On the Classification of Categories of Spatial Features

Yuxiao Li and George Benwell

Spatial Information Research Centre
University of Otago, Dunedin, New Zealand
Phone: +64 3 479-7391 Fax: +64 3 479-8311
Email: (yli, gbenwell)@infoscience.otago.ac.nz

Presented at SIRC 2002 – The 14th Annual Colloquium of the Spatial Information Research Centre
University of Otago, Dunedin, New Zealand
December 3-5th 2002

ABSTRACT

This paper discusses some issues related to the design of spatial information systems (SI systems) with consideration of real world semantics. Specifically, it concentrates on the key concept of categories for spatial features and internal structure of categories termed as taxonomy and partonomy. Here, spatial features refer mainly to geographic entities or phenomena at a large scale. The theoretical foundations of the issues are explored, including the classic philosophical and psychological views and the relatively new thoughts from experimental psychology on cognition. The main formal structures for external representation of categories and their structure, which are employed to represent categories of spatial features and their taxonomy and partonomy, are briefly discussed also. It is impossible to obtain a complete set of definitions of categories of spatial features relevant to the whole spatial information science (SI science) domain. Implementation of one single formal structure that has the potential to hold the set of category definitions and maintain their intricate relationships is too demanding. When the formation and application of categories and their internal structure can be put in the context of human cognition structure, an alternative thought may come into view. That is an operative approach, which begins from the basic level categories in an application area and deals with categories and their relationships adaptively. For facilitating the realisation of such an approach, the very top-level classification schema on categories of spatial features is needed and hence it is presented here. The classification is based on investigation of the nature of categories of spatial features, in terms of what they refer to in reality and what kinds of external forms of definitions and representations.

***Keywords and phrases*:** spatial features, categories, taxonomy, partonomy, relationships, classification of categories

1.0 INTRODUCTION

In this paper, the classification of categories of spatial features will be explored. It serves as an integrated component for design and implementation of a more intelligent spatial information system (SI system) with consideration of real world semantics. In some papers, this kind of information system, inclusive of real world semantics, is usually termed as an ontology-aware or ontology-driven information system.

Ontology is an emerging research area not only in information systems but science in general, such as the work of (Gruber, 1995; Guarino, 1995, 1998). This is also true in spatial information science (SI science) (Frank, 2001; Frank, 2001a; Frank, to appear; Smith & Mark, 1998). The terms frequently appear in studies for designing a new generation SI systems and for interoperability of conventional SI systems (Fonseca, Egenhofer, & Davis, 2000; Fonseca, Egenhofer, Davis, & Camara, 2002; Fonseca, Egenhofer, Agouris, & Camara, 2002; Rodriguez & Egenhofer, forthcoming in 2003). Ontology is usually defined or understood to be "the formal definitions of concepts and their relations in a domain". In the context of SI systems, ontology is usually interpreted as "categories of spatial features and their relations". Ontology-aware or ontology-driven SI systems refer to those existing or newly designed SI systems that have been given a knowledge-based component with explicit specifications of categories of spatial features and their relations.

In ontology-aware or ontology-driven information systems, the “formal definition or specification of concepts and relations” involved can be divided into two basic groups. One is directly related to the concepts and their relationships in the application domain itself, whether there are digital models of them or not. The other is related to the set of digital entities with structure as “concepts with relations”. They are the implementations of different abstraction levels developed for representing the real world “concepts and their relations”. The two sets of “concepts” and “relations”, as well as the relations between the two sets are both fundamentally needed and inseparable. To make a real ontology-driven system, there are at least two issues that will be of immediate concern. On the one hand, the concern is how to fit the digital implementations in general, which are based on logic and mathematical structures, into the big picture of human cognitive structure and give them the proper positions. There is a fundamental difference between a concept and its digital representation (Gruber, 1995). On the other hand, the issue is the representation of concepts and their relations involved in the application domain. In the context of SI system research, they are the spatial concepts and their relations relevant to SI science. Here a large part of the concept domain is composed of categories of spatial features and their relations. This is the focus of this paper. The structure of categories of spatial features will be, at least partially, explored from philosophical and cognitive points of view.

If a fundamental redesign of a more intelligent SI system is to be achieved it will be important to consider the categorisation of real world spatial features. It will be stated (with support from the literature) that a global taxonomy for all features is at present only a dream. It would be universally a category of the entire world, with all its complexities and is therefore impossible. Our belief is that a taxonomy for an intelligent SI system can be pragmatically dealt with and supported using the theory of intelligent agent societies. This society will demand the use of a robust framework. One way to proceed is first to develop a domain specific and detailed categorization of features. This is difficult and possibly arduous but at least, tractable. Overlaying this, a system wide ontology will include a more universal (but nonetheless incomplete) taxonomy supported by the agent society. The universal level will have to have an intelligent and flexible, adaptive bi-directional interaction with the more specific domain level taxonomy. A significant step towards addressing the difficulty is to implement these two responsive levels within the guiding adaptive management structure of an agent society. This consideration leads to a proposed top-level classification of categories of spatial features as the first step in an adaptive and operative approach towards their digital representations.

The paper is structured as follows: A brief review of categories, and structure of categories from studies in philosophy and cognitive psychology is given in Section 2. Section 3 discusses some of the representative formal structures for representing categories and their relations. The representation of categories, taxonomy and partonomy should be within a cognitive structure. Categories of spatial features and their relations should be put in the context of concepts and relations in SI science in general. These two levels of context, the principles for structuring categories, and reasoning in general are briefly discussed in Section 4. Given that categories and their structure are not static, and that the representation form of the internal structure of categories should therefore not be fixed, an alternative adaptive and operative approach for representation of categories and structure of categories as taxonomy and partonomy is considered. A classification schema of categories of spatial features serving for this adaptive and operative perspective is given in Section 5. Section 6 concludes the paper.

