The Problems of Representation

Bin Li

College of Science and Technology
Central Michigan University
Mount Pleasant, MI 48858
Tel: (517) 774-1165
Fax: (517) 774-2907
Email: bin.li@cmich.edu

ABSTRACT This paper explores the key problems in the development of GIS. It does so by considering GIS as a new and efficient way for representing the real world in the digital form. By using slope calculation as an example, the paper presents a simple model of computational geography that constitutes different levels of representations. It then expands the discussion on the following representational problems: spatial reference framework, non-Euclidean representations, attribute measurements, data model, software object model, computer languages, application software, visual representation, and interoperability.

KEY WORDS GIS, representations.

1. Introduction

In a recent reunion of college classmates, most of whom had been away from the academics for many years, I was asked to explain what GIS is and why it is bringing about fundamental changes in the way we study geographic phenomena and processes. I used the common lines of explanation to convince the curious group—that GIS is a computer system that collects, manages, analyzes, and displays geographic information; that space and time are the basic attributes of every phenomenon and GIS provides a framework to integrate the multiple dimensions of earth processes; that GIS is an instrumental device playing a role similar to that of microscopes in biological sciences, and so on. Although impressed by the technical appearance of GIS, my geography classmates, who were once unsatisfied with the lack of theoretical framework and methodological vigor in geography, were far from convinced with the theoretical implications of GIS—Is GIS just another *New Cloth of the Emperor*, or some new wine in the old bottle?

Fortunately, this was not the first time I encountered such inquiries. For several years, I had been trying to understand the reasons for the commercial success of GIS and the wide acceptance of GIS across academic disciplines. Interestingly, it was an unexpected encounter with Roman Numerals that opened up my mind on the significant implication of GIS in geosciences. One day, I was reading a book on the history of mathematics in which several earlier numerals were described. Different from a typical book on the history of mathematics, this one compares the algorithms for arithmetic operations with Arabic and Roman Numerals. I was intrigued by the complicated procedures with Roman Numerals. Below shows how a simple subtraction (CIV - XLVII), or (104 - 47), is calculated with paper and pencil:

Civ - xLvii = LLiv - xxxxvii
= (LL+iv)-(L-x+v+ii)
= L+L+iv - L+x-v-ii
= (L+L-L)+(x+iv-v-ii)
= L + (x+v-i-v-ii)
= L + (x-i-ii)
= L + (ix-ii)
= L + (viiii-ii)
= L + vii
= Lvii
= LVII

More than ten steps are required, a rather tedious algorithm compared to that with Arabic numbers, which only takes two steps. Multiplication and division are even more complicated (Heather; Luis; Rick). The knowledgeable readers may find quicker algorithms and remind us that the Roman did not use paper and pencil but abacus and other devices to do the calculations. It doesn't matter. It was the clear difference in arithmetic algorithms between numerals (e.g., Arabic vs. Roman) that helped me recognize the important problem of representation and problem solving.

It is clear how things are represented have profound impacts on how they are perceived, interpreted, and acted upon. An appropriate form of representation can facilitate more accurate interpretation and more efficient methods for problem solving. GIS provides a framework of digital representations of spatial phenomena and processes. With such a framework, spatial analytical operations used to be extremely tedious to perform, such as the calculation of slope and aspect, can now be done with simple algorithms in a very short period of time. Concepts and theories, such as spatial autoregressive modeling, that had never been tested in the real world situations can now be implemented with GIS and applied to solving practical problems. GIS can be considered as a set of new representations of spatial phenomena, on which lies the reasons for its success, for its great potentials, and for the directions for further development.

This line of argument seemed convincing to my geographer audience. Since then I have thought about more on the problem of representation. I soon realized it is a problem well explored by linguists, psycholinguists, and philosophers. I believe, however, it is a productive exercise to consider various issues on GIS and Digital Earth in the context of representation. This paper attempts to outline some of these thoughts.

2. GIS as Digital Representations of Spatial Phenomena

To demonstrate the significance of digital representations in solving geographic problems, let's consider a simple problem in surface analysis, slope calculation.

