

UC Santa Barbara

NCGIA Technical Reports

Title

NCGIA 18 Month Report (90-7)

Permalink

<https://escholarship.org/uc/item/2sq620dc>

Author

National Center for Geographic Information and Analysis (NCGIA)

Publication Date

1990-04-01

NCGIA

National Center for Geographic Information and Analysis

NCGIA

NATIONAL CENTER FOR GEOGRAPHIC INFORMATION & ANALYSIS

18 MONTH REPORT

Technical Report 90-7
April 1990

Simonett Center for Spatial Analysis
University of California
35 10 Phelps Hall
Santa Barbara, CA 93106-4060
Office (805) 893-8224
Fax (805) 893-8617
ncgia@ncgia.ucsb.edu

State University of New York
301 Wilkeson Quad, Box 610023
Buffalo NY 14261-0001
Office (716) 645-2545
Fax (716) 645-5957
ncgia@ubvms.cc.buffalo.edu

University of Maine
348 Boardman Hall
Orono ME 04469-5711
Office (207) 581-2149
Fax (207) 581-2206
ncgia@spatial.maine.edu

Preface

This document contains material requested by NSF in support of its 18-month review of the progress of the National Center for Geographic Information and Analysis (NCGIA). It is organized in three sections as follows:

- A. Narrative summaries of progress under each of the Center's Research Initiatives. For each of the initiatives currently under way, the document includes a 4-6 page discussion of progress and results, and a listing of publications with short abstracts.
- B. Narrative summary of progress in education.
- C. Narrative summary of progress in outreach.

The Annual Report for the year ended November 30, 1989 provided detailed listings of activities in outreach and education; these listings are merely summarized in this document.

Copies of the publications listed in Section A have been supplied with this document.

Background

On August 19, 1988, the National Science Foundation (NSF) awarded the NCGIA to a consortium of the University of California, Santa Barbara, the State University of New York at Buffalo, and the University of Maine, with funding of \$1.1 million a year for five years. The first year's operation officially began December 1, 1988, with funding effective as of that date. The decision to establish the Center and the selection process have been described by Abler (*International Journal of GIS* 1:303-26, 1987).

NSF's solicitation for the Center in 1987 identified "basic research on geographic analysis utilizing GIS" as the Center's primary mission and suggested five areas as possible research topics: improved methods of spatial analysis and advances in spatial statistics; a general theory of spatial relationships and data base structures; artificial intelligence and expert systems relevant to the development of geographic information systems; visualization research pertaining to the display and use of spatial data; and social, economic, and institutional issues arising from the use of GIS technology.

In addition to research, the Center was to take steps to "augment the nation's supply of experts in GIS and geographic analysis in participating disciplines; promote the diffusion of analysis based on GIS throughout the scientific community; and provide a central clearinghouse for disseminating information regarding research, teaching, and applications."

Research at NCGIA

The consortium proposed a research plan built around the concept of "impediments": that GIS technology had a significant potential, but that numerous roadblocks needed to be overcome. Some obstacles were technical, such as the unfriendliness of many user interfaces; some were institutional, such as the lack of standards for digital geographical data; and others were social and economic, including the difficulty of assessing costs and benefits of GIS, and lack of understanding of the impacts of GIS technology on organizations.

Research at the Center is organized around the systematic removal of these impediments, within the five areas suggested in the solicitation document. The full text of the research plan has been published in the *International Journal of GIS* (3:117-36, 1989).

Research takes place within the framework of a series of Research Initiatives. Each Initiative begins with a Specialist Meeting, attended by professionals from outside the Center, in which the most important problems are identified and ranked and a feasible research agenda for the initiative is defined. Research continues intensively for 18-24 months, with teams of faculty (NCGIA or other), postdoctoral fellows, or advanced graduate students, as well as representatives from private industry or governmental agencies, working in teams on specific problems. The Initiative ends with a national forum to present all of the research results. Results are also announced in articles in refereed journals, presentations at conferences, bibliographies, algorithms or models for analysis, NCGIA Technical Papers, and short courses or workshops.