2.0 THEORIES ON REPRESENTATION OF CATEGORIES

The study in cognitive science of representation and reasoning has largely developed through three stages. In the classic Aristotelian ideals, category definition followed a *set-theoretical approach*. Categories have clear boundaries. They are defined in terms of sets of attributes. These attributes are shared by all members belonging to a category. Members of a category are equally good as examples of category formation (Rosch, 1973). Reasoning is the kind of syllogism and other forms of logic (Sowa, 1995, 2000).

Another major classic school of thought on category definition is the kind of *central tendency*. This is influenced in some way by field theory in modern physics. Categories are defined as clusters of attributes. The theory may look simple. However, the implementation of such a cognitive model of central tendency or cluster model, is very demanding as the whole set of perceived attributes is hardly known (Rosch & Mervis, 1975). While psychological studies show, although some qualitative features might be perceived by one single sensor, perception and conception of most qualitative features are decided by complex patterns developed in the mind rather than a single sensor (Gleitman, 1995) (p.194-196).

Psychologist E. Rosch and associates were not satisfied with the two traditional views of category representation (Rosch, 1973, 1978). She considered first, members of a category are not judged as equally good examples. A robin is a better example of bird than an ostrich. Secondly, people tend to agree with each other on such judgments as well. She and other co-researchers, such as B. Merlin (Berlin, 1978) developed what is now termed, the *prototype theory* of category.

One of the main points in this prototype theory is, categories are not a non-structured collection, but rather a structured whole. The internal structure of categories is organised both vertically and horizontally. Vertically, there are different abstraction levels such as superordinate, basic levels and subordinate abstraction levels in the taxonomy. For example, furniture, table, and kitchen table respectively. The basic level is the most informative level and matches nicely with natural discontinuity. Horizontally, categories are defined in terms of prototypes, which can be understood as those good examples or their generalisation.

Another critical point is that the category structure is not static, but adapted with experience. A category system of a child and that of an adult are different. In both the adult's and child's category systems, the basic level categories are the most informative and crucial for the adaptation of taxonomy. It is from the basic level categories that superordinate level categories are generalised and abstracted. Subordinate level categories are formed as a specialisation with consideration of the specific properties of categories. The differences in categories and taxonomy with external representations in languages can be found across individuals and cultures (Berlin, 1978; Gleitman, 1995; Rosch, 1973). In some cultures, there are only words for "black" and "white" as the categories of color (Rosch, 1973). Their "black" is different from our perceptions of the same word.

In Rosch's experiments, the proof for internal structure of categories is based on the convergences of a set of measures, including "the attributes in common, motor movements in common, objective similarity in shape and identifiability of averaged shapes" (Rosch, 1978). For example, for the category of chair, the attributes listed by the subjects can be "the thing you sit on", "have legs", "made of wood". A later study (Tversky & Hemenway, 1984) has differentiated parts from the list of "the attributes in common" in the studies of Rosch. Measures in linguistic terms are not quite the same as their conceptions in mind. Attribute terms cannot replace the exact kinds of properties and the changes of magnitudes that can be perceived by us. They themselves are categories on other aspects or abstraction levels. Rosch herself has strongly emphasised that although prototypicality can be reliably rated by her methods, it does not mean that prototype effect should be understood as implying a representation model or a reasoning mechanism in the mind directly. As she put it,

> "The fact that prototypicality is reliably rated and is correlated with category structure does not have clear implications for particular processing models nor for a theory of cognitive representations of categories. What is very clear from the extent research is that the prototypicality of items within a category can be shown to affect virtually all of the major dependent variables used as measures in psychological research" (Rosch, p38, 1978).

3.0 FORMAL STRUCTURES FOR REPRESENTING CATEGORIES FOR SPATIAL FEATURES

3.1. List of concepts

In early approaches categories are given as a list of definitions for spatial features. The structure of categories at best is a lexical one, i.e., sorting the categories in alphabetic order. The first example is from Spatial Data Transfer Standard (SDTS) (Fegeas, Cascio, & Lazar, 1992). In part 2 of the SDTS, categories (sometimes called as entity types or classes) of spatial features are defined in terms of properties. The list of categories is a working and expandable one. There are in total 2,600 definitions considered, 200 defined entity types, 244 defined attributes, and more than 1,200 included terms (Fegeas et al., 1992). The same line of thought also underlines the early efforts in the design of object-oriented SI systems. There, each type of entity or phenomenon is expected to be formally defined and implemented as a class. Objects are the individuations of a well-defined class, which are then assumed to be the models of reality (Raper & Livingstone, 1995; Tang, Adams, & Usery, 1996; Worboys, 1994). Real world existence is called instance in this paper. Note the difference between an individuation and an instance.

3.2. Tree, ordered-tree, and semi-lattice

One formal structure for the representation of spatial features with semantics has considered the internal structure of categories. It places particular emphasis on *hierarchy* and two kinds of hierarchies are considered.