Slope measures the change of elevations on a surface. In its simplest form, slope is the ratio of "rise" and "run", i.e.,

$$Slope = \frac{dz}{dx}$$

where dz is the elevation difference and dx is the distance between the two points on the surface. Finding areas that have slopes of 30 degree or less requires calculations on a reasonable number of points on the surface. It is not hard to imagine how tedious it would be to perform such calculations by hand. An alternative strategy, called the *graphical template method*, uses slope templates to generate the slope map based on the contour map. Each slope template has contour lines with specific spacing corresponding to the range of slope (Muehrcke and Muehrcke, 1992, pp. 323). The results from such measure, however, are obvious rather crude.

We all know how easy it is to generate a slope map with a computer when the surface is represented as a grid (Burrough, 1998). Not only does the digital representation make it a trivial task to implement the rise/run formula, it also facilitates more accurate measurement of slopes. For example, we are now able to define and calculate slope as the maximum rate of change in altitude and further as the derivative of the surface plane z = a + bx + cy, where z is the surface grid, x and y are the variables in the Cartesian coordinate system, and a, b, and c are the coefficients. As a result, slope has a more appropriate but certainly more complex form, called the third-order finite difference estimator,

$\sqrt{b^2 + c^2}$, where,

$$b = \frac{\left(z_{i-1,j+1} + 2z_{i,j+1} + z_{i+1,j+1}\right) - \left(z_{i-1,j-1} + 2z_{i,j-1} + z_{i+1,j-1}\right)}{8S};$$

$$c = \frac{\left(z_{i+1,j-1} + 2z_{i+1,j} + z_{i+1,j+1}\right) - \left(z_{i-1,j-1} + 2z_{i-1,j} + z_{i-1,j+1}\right)}{8S};$$

with *S* as the grid spacing (Chrisman, 1997, pp. 166). For human to calculate cell-by-cell this form of slope is clearly unfeasible. But it is a rather trivial task for the digital computers.

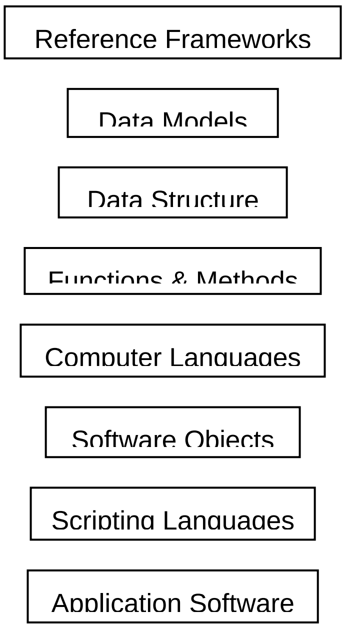
Similar to the relations between arithmetic algorithms and numerals, how the surface is represented affects the types and algorithms of the measures to be performed with surfaces. Contours can only accommodate the simplest slope calculation. When performed manually with slope templates, slope can only be a very rough estimate, resulting in limited utilities. On the contrast, grid, or raster representation of surface facilitates more advanced conceptualization of slope and makes it easy to implement with digital computers. The resulting slope maps are more accurate thus more useful in a variety of decision-making processes.

3. Computational Model

Slope is a very simple surface measure. The comparison of the analog and the digital representations, however, has general significance in our discussion on the problem of representation in GIS. To continue the exploration, let's examine the computational process with which slope is obtained.

1. We first have to determine the elevation and location of the sampling points on the earth surface. To do that, we must establish the basic measurement of the earth and the datum for horizontal and vertical positioning, i.e., the spatial reference framework. Further, when planar coordinate system is to be used, projection must be performed.
2. Next, we need to generate the digital representation of the earth surface. Either vector or raster data models can be used. The Triangulated Irregular Network, or TIN, is the common vector model for surface representation. The raster counter part is regular tessellation, or GRID. In

either case, interpolation needs to be performed to generate the digital surface.



3. Depending on the data models, we can now decide on the numerical representation (formula) to use. Assuming we are using the raster model, then we will use the *third-order finite difference estimator* for measuring slope.
4. The numerical model is implemented with a computer language and becomes a software object, which contains the abstract representation of the earth surface and corresponding attributes and functions (methods).
5. Through different levels of interface, we can now activate the software object and perform the slope calculation. The result is a slope map with different gray values representing classes of slopes.

