Chapter XIX
Fuzzy Spatial Data Types for Spatial Uncertainty Management in Databases

Markus Schneider
University of Florida, USA

ABSTRACT

Spatial database systems and geographical information systems are currently only able to support geographical applications that deal with crisp spatial objects, that is, objects whose extent, shape, and boundary are precisely determined. Examples are land parcels, school districts, and state territories. However, many new, emerging applications are interested in modeling and processing geographic data that are inherently characterized by spatial vagueness or spatial indeterminacy. This requires novel concepts due to the lack of adequate approaches and systems. In this chapter, we focus on an important kind of spatial vagueness called spatial fuzziness. Spatial fuzziness captures the property of many spatial objects in reality that do not have sharp boundaries and interiors or whose boundaries and interiors cannot be precisely determined. We will designate this kind of entities as fuzzy spatial objects. Examples are polluted areas, temperature zones, and lakes. We propose an abstract, formal, and conceptual model of so-called fuzzy spatial data types (that is, a fuzzy spatial algebra) introducing fuzzy points, fuzzy lines, and fuzzy regions in the two-dimensional Euclidean space. This chapter provides a definition of their structure and semantics, which is supposed to serve as a specification of their implementation. Furthermore, we introduce fuzzy spatial set operations like fuzzy union, fuzzy intersection, and fuzzy difference, as well as fuzzy topological predicates as they are useful in fuzzy spatial joins and fuzzy spatial selections. We also sketch implementation strategies for the whole type system and show their integration into databases. An outlook on future research challenges rounds out the chapter.
INTRODUCTION

Spatial database systems (SDBSs) are full-fledged database systems that, in addition to the functionality of standard database systems for alphanumeric data, provide special support for the storage, retrieval, management, and querying of spatial data, that is, objects in space. In particular, SDBSs are used as data management components of geographical information systems (GISs). For modeling and storing spatial data, special data types called spatial data types (see Schneider, 1997, for an introduction and survey) have been designed. The most established spatial data types are called point, line, and region. We speak of spatial objects as instances of these data types. So far, these data types are only able to represent spatial data whose extent and hence border are precisely determined and homogeneous. Spatial phenomena are represented by sharply described points with exactly known coordinates, lines linking a series of exactly known points, and regions bounded by exactly defined lines called boundaries. The properties of the space at the points, along the lines, or within the regions are given by attributes whose values are assumed to be constant over the total extent of the objects. Well-known examples are man-made spatial objects representing engineered artifacts like highways, houses, or bridges, and some predominantly immaterial spatial objects exerting social control like countries, districts, and land parcels with their political, administrative, and cadastral boundaries. We denote this kind of entities as crisp or determinate spatial objects.

In the last years, geographical applications and GIS have shown increasing interest in data models for spatial data that are characterized by the inherent feature of spatial vagueness or spatial indeterminacy (Burrough & Frank, 1996). The reason for this interest in these applications is that the current mapping of spatial phenomena of the real world to exclusively crisp spatial objects has turned out to be an inadequate abstraction process for many kinds of spatial data. Consequently, applications based on indeterminate spatial data are not covered by current GIS and SDBS. In this chapter we focus on a special kind of spatial vagueness called spatial fuzziness. Fuzziness captures the property of many spatial objects in reality that do not have sharp boundaries and interiors or whose boundaries and interiors cannot be precisely determined. Examples are natural, social, or cultural phenomena like land features with continuously changing properties (such as population density, soil quality, vegetation, pollution, temperature, air pressure), oceans, deserts, English-speaking populations, or mountains and valleys. The transition between a valley and a mountain usually cannot be exactly ascertained so that the two spatial objects valley and mountain cannot be precisely separated and defined in a crisp way. We will designate this kind of entities as fuzzy or indeterminate spatial objects.

This chapter delineates ongoing research on fuzzy spatial objects in databases and pursues several goals. First, we present an abstract, formal, and conceptual object model providing fuzzy spatial data types for fuzzy points, fuzzy lines, and fuzzy regions in the two-dimensional Euclidean space. Second, we introduce some fuzzy spatial operations and a concept of fuzzy topological predicates. The entire effort is supposed to lead to a formal specification of a fuzzy spatial algebra. Fuzzy set theory (Buckley & Eslami, 2002; Zadeh, 1965) and fuzzy topology (Liu & Luo, 1997) are the formal framework of the algebra design. Third, we show how fuzzy spatial data types, operations, and topological predicates can be integrated into database systems and their query languages, and be used as attribute types in database schemas. Fourth, we provide some implementation ideas.

The section “Related Work” explains different aspects of spatial vagueness and presents related work. The section “Abstract Definition of Fuzzy Spatial Data Types” introduces some basic concepts and notations of fuzzy set theory and fuzzy topology, as far as they are used in this chapter, and formally defines the data types for fuzzy points, fuzzy lines, and fuzzy regions. In the section “Fuzzy Topological Predicates,” we introduce a
Related Content

Online Data Mining
[www.igi-global.com/chapter/online-data-mining/11184?camid=4v1a](www.igi-global.com/chapter/online-data-mining/11184?camid=4v1a)

Cost Modeling and Range Estimation for Top-k Retrieval in Relational Databases
[www.igi-global.com/chapter/cost-modeling-range-estimation-top/52960?camid=4v1a](www.igi-global.com/chapter/cost-modeling-range-estimation-top/52960?camid=4v1a)

Assigning Ontological Meaning to Workflow Nets
Pnina Soffer, Maya Kaner and Yair Wand (2010). *Journal of Database Management* (pp. 1-35).
[www.igi-global.com/article/assigning-ontological-meaning-workflow-nets/43728?camid=4v1a](www.igi-global.com/article/assigning-ontological-meaning-workflow-nets/43728?camid=4v1a)

An Integrated Query Relaxation Approach Adopting Data Abstraction and Fuzzy Relation
[www.igi-global.com/article/integrated-query-relaxation-approach-adopting/47419?camid=4v1a](www.igi-global.com/article/integrated-query-relaxation-approach-adopting/47419?camid=4v1a)