One is the kind of *taxonomy* as in biology. The tree from *kingdoms* to *species* is organised as different abstraction levels. (This kind of hierarchy will be referred to later as an *is a* hierarchy). The other is the kind as the possible representation form of environmental knowledge theorised by Stevens and Coupe (Stevens & Coupe, 1978). Hierarchical structure as environmental knowledge representation is based on an analysis of empirical data obtained in studies on *cognitive maps*. In explaining the systematic distortions in spatial directional judgment between cities in different states, Stevens and Coupe thought the directions between cities in different states are biased by the directions between states, and proposed a hierarchy as the possible representation form of a cognitive map. This was further supported by (Hirtle & Jonides, 1985). It should be noted here, the kind of tree or ordered tree structured category hierarchy is not the same as hierarchy in taxonomy. It is largely the kind of *partonomy* as the term used in (Tversky & Hemenway, 1984). For example, in Rosch's category structure, a table is a piece of furniture. Anyway, a city is a part of a state in the tree or ordered tree, but a city is not a state. (Hereafter, this kind of hierarchy will be referred to *part-whole* hierarchy).

Practically, the preferred group of formal structures is *trees* or *ordered-trees,* such as the ordered-tree in (Hirtle, Ghiselli-Crippa, & Spring, 1993; Hirtle & Jonides, 1985). It is assumed implicitly that the tree or ordered tree structured category representation can be mapped to different situations. If this is the case, the set of categories as nodes in the tree or ordered tree should be very limited, usually restricted to a quite narrow application domain in order to balance its wide applicability in space-time. In the SI science domain, the tree is also used to represent knowledge of the "reality", where the instances of spatial entities as nodes are directly organised. These nodes are also called superordinate and subordinate units, but their difference is largely one of spatial-temporal scales such as cities and states. Vertically, there are the relations between superordinate and subordinate with links between nodes of different organisation levels, interpreted as the kind of part-whole relations (eg a state contains cities as its part). Horizontally inside a categorical unit, there are also nodes as distributions of spatial features from the same organisation level, such as the cities inside a state. There are also links between the nodes for spatial relations such as *direction* or *distances* (Stevens & Coupe, 1978). This kind of tree structure is particularly useful in network management applications with levels of management hierarchy, such as local network, or regional network.

The second group of structures for hierarchy is semi-lattices, such as those in (Kavouras & Kokla, 2002) (Kokla & Kavouras, 2001). (Hirtle, 1995) suggested that the semi-lattice is much better than tree structure for representing the hierarchical effects. Conceptually, the semi-lattice is a generalisation of an ordered tree according to (Hirtle, 1995). "It is defined formally as a collection of sets, such that for any two overlapping sets in the collection, the intersection of the sets is also in the collection" (Hirtle, 1995).

3.3. Semantic network and formal ontology approach

The most recent study in representation of categories and their structure is formal ontology research, which aims at finding and formally defining the primitive types of categories, which can be used for defining those categories and relations in various application domains. Such primitives are considered to include *frames*, *things*, *attributes*, *values* on the one hand, *is a*, *part-of*, and *whole of* relationships on the other hand. This formal research as in (Guarino, 1998) (Sowa, 1995, 2000) is a very top-down approach for "explicit specification of concepts and relations in a domain". Significant contributions have come from metaphysics, mereology and topology. There are also the very brute force approaches, as in the well-known CYC project criticized by Sowa (Sowa, 1995). The decision on what categories and relationships are the primitives and how they should be formally defined and implemented is still an ongoing work. There are some implementations of semantic networks or so-called conceptual graphs with limited types of categories and relationships. The nodes in the semantic network are the categories for spatial features, and links between the nodes are semantic relationships. Usually they are the *is a* or the *part-whole* relations. Both nodes and links have their semantics.

Contrary to the top-down approach in formal ontology research, studies on representing categories of spatial features in SI science are largely bottom-up. Ontology begins with the formal definition of spatial categories themselves and use some of the formally defined relations, in particular the *part-whole* relations. For representation of a category, the definition has considered the *parts, functions*, and *attributes* as in (Rodriguez & Egenhofer, 1999; Rodriguez & Egenhofer, forthcoming in 2003). Here *functions* refer the roles that the relevant spatial feature play, such as a school is for education. *Parts* can be the kind of distinctive properties. Informally, definition and representation of a cup can be as for drinking from a functional point of view, as with a handle and a body from the parts, and as made from metal, glass and ceramics. It can be seen that the way a category is represented has a strong influence from the method and measures used by Rosch to systematically investigate the category structure and prototype effect.

3.4. Concerns of categories of spatial features

Semantic representation is a complex issue, as both multiple levels of abstraction and multiple levels of organisation can be involved. Definitions and representation of categories and their relationships are the crucial and integrated part for semantic representation. A simple list of attributes is not the way an internal representation is made, and it does not help to construct taxonomy or partonomy.

For a group of structures as a tree, ordered tree and semi-lattice, one of the major concerns is the ad hoc way a hierarchy is formed. In these structures, it seems categories are not differentiated explicitly by their abstraction levels, and are thus treated similarly. The categories for the formation of category hierarchy can arise from any abstraction level. In the prototype theory, category representation begins from the basic levels and extends to superordinate and subordinate levels, as the consequences of generalisation and specialisation. It is hard to see how this well founded and accepted knowledge can be integrated. A second concern is the complex relationships between spatial features. There are at least two kinds of hierarchies for organising categories of spatial features. The *is a* hierarchy has relationships like those between *water coverage* and *rivers*, *lakes*, and *ponds*. The *part-whole* hierarchy has the relationships of *part-of*, *whole-of*, and *peer-to-peer*, such as a *state* contains *the city*, *a city* is contained inside *a state*, and *cities* inside *a state*. There is no explicit consideration of different types of relationships. Like abstraction and organisation levels, and types of relationships are exclusively in the mind of experts. The third concern is on its completeness as a set of formally defined concepts and their relations in a domain. It is very hard to claim a created tree or lattice is the final version without further changes. Given a larger tree, inserting of a new node may require considerable effort. It is also suspected when all domains are considered for a specific region, the domains should still stand independently or they should be integrated. The later sounds more reasonable. Given the fact that a region can be decomposed and clustered in different ways depending on points of views of the representation of the region, it is natural to expect the different views may be integrated. How the connection of different trees can be or is actually formed needs some attention.