The process of slope measurement highlights the basic layers of digital representations in GIS, as depicted in figure on the left. The integration of these layers form the computational model of geographic phenomena. Through these layers, we can put many of the research and development issues in perspectives.

4. Problems of Representation in GIS

4.1. Spatial reference framework

It is absolutely important to develop reference systems for accurate spatial measurements. The recent advances in Global Positioning System (GPS) have made surface positing highly accurate and convenient. Nevertheless, the accuracy eventually is determined by the ground control networks, or datum. Although many countries have developed rather accurate datum, improving and maintaining them remain a challenge.

More works, perhaps, are needed in developing localized planar coordinate systems. Planar coordinate systems make geometric calculations simpler but transforming the globe to the plane results in distortions. The larger the plane, the greater the distortions. One of the solutions is to develop localized planar coordinate systems, similar to the US State Plane Coordinate System. Currently, the Universal Transverse Mercator (UTM) system serves as the local plane coordinate system around the world. The problem with UTM is that the UTM zones seldom coincide with the area of interest, e.g., administrative regions. For many provinces in China, for example, it requires more than one UTM zones to cover, between which there is no simple mathematical relationship for the coordinates. Localized planar coordinate systems, therefore, should be developed to facilitate more accurate spatial referencing.

4.2. Non-Euclidean Representations

Geometric space is among the many ways for representing the world. Alternative representations can reveal characteristics that may otherwise be difficult to recognize. A commonly used non-Euclidean representation is topological space that emphasizes on connectivity and contiguity among objects. Frequency is another domain of alternative representations. A grid (raster), for example, can be represented as a Fourier series through transformation. The resulting Fourier series can be used to extract information directly that uncovers the cyclical characteristics. It can also be used as an alternative representation for the spatial autoregressive model (Griffith, 1988). Another similar transformation, wavelets, has showed great potentials in abstracting and compressing data. Recently, we have begun exploring the applications of the Kohonent Self-Organizing Map (SOM), a neural network model that is able to reduce multidimensional attribute space to a two-dimensional regular grid. The resulting SOM allocates objects into clusters that share similar attribute characteristics. Like the Fourier power spectrum, SOM can be visualized as density image and linked with its corresponding geometric representation, the cartographic map (Li, 1998).

A new form of representing the world facilitates the development of new perspectives. Different representations constitute the all possible worlds. Integrating them helps understanding the complex nature of the world objects and phenomena. Developing and integrating new form of representation are important issues in GIS.

4.3. Measurements

Measurement, either quantitative or qualitative, is another important aspect of representation. New measurements reveal attributes that are otherwise not recognized. Measuring the emission, reflection, and abortion of electro-magnetic radiation help us recognize the nature and state of objects very efficiently. This new measurement serves as the very foundation of remote sensing. The importance of measurement in the history of science and technology needs not further justification. The challenge is in identifying the areas where new measurements and corresponding instruments needed to be developed.

There are two broad groups of measurements. One group measures the physical world, e.g., the elevation, the reflectance in a particular band, and the water contents of crops. Advances in air-borne sensors systems and GPS represent the state-of-the-art of measurement and development direction in the physical domain. In the social, cultural, and economic domains, informative measurements are not easy to develop and obtain. Population census alone has been an extremely challenging matter. Various social economic indices are far from providing accurate and thorough information. New sampling methods, survey techniques, and new indices need to be developed.

No measurement is error free. It is important to be able to measure the errors and incorporate them in the process of analysis and interpretation.

4.4. Data Model

Data model is the logical representation of real world objects and their relations. There are two broad aspects in a data model. It must include data objects that correspond to the real world objects. It must also use operators to represent the object behavior and relationship between objects. How real world objects are represented has crucial implications on the capability and efficiency of the computational model.

