Spatial Information: Problems, Challenges and Directions

Peter A. Whigham
Spatial Information Research Centre
Department of Information Science
University of Otago
Dunedin
pwhigham@infoscience.otago.ac.nz

Abstract. In the spirit of David Hilbert's 1900 lecture, it would seem appropriate at the commencement of the 21st century to highlight some research areas, challenges, problems and possible directions that will be taken up by spatial scientists in the coming decades. The paper covers areas such as spatio-temporal data, data management, data models and representation, intelligent spatial systems, cognition, culture, education, mobile delivery and the world wide web. The paper does not attempt to describe how these changes will come about, nor does it go into detail with any one idea or challenge. The purpose here is to highlight the variety and diversity of fields that are part of the domain of spatial information science, and offer some thoughts on some current challenges.

1. INTRODUCTION

This short paper describes 23 areas of research and development that the author believes will be investigated during the 21st century. The problems are presented in no particular order or scale. Some concepts relate to broad fields or approaches that need extension, while others are specific concepts within a particular field. The intention is to demonstrate the diversity and variety of concepts within the field of spatial science, and to hopefully inspire current researchers to develop and extend these ideas.

2. RESEARCH CHALLENGES

2.1 Management of spatio-temporal data

Based on the passed few decades of the 20th century it is clear that the quantity and quality of spatial and spatio-temporal data is increasing, probably at an exponential rate. Current technology increases have managed to handle the increase in quantity in terms of data storage, however the techniques required to manage the data, whilst maintaining the spatial and temporal nature of the data, have been lacking. This goes back to the historical relationship between SIS and cartography, where the flat, static, 2 dimensional paper-based map still has some influence over the procedures and techniques of SIS. A clear challenge is to produce new management systems, which maintain, promote, visualise, access, represent and manage the increasingly large quantities of spatio-temporal data without loss of information.

2.2 Collection of spatio-temporal data

The future of data collection could be imagined as intelligent devices, like confetti, scattered over the landscape. Data collection devices are becoming cheaper, smaller, wireless, web-enabled, disposable, and have higher resolutions, reliability, and ease of use. These current trends imply that in the future data collection will eventually approach zero cost. The challenge is to begin developing techniques that allow this data to be visualised, managed and explored in real time. The vast quantities of data that will be collected imply that automatic procedures will be necessary to sieve through and initially describe and understand the data as it is collected.

2.3 Automatic interpretation of spatio-temporal data

The Topoclimate South project, a New Zealand based project exploring topographic microclimates, collects temperature data at 6-minute intervals, with approximately one sensor per 20 ha. Consider the amount of data generated for 48 sites (a typical single area of study):

Raw Data = 87 600 x 48 = 4.2 Million items per variable
Daily Data = 365 x 48 = 17 520
Annual Data = 1 x 48 = 48

The entire project has 2550 sites, giving approx. 223 million data items per variable. This quantity of data is both too large to study individually, and is even difficult to handle simply as a collection of relational database tables. The challenge is to develop techniques that allow the automatic detection of features defined as "interesting", within this ocean of data. This is not only the concept of spatial data mining, but must also allow the user to define patterns that they want to detect, in a language that is natural, expressive and maintains the spatio-temporal qualities of the problem. This will require a feature-based language for spatio-temporal data that allows the relationships between locations, attributes and time to be expressed.

2.4 Relational versus Propositional descriptions

Space is relational. Spatial relationships, topology and spatial correlation are essential to making spatial systems different from other forms of data. Current spatial systems allow these relationships to be explored in an incremental, step-by-step fashion, without any support for discovering useful patterns or relationships. The space is viewed in a very propositional and local manner. The challenge is to develop more relational views of space that incorporate the complex relationships that exist in the data, without deconstructing the space into simple propositions.

2.5 Passive Nature of SIS

Current SIS are passive. The data is presented in a map-like manner, the spatial features are static, change is represented as a series of layers (t, t+1, t+2, …) and the spatial features are essentially "dumb". The challenge is to integrate new technologies such as agent-based approaches that support automonous, dynamic, intelligent, active concepts into data representation and process modelling. For example, agents could represent each individual spatial feature, at any particular scale. Agents could construct themselves from the structure of initial agent features (a forest agent, or a tumour agent) based on the topology, clustering or other properties, both spatial and attribute, of simpler agents. This structuring of active, and interactive, components as fundamental spatial units will allow the previous "cartographic" framework to be extended into a more information-based and interactive view.

2.6 Intelligent Digitising

Current work has commenced on incorporating intelligence into basic data management and transformation tasks, using techniques such as Hidden Markov models to support intelligent digitising. Future work must be done to incorporate active, adaptive learning approaches to integrate the user behaviour, goals and approaches, into the basic tools of spatial systems. Techniques from artificial intelligence research, including evolving neural networks and other techniques that allow on-line learning and adaption must be integrated within the SIS framework to allow these systems to become more flexible and interactive.