Three further points need to be made. The first is a distinction between a category representation in terms of *functions, parts* and *attributes* as those in (Rodriguez & Egenhofer, 1999; Rodriguez & Egenhofer, forthcoming in 2003) with an intensional representation of categories in terms of *distinctive features*. The distinctive features function as *necessary* and *sufficient criteria* for a category. An intensional representation makes the statement that both robin and penguin are birds. A group of city examples gives the prototype of city, but does not necessarily give the intensional concept of a city. The second issue is that a prototype as the convergence of a set of measures in linguistic terms is not the same thing as the cluster of physical properties as in field theory or in the perception theory of the mind. Symbols from languages make sense beyond the pure forms and sounds. Humans cluster the attributes and their values as meanings, not as linguistic symbols. We adapt and organise the conceptions (represented by the symbols) in terms of their meaning as well. Thirdly, a collection of categories with a set of symbols does not make much sense to a computer, apart from the difference in lexicons. At present it is hardly possible for a machine to sort out the taxonomy or partonomy of spatial features, to integrate differently defined ontologies automatically, and to update the constant new stimuli from the environment. Although it may be argued that there are some *types* to restrict the digital interpretation of some symbolic statements, at present, these types are still too coarse and general.

4.0 REPRESENTATION AND ADAPTATION OF CATEGORY AND STRUCTURE

4.1. Common sense structure

It is believed that common sense structure is the foundation of science (Smith, 1995). Note that common sense structure is different from common sense. Common sense structure is the kind of framework that accounts for the assimilation of external input and results of common senses. Apart from the quite philosophical discussion in (Smith, 1995), there are also works of linguists and psychologists on exploring the forms of such cognitive framework. For example, the concept of “cognitive model” discussed by the linguist Lakoff in (Lakoff, 1987), which is explicitly considered as the kind of structure for the connection of those ideas from experimental psychology and philosophical thinking. The “cognitive model” of Lakoff can be understood as the generalised theory explaining at least the conscious part of cognition. While, the particular functioning description on what is common sense structure makes it remarkably similar to the “structuralism” concept that appeared in Piaget’s theory on cognition. Unfortunately, the concepts in Piaget’s work are often narrowed. This cognition structure is comprehensively and (sometimes) clearly unfolded in Piaget’s theory (see (Phillips, 1981) for an introduction of Piaget’s theory of cognition). Understanding “common sense structure” is non trivial. But there are several main points. In Piaget’s theory the separation between perception and conception is clearly emphasised (Piaget

& Inhelder, 1956). On the other hand, symbols from language are emphasised to be the kind of associated index for the concepts in mind, which can be arbitrarily selected but socially agreed upon (Phillips,1981). As further thought by Piaget, logic is the mirror of thought instead of other way around (Piaget, 1971). Detailed discussion of this issue is not the aim of this paper.

4.2. Concepts in SI science

In the domain of SI science, there are some particular and important issues. First, there should be a clear separation between the internal representation as concepts in the mind and the external forms of representation with symbols in natural languages. Categories defined with *parts*, *functions*, and *properties* in linguistic terms are only the external form of representation, even though such representation can be mapped to the internal representation quite accurately.

Second, there should be a clear separation between the instances and categories, such as the difference between Dunedin city and a concept of city. Instances exist in the world, they are the sources of prototype conceptions of different natural categories in mind.

Third, there are multi-facets in category structure, including *categories for spatial features*, *categories for events* and *categories of organisations kinds*. Organisations are formed as configuration of lower level spatial features and relations in a situation. They may be categories of spatial features as well. Categories for spatial features may be arranged on multi-levels of abstraction and on different scales. Different kinds of relationships, such as *is a, part-part, part of,* and *whole of* relationships can be involved. There are also quantitative and qualitative spatial relations of *distance* and *direction*. An organisation may be further generalised to suit a new context or application.

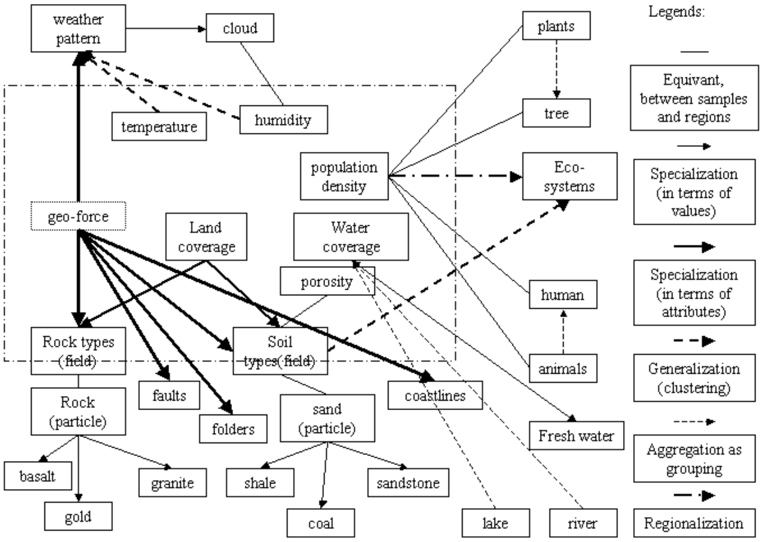

Figure 1. Complexity of structure of categories of spatial features and relationships, for illustrative purpose only

Fourth, the conceptual graph is hardly complete. The number of concept nodes in the category structure always increases, and more detailed differentiation of part-whole relations will be required.