The vector and raster data models have been serving a large number of applications quite well. The current challenge is in developing and implementing adequate data models for three-dimensional space. Hydrological, geological, and atmospheric processes are more suitable to be represented in three-dimensional spaces. Due to the lack of software support, most of the 3-D applications must be done in the 2-D environment. Transforming spaces from higher to lower dimensions often results in distortion and loss of information as well as inefficient algorithms.

The OpenGIS Consortium has proposed detailed definitions of geometric features in dimensions 0, 1, 2, and 3 (OpenGIS, 1999). The consortium indicated that representation of three-D features needs further research. A fully developed data model for three-D objects is a clear challenge for the academics and the industry.

Aside from the representation of three-dimensional objects, we should have noted the recent evolution from the entity-relations (ER) model to the object-oriented models. ESRI, in its most recent release of ARCINFO, claimed that it had adapted an object-based data model for its flagship product. The database software industry is also moving towards object-oriented systems. Unlike modeling the real world using generic geometric objects such as point, line, and polygons, abstracting the real world using high level objects such as lake, county, and highway may be more flexible and natural but may require far more sophistical data model to accommodate the customized representation (Maguire, 1999).

Another important aspect of data models lies in the understanding and conceptualization of object behavior and relationships. This is traditionally the responsibility of academic research. The measurement of slope, for example, is improved by introducing a mathematical representation, the best-fit plane, which derives the more appropriate analytical expression of slope. As another example, optimal spatial interpolation is centered upon the concept of semi-variogram that reveals the relations between the distribution of surface values and distance. In all, improvement of the data models lies in the basic understanding of the behavior and relations among real world objects.

4.5. Domain specific data models

Domain specific data models build on the basic data models to accommodate specific needs. TIGER, or Topologically Integrated Geographic Encoding and References, is an example of this kind. It is designed primary for storing population census data and is based on the topological and entity-relation data models. TIGER has been proven to be a good representation of the urban landscape in the United States.

The lack of an equivalent TIGER in China may have contributed to the slow development of GIS applications in urban and business applications. Further delay in developing a national standard for representing spatial information in the rapidly evolving urban system will result in redundancy of development effort and incompatibility among databases.

Similar to the urban and business domains, other application domains such as soil, hydrology, transportation, and utility, may also require domain specific data models. Ideally, domain specific data models would be developed on the basic data models.

4.6. Software Object Model

Software object models transform data models into components and architectures suitable for computer operations. At the lower level, component object models provide the framework on how objects should be constructed and interact with each other. There are two component object models, i.e., the Distributed Component Object Model (DCOM) by Microsoft and the Common Object Request Broker Architecture (CORBA) by the Object Management Group. Both have become the frameworks for developing application software.

The maturity of the component object models has greatly facilitated the development of software object models for such domain applications as GIS and multispectral image processing. ESRI, MapInfo, and Integraph all have incorporated the component object models in the design of their new products. The strength and potential of these products will largely depend on the quality of the corresponding software object models.

Component object approach in software development is new to the GIS software industry. Most of the software object models designed and implemented are rather primitive. Some take a transitional approach and simply repackage existing function libraries with object wrappers, which does not match well with the underneath component object models, be it CORBA or DCOM. Others attempted to build and implement a new object model from scratch but the amount of work and the problems needed to solve turn out to be overwhelming. Still there are developers who continue to build software models that are closed and not able to interact with other systems well. In all, a good software object model is crucial for problem solving and there are various challenging issues to be addressed (Li, 1999; Li and Zhang, 1998).

4.7. Computer Languages

Computer languages by themselves are models of representations. They facilitate the mapping between software object models and the computer hardware systems. Not all computer languages were created equal. In the past, many of them served particular domains of applications. For instance, FORTRAN was designed for scientific calculations; COBAL was for business applications; and C was for system development. There were even special languages designed for image processing and for computer graphics. This is no longer the case. With drastic improvement in development environments, languages that facilitate more flexible interface with the operating system and the hardware, such as C, C++, and JAVA, have become the dominating languages for software development. At the same time, high-level languages such as Visual Basic, Visual Basic Script, and Java Script, are replacing application specific macros and scripting languages. It can be panticipated the graduate diminishing of such GIS languages as AML, Map Basic, and Avenue.

