# Towards a Unified Spatial-Temporal Data Model and Query Language for Geographical Information Systems

**Patrick Browne**
*School of Computing*
*Dublin Institute of Technology*
*Kevin Street, Dublin 8, Ireland.*
*Patrick.Browne@comp.dit.ie*

**Mike Jackson**
*School of Computing and Information*
*University of Central England in Birmingham*
*Perry Barr, Birmingham BT42 2SU*
*Mike.Jackson@uce.ac.uk*

**Abstract**

*This paper outlines the principle features of a Ph.D. programme of work. Section one describes the aims of the research. Section two looks at the nature of digital maps and their importance in the broader context of geographic information. Section three looks at the two major aspects of the research. Section four discusses the role of category theory in the research. Finally the conclusion is presented in section five.*

**Keywords** – Geographic Information System, Formal System Specifications.

## 1 Introduction

The prime aim of this research is to develop a formal data model and query language, which can facilitate the management of topographic information that varies over time. Topographic data can be processed within a Geographical Information System (GIS). A further aim is the development of a formal approach that can be used to address the diverse modeling issues that arise when handling complex topographic data.

## 2 Digital maps

A map can be described as a rule-based abstraction of reality, which is intended to convey information. The map is a result of applying rules to objects on the earth's surface and translating them into a graphical and informational representation. The map itself is not a rule, but the co-domain of a particular type of rule called a function. The rules are usually written in the form of a specification, which is designed by the map author to meet an information need. With additional temporal functionality digital maps can record an archive of the past, provide an accurate view of the present and allow us to develop future plans (what-if scenarios).

## 3 General approach

Existing spatial data models [1][2][3] and the various standards such as the Open GIS Consortium (OGC) [4] have to some extent identified the salient aspects of the real world that can be processed within a GIS. The model proposed is called the Topographic Data Model (TDM). The mechanism by which the TDM will be constructed will take the novel approach of using category theory as the basis for the modeling task, see Pierce [5]. Topographic information (TI) can be considered as being derived from real world surveys and descriptions of particular spatial objects, for example houses, roads and rivers. In this research TI is regarded as a subset of geographic information. The models required to represent real world objects using digital maps are intricate and multifaceted. Within a GIS the insert, delete and update operations are considerably more complex than the analogous operations on tabular data.

### 3.1 GIS research

The primary aim of this research is to provide a data model and query language capable of addressing issues that arise when handling TI. TDM will integrate some aspects of existing models. A candidate model could be the Dual Grid from Lema, and Gutting[1] which is built on the earlier Rose Algebra[2]. Another strong candidate would be the ST-Complex model developed by Worboys[3]. These are mathematically based models that address various temporal and spatial research issues. The TDM will be a spatial-temporal data model. It will have a formal set of, axioms, functions and relations that can be used to facilitate the automatic processing of TI.

### 3.2 Software Engineering Research

The second aim of the research is to explore a formal approach to the modeling process. This formalism should permit the expression of the desired complexity and diversity that is inherent in topographic data. Abraham and Roddick[6] noted the absence of a universal spatial-temporal model and stated: *".. it is unlikely that an all encompassing model will cater for all individual requirements."*. The real world is far too complex to represent in its entirety within a single model. This has given rise to many diverse and narrowly focused models, which, although adequate for their intended use, are difficult to reuse or integrate. Rather than attempt to develop some universal model, this research will look at ways of integrating various sub-models. By using a building block approach the diverse modeling tasks and domains can be tackled individually. Then the individual models can be joined or fused using a unification mechanism. Many researchers feel that category theory has a major role to play in the modeling of complex systems, see Frank [7][8][9], Lippe et al[10], Ehrig et al[11], Johnson et al [12], Cadish and Diskin[13], and Goguen[14]. For this research category theory will be the dominant modeling technique.