At last, for a particular category of spatial features, the external form of representation is different. For example, the form of *lexicon*, the form of *template* with linguistic terms as *functions*, *parts* and *attributes*, the form of schemata for average cases or examplar with topological and geometrical details, or the form of distinctive properties under a specific context.

4.3. Representing categories of spatial features

4.3.1. Family resemblance on category representation

Apart from the basic idea of internal structure of categories presented previously, Rosch had also been working on the possible mechanism that results in the internal structure of categories (Rosch & Mervis, 1975). The idea is family resemblance. Wittgenstein criticised the classic category definition (Lakoff, 1987; Rosch & Mervis, 1975). He found that category like *games* can hardly be defined, since no single common property is shared by all games. He proposed that the category is rather, united by "family resemblances"(Lakoff, 1987). No set of properties need be shared by all family members so as long as A and B can share some properties, and B and C share others and so on. Thus, the members are similar in a wider sense.

The concept of family resemblance is accepted in (Rosch & Mervis, 1975). Rosch does not see why the two kinds of representation forms of prototype, one is in terms of cluster of attributes and the other in terms of example cases, should be contradictory. Like Tversky (Tversky & Gati, 1978), Rosch also considered the representation of category structures should be with similar and or different attributes. The hypothesis she proved is that a member judged to be the prototype shares more attributes with members belonging to the same category, and less common attributes with members belonging to other categories. The representation and reasoning on membership is based on the kind of *graded* similarity measure. The similarity measure developed by Tversky (Tversky & Gati, 1978) is widely used in reasoning.

4.3.2. Logic and prototype approaches for representation and reasoning

As was discussed in section 2, Rosch has emphasised that prototype effects should not be extrapolated (Rosch, 1978). She also stated that both mechanisms, "logical interpretation of categories in terms of definitions with necessary and sufficient criteria and prototypes in terms of clear cases and gradients of memberships" are both used in reasoning. The two mechanisms should be put in a broader context of reasoning (Rosch, 1983). Bottom-up, prototypes are needed for representation of categories of spatial features. Top-down intensional definitions with "distinctive properties" are also needed as restrictions on the memberships of categories of spatial features.

4.3.3. Adaptive and operative attitude towards representation of categories of spatial features

The contrast between the simplicity of the formerly discussed formal structures and the complexity of the real and mental reality of categories for spatial features and their relations is obvious. It has been argued that multiple formal structures are required for different levels of representation, the topological structures, the mereological structure and geometric structure (Mark & Frank, 1996), (Cohn, 1997). After all, a single generalized formal structure for representing the various kinds of concepts (includes categories of spatial features and their structures) is not pursued here, but rather an adaptive and operative approach is considered. In this approach, on the one hand the categories of spatial features and their structures are put in the cognitive structure of human mind and the concepts in SI science in general. On the other hand, the adaptive and operative framework also considers including the various formal structures for the representation of spatial concepts in general, categories of spatial features in particular. For such a purpose, the classification of categories of spatial features is one of the most important steps.

5.0 CLASSIFICATION OF CATEGORIES FOR SPATIAL FEATURES

5.1. What is meant by classifying categories of spatial features

Let's clarify what is meant by classification of categories of spatial features here. For logical reasoning to be achieved, there are the intensional level representations of categories with distinctive properties on the one hand, and a robust contextual frame to structure categories of spatial features on the other hand. In other words, there are at least two levels in a classification of spatial features. First, there are the domain level classification of spatial features restricted to the intensionally classified application, and second the top-level classification of categories of spatial features. The classification discussed in this paper relates to the second sense. In this later sense, it means the category of spatial features as a whole is the very top level superordinate. Under this superordinate, there are different kinds of categories of spatial features.

It may be noted, this very top-level superordinate category of spatial features would be similar to the category of *game* discussed by Wittgenstein. While a formal definition is elusive, judgment and wisdom can be used to decide whether something is a spatial feature or not. This does not preclude further differentiation, specialisation, and generalisation based on the properties (*as functions, parts and attributes*) of known members either. Second, the formal classification should be able to grow out of the natural way of concept formation and re-organisation in the mind. Maybe, this kind of re-organisation is based on the mechanism of family resemblance. Surely, it requires not only the ability to manipulate the symbols of natural language that have corresponding concrete existences in the world, but also the ability to "operate on operations" in mind (Phillips, 1981). Third, the dimensions considered for the classification are very abstractive. It is with consideration of the ways how the categories of spatial features can be defined and represented differently from each other. The dimensions of classification are not intended to be used as the rules directly to classify a given spatial entity or phenomenon into their corresponding categories. It is left to domain scientists to consider the detailed definitions and corresponding representation forms at this second level.

5.2. Purposes of classifying categories of spatial features

The purpose of classification relies on the answer to the question: *Is a top-level classification possible and necessary?* According to Usery (Usery, 1993), it is impossible to define a whole set of categories for spatial features and their relations, but a reasonable subset of categories and their relations for a particular application area is. This is true if a complete enumeration of the set members is what is meant. But, the impossibility to enumerate every biological type does not prevent biologists from giving the taxonomy of biology. This is similar to the category of spatial features. Also the bottom up approach to define categories may pose potential problems later when the data from different application domains really needs to be integrated as discussed early in section 3.4. There will be remaining issues on how to integrate these application-oriented and different ontology specifications as well. This is especially important for any implementation of a general purpose SI system.