From the perspective of representation, there are trade-offs for substituting application specific languages with general purpose ones. Clearly, a general-purpose language such as Visual Basic does not provide a straightforward mapping to the application software object model as a specifically designed language would do. For example, one would feel AML and Avenue more convenient in expressing GIS operations than Visual Basic. On the other hand, designing, implementing, and maintaining a macro language are expensive. Making it to be compatible with other application software is even more costly and difficult. With the support of component object architectures and the wide availability of high level languages, there is little incentive for venders to develop domain specific languages. Such drawback can be overcome by a well design user interface in the application software, which is made easy by the high-level general-purpose languages.

Today, most of the computer languages implement the basic concepts of the Tuning Machine, which operates in a sequential manner. Many natural phenomena as well as aspects of human thinking have sequential characteristics, which maps well onto sequential representations. There are however equally important portions of the real world processes that are not sequential but parallel. Human vision, for example, is largely a parallel system. The real world clearly operates in a combination of sequential and parallel processes. Using sequential representations alone would force two or higher dimensional processes into a single dimension, i.e., time.

Take image processing as an example. A multispectral image represents the electro-magnetic characteristics of ground objects with a two dimensional grid. Such a grid is typically represented as a two dimensional array in a computer language, which corresponds to the memory slots in the hardware. When performing operations on the two-dimensional array, a sequential language would require a double loop to process one pixel after another, sequentially. Such a transformation from space to time has consequences in computational efficiency and problem solving. In sequential representation, operations typically are a function of the number of the data element. The more data elements, the more steps it would require. Further, since sequential representation may contradict with human intuition, our ability to devise better algorithms may be impeded by the sequential language. It is clear that parallel representations are superior in many spatial operations (Li, 1993; Fox, 1988). The seemingly lack of research and development activities in parallel computing, particularly in the spatial sciences, should not lessen the effort in pursuing better modes of representations at the programming language levels.

4.8. Application Software

Application software is the final level of digital representations at which human directly interact with the digital model of the real world. At this level, the key is in user interface design. A good user interface should facilitate an efficient mapping of the user's view and behavior onto the software object model through interface elements. On contrast, a poor user interface design typically reflects the software engineer's view and behavior, resulting in software packages that are "unfriendly" and difficult to use. No matter how advanced the underneath data model and software object model, and how efficient those are implemented, the computational process will be impeded without a good user interface.

4.9. Visual Representation

Visual representation turns digital data into graphic symbols that can be perceived by the human vision system, which is far superior to computers in synthesizing and extracting visual information. Geography has a long tradition in portraying the world as pictures. The entire field of Cartography, a sub-discipline of geography, is devoted to the visual representation of earth objects and phenomena. With the development GIS and other Information Technology, Cartography has evolved rapidly and has made some important progresses.

The most important changes are in the transitions from static to dynamic representation and from a means of communication to a central tool for exploration and information discovery (Kraak, 1998). Dynamic Cartography, or, Exploratory Cartography promises to facilitate the graphic representation of time and space. Time series no longer have to be represented as small multiples but as user-controlled animated maps, which are more efficient for human to perceive temporal variations. Map users should no longer be passive receivers of cartographic information but in full control over what, when, and how to view the information. To realize the ideals of Exploratory Cartography, there is much to accomplish. Below is just a subset of the problems needed to be addressed:

- Visual variables are at the core of cartographic symbolization. How can existing visual variables be extended to accommodate the representation of dynamic information?
- In traditional cartography, when map objects are all arranged on a single sheet, cartographic design needs only focus on the relations among these objects. In Dynamic Cartography, when more than one map need to be organized efficiently, the focus of design has a new and perhaps more important dimension—how the maps should be structured and what types of navigational controls must be provided?
- Turning data into graphics alone does not help a great deal in revealing patterns, anomalies, and other useful information embedded in the data. New methods of data mining need to be developed and integrated with cartographic systems.

4.10. Interoperability

The essence of interoperability is the ability to map one representation to another accurately and efficiently. When a task has to use multiple models at the same level of representation, there are needs for interoperation. Data format, for example, is a kind of digital representation that can vary among software programs. Format conversion is a familiar operation for most of us. Obviously, the need for interoperations is not limited to data format. Interoperability is such a crucial problem that it requires an industrial consortium to tackle—the OpenGIS Consortium was formed to address the problem of interoperability.