2.7 Interface Design and interaction

Current user interaction with a SIS proceeds along the lines of most other forms of modelling or database system: a command is given in a text-based format (often SQL), some operation on a set of tables are performed, typically producing another set of tables, and these results are presented. Although a SIS can present the final tables in a spatial context, the user generally feels that the data being manipulated is similar to a database table. Some progress has been made towards other concepts of interaction, such as sketching spatial concepts rather then defining accurate geographically referenced data, although these attempts are still very much based on "map-based" concepts. Future SIS will be multi-modal: no one form of interface will dominate, and a graphical, visual and spatial interface will dominate the normal user's view of spatial systems. Spatial data is inherently "spatial", and computing-based technologies are now entering an age of visualisation and graphical exploration. Future SISs will have to take these computer science techniques and integrate them into the basic query, representation and interaction of their systems.

2.8 Spatial Similarity

Image-based similarity measures are currently a major research topic within the field of multi-media and image processing. These techniques are being developed to support image data mining and searches for complex image structures within large data sets. The use of similarity measures for exploring spatial information will become a fundamental technique with spatial systems. In particular, these approaches will be integrated with distributed approaches to data sharing and metadata technologies, allowing simple, descriptive and visual definitions of data requirements to be given.

2.9 Spatial Meta-data and Data sharing

Current data sharing and discovery systems are often based on direct mappings between meta-data categories, usually without taking into consideration the explicit spatial nature of the data being converted or compared. The continued promotion of data sharing will push the development of new approaches to data sharing and meta-data representation. For example, GML (geographical markup language, GIS consortium), based on the eXtensible Markup Language, is one approach to developing open-ended techniques for data description and data sharing. These and other approaches, in hand with distributed system development, will address many of the data sharing and data discovery problems that currently are being researched.

2.10 Mobile SIS

Hand-held technologies have allowed SIS's to move into the field for the collection, management and visualisation of spatial information in real time. The future will produce more mobile technologies that integrate the basic concept of "You are here" into devices. What will these systems look like? Current trends imply that they will not look like a computer. In the near future such concepts as the wearable computer, and intelligent devices, will integrate with mobile locational information to support intelligent, positional based, machines. The question will not be "Where am I now?", but "How many devices am I wearing?"

2.11 Web-based systems – Delivery and Collection of spatial information

The WWW is fast becoming the defacto information medium, and clearly the 21st century will build on this development. Spatial information delivery will become mobile, handheld, real-time and web-based. How the technologies handle this downsizing in transmission, and the modes of presentation that will be required as devices decrease in size, willl have to be addressed. Clearly there are issues with multi-modal forms of interactivity, including hands free, verbal interactions with spatial systems, and methods of expressing and manipulating spatial data, that will be required.

2.12 N-dimensional integration and scale

Fields such as genetic research require the exploration of patterns in many-dimensional space. Although SIS can be used in a multi-dimensional fashion, future systems will be designed with the original n-dimensional framework as a basic assumption. Additionally, spatial systems were originally designed with geographic scales in mind. This scale assumption will become less relevant and the systems will move towards generalized scale-free systems. This will allow spatial systems to be used in disciplines that currently require specialist software. For example, concepts such as protein folding or the relationship between base pairs in a DNA sequence are currently not explored using spatial systems, although they are clearly spatial in nature.

2.13 Virtual reality and Animation

Advances in hardware and visual image rendering will be integrated into spatial systems. The ability to explore space in an immersed, interactive environment will initially be used in fields such as architecture, art and design, although this approach will be expanded into all aspects of spatial information research and education.

2.14 Theories of Scale

Scale is currently studied as both a mathematical and cognitive concept. The concept of scale in a spatial system is currently viewed explicitly as a relationship between distance on a “map”, and distance in the “real” world, and implicitly as the objects represented by the digital representations. Future research will develop theories of how scale integrates generalization behaviour, and use this to produce systems that explicitly represent and understand the objects, scales, and their relationships within the system. This will allow more realistic and intelligent interaction between the system and user.

2.15 A Single Model of Space

There is currently a clear split between raster and vector representations, with separate tools, operations and database requirements. This distinction between data models will eventually be resolved, and the user will see a single, uniform representation of space. Although the underlying techniques may still handle objects or fields, the user will not be aware of these changes in representation while interacting with the system.

2.16 OpenGIS Systems

Currently within computer science there is a trend towards open systems architectures, where it is possible to integrate software components, written in various languages, into an existing system. This trend will continue within the spatial systems environment, driven by the extension of fields with a diverse range of models that incorporate spatial descriptions. For example, modelling spatiotemporal dynamics in ecology requires models such as home range models, habitat fragmentation, nonlinearities, scale variations, sparse and messy data, diffusion-driven patterns, and other spatial intra- and inter-actions between and within species. No SIS can offer all of these models, and therefore the open architectures of future SIS will be required if these systems are to be truly useful to a variety of fields.