Second, an adaptive and operative attitude is required for "explicit specification of concepts and relations", given categories of spatial features are of different abstraction levels and organization scales and different relationships between and among them. The integrated structure for taxonomy and partonomy of spatial categories needs to maximise the ability to assimilate digital representations of those instances of spatial features. It also needs to assimilate new definitions, adaptation of definitions for individual categories and their representation schemata. Sometimes, adaptation of the internal structure of categories may also be involved. It will be an advantage overall if the change of representation structure of categories can be minimised. In the short term, maximum flexibility should be allowed and in the future the assimilation can be more restricted.

Third, when making the classification, there is also the consideration of logico-mathematical formalisation and further digital implementations as the models of categories of spatial features and relations. Categories may be represented in natural languages or intensional representation in logic/mathematical form. There may only be partial knowledge of definitions of some categories and therefore the external forms of representation for such categories will be restricted. The classification will help to outline what kinds of categories are definable at which levels of details, how the definitions are represented and adapted from one form into another.

Finally, relations between and among different categories are more complicated than the well-known relationships as *is a* relation or *part-whole* relation. A clear exploration on the nature of categories of spatial features can make the relationships more specific and more easily implementable. Even if the relations between concepts are of the same name, there may be good reason to believe the interpretation and digital implementation could be different. For example, there are possible different interpretations of *part-whole* relations between a city and a state and between branches of a river and a river. The consideration of exact interpretation and implementation of relevant relations is beyond the scope of the paper.

5.3. Classification of categories for spatial features

5.3.1. Typology of users of SI systems

There are two main starting points for the consideration of classification. One of them is the very promising approach of classifying the application areas of SI systems is presented in (Burrough & Frank, 1995) (Couclelis, 1996). This typology classifies users of SI systems into five main categories:

1. public at large
2. managers of defined objects
3. modellers of space-time
4. planners and resource managers, and
5. politician and entrepreneurs

Here, it is implicitly or explicitly acknowledged that different categories of spatial features are dealt with by users of different application domains. Differences of categories in these different application domains should be clarified further, in terms of the nature of the categories for their conceptual existences, external forms of representation, and adaptation of the representation forms.

5.3.2. Spatial features in different spaces

Apart from typology by users, another starting point for classification considers the idea of different spaces (Couclelis, 1992; Montello, 1993). Spaces can be divided into four groups. The first is the small palm-top or table-top space, which includes categories like *cups*, *books*, *chairs*, *computer monitors*, *human body*, *small animals*, and *small plants*. This space was called A-space (Couclelis, 1992). Objects in this space are small enough to be manipulated and to be viewed without movement. The conceptions for the objects can be formed with little mental abstract generalisation. The other three groups include spatial features that are much larger, and cannot be viewed with one snapshot. B-space is also termed as cityscape, which refers to those spatial features like buildings, road and probably vehicles[1]. C-space is the landscape scale, which includes spatial features, whose conceptions need some level of abstract generalisation over the physical properties perceived. For example, a distant mountain , built-up area for cities, and other things like soil types. They can still be viewed from some vantage point. D-space includes features which are not included in the former spaces and which are usually regions and realms that are beyond the direct personal experiences. This space is the least understood (Couclelis, 1992).

5.3.3. Classification of categories of spatial features

This paper concentrates on those features within B-, C-, and D-spaces. As discussed in (Usery, 1993), the exact meaning of spatial features is similar to spatial regions in geography. Although classification of spatial features into different spaces is convenient and easy to understand, it is not sufficient for an adaptive and operative schema. Five other dimensions are required:

1. Object-like or field like
2. Spatial features of fixed or floating natures
3. Spatial features as conception of individual and of population
4. Spatial features with physical and social variables
5. Spatial features of natural or human-made spatial-temporal boundaries

These dimensions for classification need to be used in combination. The notion of spatial features as objects or fields is the classic dichotomy. An *object* is an existence in the void of space-time, or a perceptually contrasting background. A *field* means continuity itself over space-time. In this sense, a *field* might be understood as a situation of spatial-temporal frame. With differentiation of objects and fields, it is possible to differentiate between a soil particle (an object) and the bordered patch that is named as a soil type (a field). The field-object differentiation is on the nature of the categories, and should not be confused with the form of representation in terms of vector and raster.

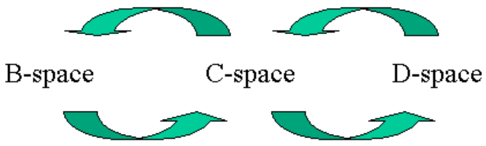

Figure 2. Relations between spatial features on spaces of different scales. Following the top arrows, hierarchy of formal definition for spatial regions (as agreement among population groups) is with decreasing spatial resolutions from right to left. Following the bottom arrows, hierarchy of natural boundaries is constructed in the mind of individual with less and less direct perceptions can be associated from left to right.

The conception of spatial features is both a private and public. The concept differences are in terms of individual-individual, individual-group, group-group. For example, the *commerce building* at University of Otago is a place to work for some, and a place to buy lunch for others. The role played by particular spatial features to an individual and to the public should be separated. Categories of spatial features with pure physical properties are perceptually definable. Categories for spatial features with both physical and social properties are more complex. The exact boundary of their existence in space-time is rather more a social agreement under differing forces. With consideration, a category may be defined with pure physical variable or with both physical and social variable. For example, a city can be strictly differentiated from a built-up area. First, the concept of built-up area is the superordinate category for human-made fixed spatial features, while the concept

---

[1] It is hard to see whether the categories of trees and elephants should be included in A-space or in B-space. The image of a big tree or an elephant perhaps could not sometimes be obtained in one snapshot and could not be manipulated unless to cut a chunk off them. But, it looks very unnatural to include them in B-space, and separate them with other small plants and animals. This exact example perhaps tells how weak it may be to use such a classification of space directly for a classification of categories of spatial features which are pursued here.

of a city is a specialisation. Second, the spatial boundary of a built up area as an object does not match exactly with the spatial boundary of a city, though there may be high degree of correlation.

