Spatial Knowledge-Based Applications and Technologies: Research Issues

Elisa Bertino[1] and Maria Luisa Damiani[2]

[1] Department of Computer Sciences, Purdue University, West Lafayette, Indiana (USA)
[2] Dipartimento di Informatica e Comunicazione, Università degli Studi di Milano, Via Comelico 39/41, 20135 Milano (Italy)

Abstract. Spatio-temporal data and knowledge have become increasingly important in the last years. In this paper, which introduces the session on *Knowledge Engineering Techniques for Spatio-temporal Applications*, we outline some issues in geo-spatial research and applications for which the use of innovative knowledge, tools and methodologies is crucial.

1 Introduction

In the last decade we have witnessed significant evolutions of data, knowledge and applications in the geo-spatial domain. Such evolutions have been pushed by the availability of wireless networks and portable devices that have enabled the development of significant location-based services. Moreover, new important application areas have emerged, such as homeland security and mobile-commerce, which pose new interesting challenges. In particular, what is most striking in many of these innovative applications is the increasing pervasiveness of spatial information. Whereas until few years ago, only a limited number of organizations were able to afford a geographical information system (GIS) because of its high cost and complexity, nowadays geo-spatial information is going to be integrated into the business processes of a variety of different enterprises, even not directly involved in geographical-related sectors. Geo-spatial information has become the basic ingredient of applications that are or are going to be of common use, such as location based services, user navigation and ubiquitous computing. Increasingly thus geo-spatial information is deployed in heterogeneous contexts and used by people that often have little familiarity with geo-spatial concepts.

It can be also observed that current spatially-aware applications often deal with increasingly larger geographical areas. In contrast with early applications that were mostly concerned with local geographical phenomena, such as cadastre management and urban planning, new emerging applications may scale up to a continental or worldwide dimension and involve a large number of players spread across large geographical areas. Applications related with the development of Spatial Data Infrastructures are a significant example of such a trend. A Spatial Data Infrastructure (SDI) can be defined as infrastructure consisting of IT systems and policies that facilitate the availability of and access to spatial data. In the SDI Cookbook [1], a SDI is defined as "a basis for spatial data discovery, evaluation, and application for users and providers within all levels of government, the commercial sector, the non-profit sector, and academia and

by citizens across different jurisdictions ". A SDI thus aims at establishing a bulk of rules and knowledge to be shared across cultural and administrative boundaries. The establishment of an infrastructure for spatial information at European scale has been recently approved by the European Commission [3]. For the realization of a SDI, however, like for all the applications based on the cooperation of different communities of users, it is fundamental to ensure the interoperability among the sources of geo-spatial information, that may differ for example in resolution and spatial and thematic content.

Another characteristic of current applications is that the notion of space has also evolved. We can distinguish among different kinds of space: besides the *data space*, that is, the space populated by observations of the real world, there is a *meta-data space* populated by meta-objects, representing spatial resources such as geo-spatial data sources and user communities. Since the spatial resources may be characterized by a position and evolve in time, they can be considered as themselves embedded in a space and be possibly spatially interrelated. An example of spatial relationship at the metadata level is that of nearness between cadastral data sources. Finally, space can also consist of the *artificial space*. Examples of artificial spaces are the virtual landscapes built using virtual reality and videos. Note that there are applications in which all these spaces need to coexist.

For the development of spatially aware applications characterized thus by novel kinds of users and services, extended geographical scope and information dissemination, the technology has to deal with a number of challenges. In what follows we briefly discuss some research themes in which spatial knowledge can play an important role. Finally we focus on a possibly reference application area in which these technological advances would find a social and economic motivation.