2.17 Education

Spatial data, and the representation and interaction with spatial data, will become far easier in the future, due to delivery systems such as the web and the extension of data sharing and meta-data techniques. This will require developments in education and system structure to allow a more diverse range of skilled people to interact with the spatial data and information. For example, in computer science the language Logo, based on a turtle wandering around space, was developed to explain simple algorithmic concepts to children[1].

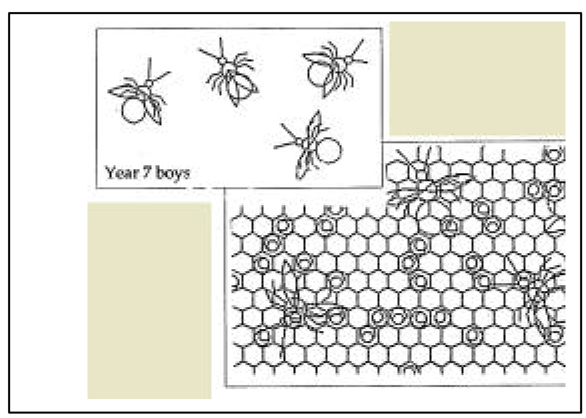

Figure 1. The Bee Dance – Logo Patterns

[1] http://www.ecu.edu.au/pa/ecawa/sig/logo/paul_dench/turtle/index.html

In a similar fashion, software that supports interactive and graphical means of learning spatial systems concepts will be developed. Based on these educational advances, spatial data will become one of the most relevant, visualised and manipulated forms of data during the 21st century.

2.18 Cognitive Aspects

The integration of positional information, based on GPS and mobile technologies, will drive a shift in the importance of spatial systems for wayfinding and how people view their use of space. Concepts such as space syntax, and how people interact with space, will integrate psychology, robotics and planning to produce spatial systems that represent how and why people interact within particular spatial domains. For example, sophisticated models of where people go when they are lost, in both urban and native environments, will be produced. These will be extended to become sophisticated models of behaviour that will support industries such as retail, marketing and tourism.

2.19 Cultural Aspects (Fuzzy Boundaries)

Land rights, land ownership and the rights of people have all become notable topics towards the end of the 20th century. Current concepts of land ownership, and the meaning of space in a cultural sense, will become even more prevalent during the 21st century. This will drive other representations of space that attempt to more realistically represent how people interact and respond to the space in which they live. The concept of Voronoi space is just one example where the boundaries between spatial elements are developed due to the objects that exist in the space. Other approaches will be developed, for example fuzzy Voronoi boundaries, and formally integrated into SIS. This will also help to move the concept of a “spatial system” away from “geographical maps” and more towards “mental maps” and “cognitive space”.

2.20 Global versus local pattern detections

Currently there are few tools that allow both global and local spatial patterns to be detected. Future systems will incorporate a language and a set of fundamental operations that can be combined to describe the properties of space. For example, topoclimate queries (i.e. climatic affects driven by topography) that a user may want to detect would include such effects as where does wind funnelling occur, evidence for valley winds produced due to warm slopes, cold air ponding at night and radiation loading. Clearly there is required to be a definition of how these areas can be detected, how the relationships between temperature and space are related, and how the varying spatial and temporal scales can be used to detect and control these patterns. These global and local patterns must be integrated into a common framework.

2.21 Spatiotemporal Logics

Currently there is a large body of work dealing with temporal logics, but few solid foundations with spatio-temporal concepts. The field of spatio-temporal logic will be developed as a stepping-stone for pattern detection algorithms, global versus local concepts, and the consistent handling of various scales and complexities for multi-dimensional data.

2.22 Precision Agriculture

An example application that relies on a spatial system is that of precision agriculture. Location, and the relationship between soils, geology, management practice, micro-climates and crop type are integrated within a spatial system. This allows tasks to be performed that would not otherwise be possible, producing more intelligent business decisions and better management. This type of market will continue to increase, as location becomes a fundamental attribute of many business systems.

2.23 Integrator of Technologies

Spatial systems have been previously labeled as the great integrators of data, due to the common foreign key of location. The ability to integrate disparate data will become more prevalent and allow spatial technologies to become one of the most common commercial backbones of the future.

3. CONCLUSION

Many of the concepts presented here are already in development, or are well recognised as important next developments. However, to be truly successful at many of these problems I believe that the research community must start by assuming the data exists (if it is a data problem), and commence building and exploring concepts before the hardware or delivery systems are in place.

Other developments require the integration of various techniques from other fields. This is not surprising – many scientific advances occur through the integration of disparate disciplines, or through the analogical use of ideas from one field to the next. The future SI scientist must be willing to be broad and deep (but focused) in their knowledge. It is no longer adequate to just be part of a single field, and these developments will occur by those with skills that cross boundaries.

It is the planning for the future expansion of these concepts that will drive our progress in both academic and commercial approaches. I look forward to the developments of spatial information systems in the next few decades.