Acknowledgments

This research was supported in part by US Environmental Protection Agency grant R825195-01-0. The authors would like to thank Agnar Renolen, Monica Wachowicz, and Alan MacEachren for their contributions to the conceptual development of the paper. Thanks also to the two anonymous reviewers for their helpful comments.

References

Anderson, R. A., 1987, Inferior Parietal Lobule Function in Spatial Perception and Visuomotor Integration. In *Handbook of Physiology*, edited by F. Plum and V. B. Mountcastle (Rockville, MD: American Physiological Society), pp. 483–518.

Booch, G., 1994, *Object oriented Analysis and Design* (New York: Benjamin/Cummings Publishing Co.).

Borgida, A., 1991, Knowledge representation, semantic modelling: similarities and differences. In *Entity-Relationship Approach: The Core of Conceptual Modelling,* edited by H. Kangassollo (North-Holland: Elsevier Science Publishers), pp. 1–24.

Burrough, P. A., and Frank, A. U., 1995, Concepts and paradigms in spatial information: are current geographical information systems truly generic? *International Journal of Geographical Information Systems*, **9**(2), 101–116.

Burrough, P. A., and Frank, A. U. (eds), 1996, *Geographic Objects with Indeterminate Boundaries* (London: Taylor and Francis).

Carleton, A., 1991, *Satellite Remote Sensing in Climatology* (London: Bellhaven Press).

Chappell, C. F., 1986, Quasi-stationary convective events. In *Mesoscale Meteorology and Forecasting*, edited by P. S. Ray (Boston: American Meteorological Society), pp. 289–310.

Chen, P. P.-S., 1976, The entity-relationship model—toward a unified view of data. *ACM Transactions on Database Systems*, **1**(1), 9–36.

Codd, E. F., 1970, A relational model of data for large shared data banks. *Communications of the ACM*, **13**, 377–387.

COUCLELIS, H., 1993, People manipulate objects (but cultivate fields): beyond the raster-vector debate in GIS. In *International Conference GIS—From Space to Territory: Theories and Methods of Spatio-Temporal Reasoning*, edited by A. U. Frank, I. Campari, and U. Formentini (Berlin: Springer-Verlag), pp. 65–77.

DOWNS, R., and STEA, D. (eds), 1973, *Image and Environment* (Chicago: Aldine).

EASTMAN, J. R., 1985, Graphic organization and memory structures for map learning. *Cartographica*, **22**(1), 1–20.

EGENHOFER, M. J., and FRANK, A. U., 1989, Object oriented modelling in GIS: inheritance and propagation, *Auto-Carto 9 Proceedings*, 588–598.

FARAH, M. J., HAMMOND, K. M., LEVINE, D. N., and CALVANIO, R., 1988, Visual and spatial mental imagery: dissociable systems of representation. *Cognitive Psychology*, **20**, 439–462.

FRANK, A. U., 1992, Spatial concepts, geometric data models, and geometric data structures. *Computers & Geosciences*, **18**(4), 409–417.

FREKSA, C., and BARKOWSKY, T., 1996, On the relations between spatial concepts and geographic objects. In *Geographic Objects With Indeterminate Boundaries*, edited by P. A. Burrough and A. U. Frank (London: Taylor and Francis), pp. 109–121.

GOLLEDGE, R., 1993, Geographical perspectives on spatial cognition. In *Behaviour and Environment: Psychological and Geographical Approaches*, edited by R. Garling and R. G. Golledge (Amsterdam: Elsevier), pp. 16–46.

HERRING, J. R., 1992, TIGRIS: a data model for an object-oriented geographic information system. *Computers and Geosciences*, **18**(4), 443–452.

HULL, R., and KING, R., 1987, Semantic database modelling: survey, applications and research issues. *ACM Computer Surveys*, **19**(3), 201–260.

KOSSLYN, S., CHABRIS, C., MARSOLEK, C., and KOENIG, O., 1992, Categorical versus coordinate spatial relations: computational analyses and computer simulations. *Journal of Experimental Psychology: Human Perception and Performance*, **18**(2), 562–577.

KOSSLYN, S. M., and KOENIG, O., 1992, *Wet Mind: The New Cognitive Neuroscience* (New York: The Free Press).

LAKOFF, G., 1987, *Women, Fire, and Dangerous Things: What Categories Reveal about the Mind* (Chicago: University of Chicago Press).

LAKOFF, G., and JOHNSON, M., 1980, *Metaphors We Live By* (Chicago: University of Chicago Press).

LANDAU, B., and JACKENDOFF, R., 1993, 'What' and 'where' in spatial language and spatial cognition. *Behavioural and Brain Sciences*, **16**, 217–265.

LANDGREBE, D., 1998, Information extraction principles and methods for multispectral and hyperspectral image data. In *Information Processing for Remote Sensing*, edited by C. H. Chen (River Edge, NJ: World Scientific Publishing), pp. 1–30.

LANGRAN, G., 1993, Manipulation and analysis of temporal geographic information, *Canadian Conference on GIS* (Ottawa: The Canadian Institute of Geomatics), pp. 869–879.

LLOYD, R., PATTON, D., and CAMMACK, R., 1996, Basic-level categories. *Professional Geographer*, **48**(2), 181–194.

MADDOX, R. A., 1983, Large scale meteorological conditions associated with mid-latitude, mesoscale convective complexes. *Monthly Weather Review*, **111**, 1475–1493.

MADDOX, R. A., HOWARD, K. W., BARTELS, D. L., and RODGERS, D. M., 1986, Mesoscale convective complexes in the middle latitudes. In *Mesoscale Meteorology and Forecasting*, edited by P. S. Ray (Boston: American Meteorological Society), pp. 390–413.