1. Initiative 1: Accuracy of Spatial Databases

1.1 Narrative summary of progress to date

GIS processes operate with high precision, but frequently must process data of strictly limited accuracy. On the other hand conventional, manual methods of spatial data analysis operate with a precision which is compatible with the inherent accuracy of the data. In essence, GIS features such as scale change, and the ability to overlay data from different sources and of different quality are often regarded as the technology's most significant advantages, but in practice may represent significant problems. During the proposal-writing stage, data errors and uncertainties were identified as one of the most important impediments to the successful implementation of GIS. The topic of spatial database accuracy was given highest priority as the first Center initiative.

The specialist meeting for this first research initiative of the NCGIA was held at La Casa de Maria in Montecito, CA from December 13 through 16, 1988, to lay out the initiative's specific research agenda. Participants were drawn from the three Center sites, from universities in North America and Europe, and from numerous federal agencies and companies active in the GIS field. Disciplines represented included Geography, Mathematics, Statistics and several branches of Engineering.

The position papers of the meeting have since been published as an edited volume. The following sections discuss each of the research issues identified at the meeting and the research progress achieved in the sixteen months since the meeting. In many cases the research is still continuing, and much remains to be published. In other cases there are clear needs for research which the Center has been unable to fulfil because of limited resources.

Current plans call for the initiative to be formally terminated by a series of presentations at GIS/LIS '90 in Anaheim in October. However many of the results will be directly relevant to the research agenda of Initiative 7, Visualization of Spatial Data Uncertainty, which will begin late in 1990 or early 1991. In particular we hope that the models and prototypes developed in I1 will form the basis for methods of display to be developed in I7.

1.1.1 Data structures and models

This research theme has concerned the development of models of spatial data and database structures which are sensitive to data accuracy, can represent accuracy explicitly and can support the tracking of error through spatial database operations. Many object-based models have no such sensitivity to data accuracy, resulting in numerous forms of artifacts. Two papers at the Specialist Meeting directly addressed this issue by recommending a finite resolution approach in which the precision of computation is adjusted to the accuracy of the data. Geoffrey Dutton in particular discussed the importance of this approach for global data, where accuracy is highly variable, and where no suitable finite element schemes existed.

Yang Shiren and Michael Goodchild began research on global tessellation schemes in January 1989 at Santa Barbara. Working with a modified version of Dutton's proposal, they have succeeded in developing a finite element scheme for the globe based on projection onto an octahedron, followed by recursive subdivision of triangular facets. The scheme was included in EPA's recent review of possible sampling schemes for the EMAP project, as it provides a method of achieving approximately uniform sampling over a spherical surface. Research to date at Santa Barbara has determined the mathematics of the necessary projections, and algorithms for such basic functions as conversion to and from latitude/longitude, chain coding of lines, dilation and region filling. Yang is currently at Boston University and continuing this work with Dutton until July 1990, when he will return to Santa Barbara. Two graduate students, Shu Shurong and Sun Guoqing, are currently moving the algorithms to the IBM RS/6000 workstation, as the first step in implementing them in 3D display technology. A proposal for a two-year project to develop a prototype global GIS workstation based on this scheme has been submitted to NASA with Goodchild as PI. The basic details of the scheme were published as a Center Technical Report and a paper was submitted to *Computer Vision, Graphics and Image Processing* in December 1989. Two other papers on the algorithms are in draft form.

GIS provides a number of ways of converting real geographical variation into finite, discrete digital representations. This issue of data modeling is critical from an accuracy perspective, as any model of uncertainty must be specific to data model, and thus can inform the choice. Since December 1989 the Santa Barbara site has held regular seminars on the conceptual design of the "Next Generation GIS", in part through its joint study agreement with IBM. A paper by Goodchild, developing a perspective on GIS data modeling consistent with the findings of Initiative 1, has been presented at a seminar on Data Modeling sponsored by the Canadian Inter Agency Committee on Geomatics in Ottawa in March 1990, and at the GIS Design Models Conference sponsored by the Midlands Regional Research Laboratory in Leicester, UK, also in March 1990, and submitted to *Computers and Geosciences*.