2 Research Challenges

2.1 Geographical Ontologies

A classical and well-known issue in geo-spatial data management is data integration. Geo-spatial data are inherently heterogeneous simply because the real world can be observed from multiple points of views. However, nowadays the problem is further exacerbated by the need of integrating geo-spatial information across cultural and administrative boundaries, as requested for example by SDIs. Whereas the standardization initiatives that started in 1994 with the constitution of OpenGIS Consortium (now Open-Geospatial Consortium), has contributed to solve the problem of physical heterogeneity by defining common software specifications and interfaces, the issue of semantic heterogeneity of spatial information is still open. One aspect of such heterogeneity is the different conceptualization of spatial objects. An interesting example of such a diversity in conceptualization is reported by Bennet in [2]. In particular, this paper discusses the concept of *forest* and presents different interpretations of this term and how these interpretations can affect the determination of the spatial extensions of the spatial objects classified as forests. In order to deal with the semantic heterogeneity of spatial information, the development of geographical ontologies is crucial. A geographical ontology (geo-ontology) defines the spatial concepts that are shared by a community of users, representing thus the semantic reference for users and applications that accept to align

their spatial data with such an ontology [10][6]. The formalization of geo-ontologies either for general purpose or for specific domains is thus a major technological challenge. In this view, a notable case study is the GEON project (Geo-sciences Network) currently under development in the US [5]: a major goal of the project concerns the development of domain-specific ontologies for geology and more in general for geo-sciences. Research issues related to the definition of ontologies include how to fuse information compiled using different ontologies possibly developed by multiple communities. In general, the mapping between ontologies is not automated and relies on the application of domain expertise. Recent approaches, however, try to automate such a process of geospatial information fusion [4].

2.2 Distributed Architectures for Geo-spatial Information Management

A second challenge concerns the development of distributed architectures for the management of geo-spatial information. Because the processing of spatial data and knowledge is complex and expensive, many computational functions that once were centralized in a GIS can now be made available in form of Web Services, thus as programs that can be accessed by clients through the Web. Web services may consist of simple functions such as those performing coordinate conversion or more complex ones such those simulating spatial phenomena. At the same time, the development of portable devices has encouraged the development of spatially-aware applications for mobile users. Therefore mobile users can access Web services and download on portable devices geo-spatial information that can be further managed locally. In all these cases, however, the spatial information mainly flows from the centre, i.e. the server, towards the periphery, i.e. the client. A more innovative distributed architecture in which the information, being collected remotely, flows instead from the periphery to the centre is represented by geo-sensor networks. Geo-sensors, such as GPS-based devices, biosensors, cameras on board of static or mobile platforms typically acquire streams of geo-spatial data. The aggregation of a number of sensor nodes into a computational structure forms a geo-sensor network. Raw data is thus collected by the nodes of the network and possibly aggregated and analyzed. Data aggregation and analysis can be locally executed in real-time by sensor nodes, or off-line in several distributed or centralized repositories [9]. Because this area of research is recent, there is no a common accepted framework. From a knowledge engineering perspective, however, relevant open research issues include modelling and reasoning with geo-sensors.

2.3 Spatial Data Warehouses and Spatial Data Mining

In the near future large amounts of raw spatial data will become available, especially in the form of sensor data. At the same time the complexity of the decision making processes, involving geo-spatial data and knowledge, will increase . It is thus reasonable to expect that technologies for spatial knowledge discovery will become more and more relevant. Spatial data warehousing and spatial data mining are two complementary techniques for knowledge discovery. Spatial data warehousing mainly focuses on

the development of extended and spatial-based multidimensional data models to support spatial data aggregation and user navigation. Spatial data aggregation is an operation providing a synthetic view of a geographical phenomenon, whereas navigation is a functionality allowing the user to interactively inspect and analyze data through a set of spatially aware operators (Spatial OLAP) [7]. In contrast, spatial data mining deals with the development of algorithms for the discovery of complex spatial knowledge, such as spatial clustering and spatial association rules. Because of the increasing number of mobile communication devices, a challenging knowledge discovery task is the extraction of knowledge in the form of space-time trajectories of these personal devices [8]. These trajectories contain detailed information about personal and vehicular mobile behaviour, and therefore offer interesting practical opportunities to find patterns to be used, for instance, in traffic and sustainable mobility management, e.g., to study the accessibility to services. A further research issue is the integration of multidimensional data modelling, OLAP and data mining in a comprehensive system able to deal with the different aspects of data modelling, user interaction and complex knowledge extraction.