The differentiation of representations of spatial features with natural boundaries and human-made boundaries is another important dimensions for classification. Spatial regions can be defined by their natural boundaries or in terms of formal spatial boundaries of geo-coordinates. One concern is unavoidable arbitrary aspect of the exact spatial boundary because of the inherent connection of spatial features with environment. The other is the offset between the natural boundary as a system of interaction and the human-made boundary defined by geo-coordinates (Couclelis, 1992). Present knowledge is generally limited to the notion of a rigid spatial boundary for the regions of larger scale features. The human-made boundaries of spatial features with both physical and social variables evolve via social events. (see figure 2).

Applying these classification dimensions should give the category of interest a restricted position in multidimensional conceptual space. The potential forms of representation for the kind of categories and consequently the forms of representation for instances and their adaptation rules and mechanisms can be modelled. The issues on how the classification schema can be used and quantitative discussion on the classification are not covered here. There are no discussions on how the categories of spatial features interested in by different user types are defined and represented, how structured categories are connected with each other, and how they are put in the context of concepts in SI science and the context of cognitive structure as a whole.

6. CONCLUSION

To design the next generation of spatial systems, it will be necessary to utilise advanced concepts on categories. This paper has focused on this issue of categories of spatial features. First, it explored the main theories on category representation in general. Then there was a brief description of some formal structures used in representation of categories of spatial features and the structure of categories as a whole. Existing approaches are like: definitions of a list of spatial features without structure to relate them; the attempt to use single representation structure of taxonomy or partonomy of spatial features as tree, ordered tree, semi-lattices; representation in semantic networks model with limited and too general primitives; and indiscriminate imitation of methods used in cognitive experiments without considering that symbols are different from their references in reality and conceptions in the mind. It should be noted that spatial features are on different abstraction levels, belong to different levels of organisation, and categories can be represented in different forms. An alternative adaptive and operative approach is proposed, similarly as suggested in (Couclelis, 1996), which is termed as an operational approach to represent spatial features. While it should not only be *operative* towards boundaries of instances of spatial features, but also towards the conceptual definitions of categories themselves and the structure of categories as taxonomy and partonomy, where exact spatial boundaries may not be included.

For such an operative and adaptive approach to be realised, the first step will depend on a classification of the categories in the SI science domain in general, which based on the nature of categories in the relevant application domains. The expectation of the classification is: it needs to consider both the structuring mechanisms as generalisation and specialisation and structured results as categories on different levels of abstraction and organisation scale. It also needs to consider that since different definitions in terms of accuracy and robustness can be given for a same category, a category can have different forms as its external representations correspondingly, such as in terms of natural or in logico-mathematical language. The exploration on the nature of categories will help the judgment of availability, evaluation of the adequacy, and guide the adaptation of spatial boundary representation for instances of spatial features in the context of an application. This will be an advancement of the present generation of spatial systems.

ACKNOWLEDGEMENTS

The work is partially supported by a scholarship from University of Otago, School of Business.

REFERENCES

Berlin, B. (1978). Ethnobiological classification. In E. Rosch & B. Lloyd (Eds.), *Cognition and Categorization* (pp. 79-98). Hillsdale N.J.: Lawrence Erlbaum Associates Inc. Publishers.

Burrough, P. A., & 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.

Cohn, A. G. (1997). Qualitative spatial representation and reasoning techniques. In G. Brewka & C. Habel & B. Nebel (Eds.), *KI-97: Advances in Artificial Intelligence* (pp. 1-30). Berlin: Springer.

Couclelis, H. (1992). People manipulate objects (but cultivate fields): beyond the raster-vector debate in GIS. In A. U. Frank & I. Campari & U. Formentini (Eds.), *International Conference GIS - from Space to Territory: Theories and Methods of Spatio-temporal Reasoning* (pp. pp. 65-78). Pisa Italy: Springer-Verlag.

Couclelis, H. (1996). Towards an Operational Typology of Geographic Entities with Ill defined Boundaries. In P. Burrough & A. Frank (Eds.), *Geographic Objects with Indeterminate Boundaries* (pp. 45-55): Taylor & Francis.

Fegeas, R. G., Cascio, J. L., & Lazar, R. A. (1992). An overview of FIPS 173 the Spatial Data Transfer Standard. *Cartography and Geographic Information Systems, 19*(5), 278-293.

Fonseca, F., Egenhofer, M., & Davis, C. (2000). Ontology-driven information integration, *AAAI-2000 Workshop on Spatial and Temporal Granularity*. Austin TX, Aug., 2000.

Fonseca, F., Egenhofer, M., Davis, C., & Camara, G. (2002). Semantic granularity in ontology-driven geographic information Systems. *Annals of Mathematics and Artificial Intelligence, 36*(1-2), 121-151.

Fonseca, F. T., Egenhofer, M. J., Agouris, P., & Camara, G. (2002). Using ontologies for integrated geographic information systems. *Transactions in GIS, 6*(3), 231-257.

Frank, A. U. (2001). Computational models for ontology, *Proceedings of FOIS 2001 International Conference on Formal Ontology in Information Systems, October 17-19*. Ogunquit Maine.

Frank, A. U. (2001). Tiers of ontology and consistency constraints in geographical information systems. *International Journal of Geographical Information Science, 15*(7), 667-678.