1.1.2 Models of error and distortion

In this area we have concentrated on developing models which can successfully describe, characterize and parametrize error, both for spatial fields and for complex spatial objects. Several methods of simulation have been developed for modeling error in different types of spatial data. In essence, each assumes that uncertainty is described by a raster in which each pixel carries not a single deterministic class, but a vector of probabilities of belonging to each of a set of classes. We have explored methods for generating a population of distorted "maps" from such input, under two constraints. First, the proportion of occurrences of each class in a pixel across realizations must be as input; second, the outcome within one realization must be spatially dependent. A method based on a spatially autoregressive process has been programmed by Sun Guoqing at Santa Barbara, and implemented using a grant of computing time from IBM on the 3090 facility at Palo Alto Research Center. A second method based on swapping values between realizations was implemented by Fang Chiuwen. We have worked out the conceptual design of a GIS capable of encoding the level of uncertainty in each of its layers, and tracking the propagation of uncertainty through GIS operations into confidence limits on GIS products. The results of this work have been described in several presentations, but not yet submitted for publication.

These methods are defined entirely within the field domain. We have concluded based on this work that in general it is feasible to build models of uncertainty in the field domain, but much more difficult in the object domain. Models of uncertainty for objects are more readily constructed by distorting underlying fields from which the objects have been derived, and we have devised methods for doing this. This serves to point up an interesting and fundamental distinction between Initiatives 1 and 2. Whereas the field domain is more significant in I1, I2 has dealt almost entirely with cognitive views of a space populated by objects. Indeed, our vocabularies for describing continuous variation in fields are quite limited when compared to those for describing relationships between objects. But whereas the object view is clearly implied by human perception and cognition, the field view is preferred for describing uncertainty and for modeling many physical processes.

At Buffalo, David Mark and Ferenc Csillag have investigated the nature of boundaries on area-class maps, and have argued in a paper published in *Cartographica* that the process of generalization of such boundaries requires a continuous view of space, using surfaces representing probabilities of class membership.

At Maine, David Pullar is doing a doctoral thesis on the computation of the geometrical intersection of two or more polygon-nets having imprecise numerical data. Imprecision in the numerical data reflects uncertainty in the geometry of objects, and it is measured by a variance parameter. A variance is incorporated into the geometrical intersection procedure to prevent slivers, gaps, and other artifacts. The results from completed work are described in three papers. First, a comparative study was carried out on the complexity and performance of algorithms for reporting geometrical intersections. A fast pragmatic algorithm is proposed under the conditions that line segments are randomly distributed over the subject space. Second, an algorithm was developed to decide if two line segments with imprecise numeral data, cross, nearly cross, or are separate. It is referred to as a "fuzzy-intersection" algorithm. Third, the previous work is extended to treat intersections among arrangements of polylines. An algorithm for computing the geometrical intersection of a number of polylines, with associated variance parameters, is proposed and is called "fuzzy-chain-intersection". The algorithm uses clustering techniques to: i) resolve positional ambiguities, and ii) produce a valid planar topology within specified tolerance bounds. Future work will describe the way geometrical variance propagates through the process of fuzzy-chain-intersection.

1.1.3 Error propagation, product uncertainty and sensitivity

GIS processes combine data from different sources with different levels of spatial resolution, using rules which are often complex. There is a growing interest in modeling using such diverse spatial datasets, and in many cases models are highly nonlinear, leading to error effects which are difficult to control. Initiative research in this area has concentrated on understanding how error propagates through each GIS process, particularly overlay. Lodwick described existing work on this topic at the Specialist Meeting and has continued this line of research, in part with funding provided by the Center. Funding was obtained from USGS to support work on this topic at Santa Barbara in the Fall of 1989 by Giuseppe Arbia (University of Rome) and Robert Haining (University of Sheffield). Arbia and Haining have developed a general linear model which includes numerous forms of spatial data error as special cases, and provides a rich framework for analysis and simulation. The first set of results has been described in a paper submitted to a statistics journal.