The success of ODBC in providing database connectivity has led to the thinking that the key to interoperability is to define a universal representation. This is the strategy OpenGIS adapted. It hopes that by providing standard definitions for the geometric objects and interfaces, interoperation can be achieved. The problem with such strategy lies in the assumption that somehow one can find a universal representation that can encompass all other types of representations. Interoperation is then a matter of first transforming the source model to the universal model, which in turn converts to the destination model. This approach may work for certain period of time. But as new models developed, the universal model will become obsolete because its construction is likely based on the generalization and abstraction of existing models at a particular period of time. One may argue further the validity of universal representations from deeper theoretical grounds or simply anticipate complications similar to those attempts in developing a universal human language, Esperanto.

5. Conclusion

The potential and challenges of GIS can be analyzed through the perspectives of representation. The success of GIS has been largely due to its effectiveness in facilitating digital representations. The limits and obstacles in the development path of GIS also exist in every level of representations. New and better representations will bring about breakthroughs in GIS. This paper has outlined some of the research agendas centered at the problem of representations. They are more or less reflections of attempts to systematically understand the theoretical aspects of GIS.

References

Burrough, P. A. and McDonnell, R. A., 1998, Principles of Geographical Information Systems, New York: Oxford University Press.

Fox, G., 1988, Solving Problems in Concurrent Computers, General Techniques and Regular Problems, Vol. 1, Englewood Cliffs, NJ: Prentice-Hall.

Griffith, D., 1988, Advanced Spatial Statistics, Dordrecht: Kluwer Academic Publishers.

Heather, *Addition and Multiplication with Roman Numbers*, Ask Dr. Math, http://forum.swarthmore.edu/dr. math/problems/holtby27.html. (Note: The full name for the author is not available.)

Kraak, M. J., 1998, *Exploratory Cartography, Maps as Tools for Discovery*, Inaugural address, Division of Geoinformatics, Cartography and Visualization International Institute for Aerospace Survey and Earth Sciences (ITC), The Netherlands, http://www.itc.nl/~carto/kraak/

Li, Bin, 1992. "Opportunities and Challenges of Parallel Processing in Spatial Data Analysis: Initial Experiments with Data Parallel Map Analysis," *GIS/LIS'92*, Bethesda, Maryland: American Society for Photogrammetry and Remote Sensing, pp. 445-458.

Li, Bin, 1998. "Exploring Spatial and Temporal Patterns with Self-Organizing Maps," *GIS/LIS'98,* Bethesda, Maryland: American Society for Photogrammetry and Remote Sensing, pp. 89-96.

Li, Bin, 1999. "A component perspective on geographic information services," *Cartography and Geographic Information Systems*, to appear in the October Issue.

Li, Bin and Zhang, Li, 1997. "A model of component-oriented GIS," *GIS/LIS'97,* Bethesda, Maryland: American Society for Photogrammetry and Remote Sensing, pp. 523-528.

Luis, 1999, *Adding and Subtracting Roman Numerals*, Ask Dr. Math, http://forum.swarthmore.edu/dr.math/problems/meghan10.7.97.html. (Note: The full name for the author is not available.)

Maguire, David J., 1999, "Object-Component GIS," ARCNEWS, Winter 98/99, Redland, CA: ESRI, http://www.esri.com/news/arcnews/winter9899articles/02-arcinfov8.html.

Muehrcke, P.C. and Muehrcke, J. O., 1992, *Map Use, Reading, Analysis, Interpretation*, third edition, Wisconsin, Madison: JP Publications.

Rick, 1999, *Subtracting Roman Numerals*, Ask Dr. Math, http://forum.swarthmore.edu/dr.math/problems/heather03.14.99.html. (Note: The full name for the author is not available.)

OpenGIS Consortium, 1999, *OpenGIS Specifications*, a series of technical papers published by the OpenGIS Consortium, available at http://www.opengis.org/techno/specs.htm.