Frank, A. U. (to appear). Ontology for spatio-temporal databases. In T. Sellis, et al (Ed.), *Spatiotemporal Databases: The Chorochronos Approach*.: Springer-Verlag.

Gleitman, H. (1995). *Psychology* (4th ed.). New York, London: W. W. Norton & Company.

Gruber, T. R. (1995). Toward principles for the design of ontologies used for knowledge sharing. *International Journal of Human-Computer Studies, 1995*(43), 907-928.

Guarino, N. (1995). Formal ontology conceptual analysis and knowledge representation. *International Journal of Human-Computer Studies, (1995)43*, 625-640.

Guarino, N. (1998). Formal ontology and information systems. In N. Guarino (Ed.), *Formal Ontology in Information Systems(Proceedings of FOIS98 Trento Italy 6-8 June* (pp. 3-15). Amsterdam Berlin Oxford Tokyo Washington DC: IOS Press.

Hirtle, S. C. (1995). Representational structures for cognitive space: trees ordered trees and semi-lattices. In F. A. U. a. K. W. (Ed.), *Spatial Information Theory: A Theoretical Basis for GIS* (pp. 327-340). Berlin: Springer.

Hirtle, S. C., Ghiselli-Crippa, T., & Spring, M. B. (1993). The cognitive structure of space: an analysis of temporal sequences, *Spatial Information Theory: A Theoretical Basis for GIS, Proceedings European Conference COSIT'93 Marciana Marina Elba Island Italy September 19-22 1993* (Vol. 716, pp. 177-189): Springer-Verlag.

Hirtle, S. C., & Jonides, J. (1985). Evidence of hierarchies in cognitive maps. *Memory and cognition, 13*, 208-217.

Kavouras, M., & Kokla, M. (2002). A method for the formalization and integration of geographical categorization. *International Journal of Geographical Information Science, 16*(5), 439-453.

Kokla, M., & Kavouras, M. (2001). Fusion of top-level and geographical domain ontologies based on context formation and complementarity. *International Journal of Geographical Information Science, 15*(7), 679-687.

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

Mark, D. M., & Frank, A. U. (1996). Experiential and formal models of geographic space. *Environment and Planning B: Planning and Design, 23*, 3-24.

Montello, D. R. (1993). Scale and multiple psychologies of space, *Spatial Information Theory, A Theoretical Basis for GIS, Proceedings, European Conference, COSIT'93* (pp. 312-321). Marciana Marina, Elba Island, Italy, September 19-22, 1993: Springer-Verlag.

Phillips, J. L. J. (1981). *Piaget's Theory: A Primer*. San Francisco: W.H. Freeman and Company.

Piaget, J. (1971). *Structuralism*. New York: Basic Books.

Piaget, J., & Inhelder, B. (1956). *The Child's Conception of Space*. London: Routledge and Kegan Paul.

Raper, J., & Livingstone, D. (1995). Development of a geomorphological spatial model using object-oriented design. *International Journal of Geographical Information Systems, 9*(4), pp. 359-383.

Rodriguez, A., & Egenhofer, M. (1999). *Assessing semantic similarities among geospatial feature class definitions.* Paper presented at the Interoperating Geographic Information Systems, Proceedings of the 2nd International Conference INTEROP'99, Zurich, Switzerland.

Rodriguez, A., & Egenhofer, M. (2003). Determining semantic similarity among entity classes from different ontologies. *IEEE Transactions on Knowledge and Data Engineering, 12*(2).

Rosch, E. (1978). Principles of categorization. In E. Rosch & B. Lloyd (Ed.), *Cognition and Categorization* (pp. 27-48). Hillsdale N.J.: Lawrence Erlbaum Associates Inc. Publishers.

Rosch, E. (1983). Prototype classification and logical classification: the two systems. In E. K. Scholnick (Ed.), *New Trends in Conceptual Representation: Challenges to Piaget's Theory* (pp. 73-86). Hillsdale N.J.: Lawrence Erlbaum Associates Publishers.

Rosch, E., & 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 M. T. E. (Ed.), *Cognitive Development and Acquisition of Language* (pp. 111-144). New York: Academic Press, Inc.

Smith, B. (1995). Formal ontology common sense and cognitive science. *International Journal of Human-Computer Studies, 1995*(43), 641-667.

Smith, B., & Mark, D. M. (1998). Ontology and geographic kinds, *Proceedings International Symposium on Spatial Data Handling(SDH'98)*. Vancouver Canada, July 12-15.

Sowa, J. (1995). Top-level ontological categories. *International Journal of Human-Computer Studies, (1995)43*, 669-685.

Sowa, J. (2000). *Knowledge representation: logical, philosophical, and computational foundations*. Pacific Grove, CA: Brooks Cole Publishing Co.

Stevens, A., & Coupe, P. (1978). Distortions in judged spatial relations. *Cognitive Psychology, 10*, 422-437.

Tang, A. Y., Adams, T. M., & Usery, E. L. (1996). A spatial data model design for feature-based geographical information systems. *International Journal of Geographical Information Systems, 10*(5), pp. 643-659.

Tversky, A., & Gati, I. (1978). Studies of similarity. In E. Rosch & B. Lloyd (Ed.), *Cognition and Categorization* (pp. 79-98). Hillsdale N.J.: Lawrence Erlbaum Associates Inc. Publishers.

Tversky, B., & Hemenway, K. (1984). Objects parts and categories. *Journal of Experimental Psychology: General, 113*(2), 169-191.

Usery, E. 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 approaches to geo-referenced information. *International Journal of Geographical Information Systems*, *8*(4), 385-399.