## 4 Nature of research.

Category theory (CT) can be used as a unifying framework for specifying complex systems, see Goguen [14]. Category theory characterizes the internal structures of objects through mappings between them, see Cadish and Diskin [13]. CT is quite abstract, it provides rules on how objects may be joined but it does not make any assumptions about their internal structure. Within CT the term *object* is very general. For example in the category of ***Set*** the objects are sets (e.g. a set of houses may be related to a set of streets) while in the category of ***Algebra*** the objects are algebras. Objects are never considered in isolation, the rules by which they are connected (called morphisms) must be included. In the category ***Set*** the morphisms are functions while homomorphisms are the morphisms in the category ***Algebra***. On the other hand CT does provide rules that preserve the structure of objects. So while CT is highly abstract it is appropriate for handling structured objects, such as algebraic specifications. In this research CT will be used as a technique and as a meta-technique, which sometimes will be used directly, but will also be used indirectly in that it forms the basis of some other formalism such as algebraic specifications.

The high level (or abstract) nature of CT means that commonality between apparently distinct objects can often be revealed and precisely described. In the general case many formal systems have objects in common, for example, different programming languages often share the concepts of sequence, condition and iteration. The various implementations of sequence, iteration and condition may differ considerably between languages. In the special case of spatial-temporal models there will be common concepts occurring in different spatial models (e.g. representation of intervals). CT offers a way of discovering and describing these commonalities. Further CT offers techniques to join different objects in various ways (such as categorical products and sums). By using the concepts of limit and colimit these joins can be shown to be canonical. These ideas offer a formal way to unify objects. Some computer science researchers use CT as a uniform notational framework to manage masses of relatively elementary detail, others use deep theorems from CT applied to complex computations, see Pierce [5]. This research will concentrate on the former approach rather than the latter.

## 5 Conclusion

The aim of the research is to develop a spatial-temporal data model capable of supporting queries over topographic information by combining various sub-models. The objects of interest are topographic objects and the time domain is calendar time. The research will explore the current approaches to model unification. We will examine appropriate existing models and adapt them as necessary using existing methodologies and formalisms. The main research effort centers on fusing or unifying these models in a coherent fashion. The TDM will take a holistic approach and should be capable of modeling the metric, directional, topological, network, set-theoretic, logical and temporal aspects of geographic space and its embedded topographic objects. The central academic challenge will be to establish the degree to which CT can be used to unify these diverse concepts.

## 6 References

[1] Jose Antinio Cotello Lema, Ralf Hartmut Gutting. Dual Grid: A new approach for robust spatial algebra implementation. GeoInformatica 6:1, 57-76 (2002), Kluwer Academic Publishers.

[2] R.H. Gutting and M. Schneider: "Realms: A Foundation for Spatial Data Types in a Database System" Advances in Spatial Databases: Springer Verlag vol. 692, 1993

[3] M.F. Worboys, "A Unified Model for Spatial and Temporal Information." The Computer Journal Vol 37 Number 1 The British Computer Society.

[4] The Open GIS Consortium develops open publicly available geo-processing specifications. See http://www.opengis.org

[5] Basic Category Theory for Computer Scientists, by Benjamin C. Pierce, ISBN 0262660717

[6] Survey of Spatio-Temporal Databases Tamas Abraham And John F. Roddick, GeoInformatica, 1999 Kluwer Academic Publishers.
[7] Andrew. Frank: Specifying Open GIS with Functional Languages, 1995, Department of Geoinformation, Technical University Vienna, Austria.
[8] A. Frank and D Medak, Executable axiomatic specification using Functional language. From proceedings of EJCIMKB'97: Executable Axiomatic Specification 1997.
[9] A. Frank: One Step up the Abstraction Ladder, from COSIT 1999, Springer Lecture Notes in Computer Science, 1999, ISBN 3-540-66365-7
[10] E. Lippe and A.M.H. ter Hofstede. A Category Theory Approach to Conceptual Data Modelling, Computer Science Institute University of Nijmegen, The Netherlands.
[11] H. Ehrig, M. Grobe-Rhode, U. Wolter, The role of Category in the area of Algebraic Specifications, Technical University of Berlin, Germany.
[12] Entity-Relationship-Attribute Designs And Sketches by Johnson Michael Johnson, Robert Rosebrugh And R.J. Wood in Theory and Applications of Categories, Vol. 10, No. 3, 2002, pp. 94-112.
[13] B. Cadish and Z. Diskin. Algebraic Graph-Oriented = Category Theory Based. Technical Report 06, Frame Inform Systems, Riga,Latvia
[14] A Categorical Manifesto, by Joseph Goguen, in Mathematical Structures in Computer Science, Volume 1, Number 1, March 1991, pages 49-67.

## 7 Copyright