3 An Application Scenario

Finally we would like to pose the question of which kinds of application domains are best representative of the next generation spatial applications and thus where the above technologies could be best exploited. We believe that the recent natural disasters should make researchers reflect about the relevance of issues such as natural risk and disaster management. Disasters like floods, storms, earthquakes, represent a serious disruption for the human society and can compromise the social and economic stability of several countries. To mitigate their effects, disaster management strategies are needed based on an interdisciplinary approach and accounting for organizational, social and technological aspects. The use of spatial technologies in disaster management is a growing field with limited published information to date. From a technological and an information engineering point of view, this domain offers several challenges. First, this is an application area in which many types of information have both a spatial and a temporal component. Moreover, relevant data are spread on the territory, and are acquired by many diverse technologies, including real-time sensor data from satellites and networks on the Earth. Further, in these applications the interaction with the available data spatial infrastructures is essential. The general architecture that can be envisaged is thus a highly distributed architecture enabling the cooperation among geo-sensor networks, infrastructures such as SDIs and decision support systems. For example, sensor data related with other types of data, such as mapping and cadastral information from different countries, may provide a database from which hazard maps can be generated, indicating which areas are potentially dangerous. Similarly, knowledge patterns extracted from raw spatial data can be used together with simulation models for environmental surveillance and monitoring purposes. Geo-ontologies instead would represent the semantic glue for connecting all the application resources in some metadata space.

4 Conclusions

In this paper which introduces the session on knowledge engineering techniques for spatio-temporal applications we have outlined current trends in geo-spatial applications for which the use of innovative knowledge tools and methodologies is crucial. We have also briefly discussed some relevant research issues; certainly our discussion is not exhaustive and many other relevant challenges can be identified, such as for example the development of security and privacy techniques for geo-spatial applications and location-aware services. We trust that the attendees will find many interesting ideas, concepts and research results in the papers presented in this session as well as challenging research directions.

References

1. Global Spatial Data Infrastructure: The SDI Cookbook, V2, Jan. 2004. http://www.gsdi.org/docs2004/Cookbook/cookbookV2.0.pdf
2. Bennett B.: What is a Forest? on the vagueness of certain geographic concepts. Journal Topoi, 20 (2), pp 189-201, 2001.
3. Commission of the European Communities: Proposal for a Directive of the European Parliament and of the Council establishing an infrastructure for spatial information in the Community, 23 Jul. 2004. http://inspire.jrc.it/proposal/COM_2004_0516_F_EN_ACTE.pdf
4. Duckham, M. and Worboys, M. F.: An algebraic approach to automated geospatial information fusion Accepted for publication in International Journal of Geographical Information Science, 2005
5. Kai L., Ludaescher B.: A System for Semantic Integration of Geologic Maps via Ontologies. In the Proc. of the Workshop on Semantic Web Technologies for Searching and Retrieving Scientific Data (SCISW), Sanibel Island, Florida, 2003
6. Kuhn W.: Semantic reference systems. International Journal of Geographical Information Science 17(5): 405-409, 2003
7. Rivest S., Bdard, Y. and Marchand P.: Towards better support for spatial decision-making: defining the characteristics of spatial on-line analytical processing (SOLAP) Geomatica, the Journal of the Canadian Institute of Geomatics, 55:2001, 539–555, 2001
8. Meratnia, N. and de By, R.: Aggregation and Comparison of Trajectories. In the Proc. ACM GIS'02
9. Nittel S., Stefanidis A., Cruz I., Egenhofer M., Goldin D., Howard A., Labrinidis A., Madden S., Voisard A. and Worboys M. : Report from the First Workshop on Geo Sensor Networks. SIGMOD record, Special Issue on Sensor Network Technology Infrastructure, Security, Data processing, and Deployment, Ed.Vijay Kumar, March, 2004
10. Spaccapietra S., Parent C., Cullot N., and Vangenot C.: On Spatial Ontologies. In the Proc.of GeoInfo 2004, VI Brazilian Symposium on GeoInformatics, Campos do Jordao, 2004