MARK, D. M. (ed.), 1999, Geographic information science: critical issues in an emerging cross-disciplinary research domain. Report from the Workshop on Geographic Information Science and Geospatial Activities at NSF, Arlington, VA, 14–15 January.

MARK, D. M., and FRANK, A. U., 1996, Experiential and formal models of geographic space. *Environment and Planning B: Planning and Design*, **23**, 3–24.

MARR, D., 1982, *Vision* (New York: W. H. Freeman Co.).

MILNE, P., MILTON, S., and SMITH, J. L., 1993, Geographical object oriented databases—a case study. *International Journal of Geographical Information Systems*, **7**(1), 39–55.

MINSKY, M., 1975, A framework for representing knowledge. In *The Psychology of Computer Vision*, edited by P. H. Winston (New York: McGraw-Hill), pp. 211–277.

MURPHY, G., and WISNIEWSKI, ?., 1989, Categorizing objects in isolation and in scenes: what a superordinate is good for. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, **15**(4), 572–586.

NYERGES, T. L., 1991, Geographic information abstractions: conceptual clarity for geographic modelling. *Environment and Planning A*, **23**, 1483–1499.

PECKHAM, J. and MARYANSKI, F., 1988. Semantic data models. *ACM Computing Surveys*, **20**(3), 153–189.

PEUQUET, D. J., 1988, Representations of geographic space: toward a conceptual synthesis. *Annals of the Association of American Geographers*, **78**(3), 375–394.

PEUQUET, D. J., 1994, It's about time: a conceptual framework for the representation of temporal dynamics in geographic information systems. *Annals of the Association of American Geographers*, **84**(3), 441–461.

PEUQUET, D. J., and QIAN, L., 1996, An integrated database design for temporal GIS. In *Seventh International Symposium on Spatial Data Handling*, edited by M. J. Kraak (London: Taylor & Francis), pp. 2-1–2-11.

PORTUGALI, J. (ed.), 1996, *The Construction of Cognitive Maps* (Dordrecht: Kluwer Academic Publishers).

RAPER, J., and LIVINGSTONE, D., 1995, Development of a geomorphological spatial model using object oriented design. *International Journal of Geographical Information Systems*, **9**(4), 359–383.

RENNISON, E., and STRAUSFELD, L., 1995, The millenium project: constructing a dynamic 3D virtual environment for exploring geographically, temporally and categorically organized historical information. In *Spatial Information Theory: A Theoretical Basis for GIS*, edited by A. U. Frank and W. Kuhn (Berlin: Springer-Verlag), pp. 69–91.

ROSCH, E., 1978, Principles of categorization. In *Cognition and Categorization*, edited by E. Rosch and B. Lloyd (Hillsdale, N.J.: Lawrence Erlbaum Associates), pp. 27–77.

ROSCH, E., 1983, Prototype classification and logical classification: the two approaches. In *New Trends in Conceptual Representation: Challenges to Piaget's Theory?*, edited by E. K. Scholnick (Hillsdale, NJ: Lawrence Erlbaum Associates), pp. 73–86.

ROSCH, E., and MERVIS, C. B., 1975, Family resemblances: studies in the internal structure of categories. *Cognitive Psychology*, **7**, 573–605.

ROSCH, E. H., 1973, On the internal structure of perceptual and semantic categories. In *Cognitive Development and the Acquisition of Knowledge*, edited by T. E. Moore (New York: Academic Press), pp. 111–297.

RUMBAUGH, J., BLAHA, M., PREMERLANI, W., EDDY, F., and LORENSEN, W., 1991, *Object oriented Modelling and Design* (Englewood Cliffs, NJ: Prentice-Hall).

SERGENT, J., 1991, Judgements of relative position and distance on representations of spatial relations. *Journal of Experimental Psychology: Human Perception and Performance*, **17**, 762–780.

SINTON, D. F., 1978. The inherent structure of information as a constraint to analysis: mapped thematic data as a case study. Harvard University, Cambridge.

STEVENS, A., and COUPE, P., 1978, Distortions in judged spatial relations. *Cognitive Psychology*, **10**, 422–437.

TANG, A. Y., ADAMS, T. M., and USERY, E. L., 1996, A spatial data model design for feature-based geographical information systems. *International Journal of Geographical Information Systems*, **10**(5), 643–659.

THORNDYKE, P., 1984, Applications of schema theory in cognitive research. In *Tutorials in Learning and Memory: Essays in Honor of Gordon Bower*, edited by J. Anderson and S. Kosslyn (San Francisco: W. H. Freeman), pp. 167–192.

TVERSKY, B., and HEMENWAY, K., 1983, Categories of environmental scenes. *Cognitive Psychology*, **15**, 121–149.

TVERSKY, B., and HEMENWAY, K., 1984, Objects, parts, and categories. *Journal of Experimental Psychology: General*, **113**(2), 169–193.

UCGIS, 1996, Research priorities for geographic information science. *Cartography and Geographic Information Systems*, **23**(3).

UNGERLEIDER, L. G., and MISHKIN, M., 1982, Two cortical visual systems. *In Analysis of Visual Behaviour*, edited by D. J. Ingle, M. A. Goodale, and R. J. W. Mansfield (Cambridge, MA: MIT Press), pp. 549–586.

USERY, L., 1993, Category theory and the structure of features in geographic information systems. *Cartography and Geographic Information Systems*, **20**(1), 5–12.

WORBOYS, M. F., 1994, Object oriented aproaches to geo-referenced information. *International Journal of Geographic Information Systems*, **8**(4), 385–389.