Propagation effects may be seen in terms of sensitivity, or the relationship between the product or output of a GIS process and uncertainty in the corresponding inputs. At the Buffalo site, Rajan Batta has addressed the relationship between data error and the results of spatial decision-making, particularly in the context of location. One way to represent uncertainty is through the use of a stochastic location model, and objective functions which explicitly recognize uncertainty. Another is to look at the database as uncertain, and to examine the propagation of database uncertainty into the results of optimization. The references to Batta's work below include both of these approaches. Also at Buffalo, Rogerson has researched the link between region size and shape, and the

migration process, particularly in the estimation of migration distances.

1.1.4 Accuracy and risk

Criteria for acceptable levels of uncertainty in GIS products must be obtained ultimately from an analysis of the risks associated with decisions based on those products. Risk analysis is therefore an important link in the chain which stretches from error models and database concerns through to decision-making. Although analysis of risk is an accepted part of decision theory, it has not yet been applied to GIS in any systematic way. This research area forms a link between this initiative on spatial database accuracy and Initiative 4 on the use and value of spatial information. We hope that the error models developed as part of this research effort will ultimately lead to a better-informed approach to risk. Moving away from basic toward applied research, there is a need to understand better the concerns of GIS users, data providers and land management agencies for data accuracy. What minimal standards for accuracy are being developed by agencies, what methods are being used to define and document data quality, and what interactions might be developed between these concerns and basic research on error modeling? There is a need to develop standard measures of quality which are compatible with error models, and can be determined and monitored at reasonable cost for standard data types. It would be useful to have standard benchmark datasets which could be used to measure the accuracy of various data entry processes. Work by Amrhein at the Buffalo site has focused on the specification of accuracy requirements for data collecting agencies. Two reports sponsored by the Center and aimed at increasing awareness of accuracy issues have been written by Howard Veregin and published in the Center's technical report series: a bibliography of accuracy literature and an associated taxonomy of errors. We are currently seeking a publisher for a monograph based on these two reports.

Once risk is measurable, it is necessary to develop strategies which deal with it appropriately, and incorporate it in decision-making. Work at the Buffalo site by Batta has centered on location models which explicitly recognize risk, and distribute it equitably over a dispersed population. These have been applied to the specific cases of liquified-gas transportation, and the location of mobile ambulance units.

1.1.5 Experimentation and measurement

There is an important area for research in the development of methods for measuring accuracy empirically. This includes methods for measuring the quality of digitizing and scanning, relative to source documents, as well as standard techniques for measuring accuracy of databases with respect to ground truth. Little is known about the design of efficient sampling schemes. Measures of accuracy developed from experiment should be designed to be easily interpreted. It would be useful to have software, which might be simple additions to standard GIS products, which could be used to monitor and measure data accuracy. The Santa Barbara site of the Center is currently involved in a project with the California Department of Forestry to measure accuracies in GIS data used to estimate the State's forest resources. Estimates are currently made from a variety of sources including vegetation maps, air photos and ground truth. A model is being built to optimize the use of these sources, and a demonstration data set has been built using ARC/Info as part of the project. The final reports for the project are due in summer 1990.

1.1.6 Interpolation and surface modeling

The choice of data model has a significant impact on accuracy, but has more often been regarded as an issue of storage and processing efficiency. For example, the choice between DEM and TIN for storage of a topographic surface has a poorly understood impact on accuracy, as does the choice between field and object models, or raster and vector. In each of these cases the question of accuracy is linked to the nature of the spatial variation being modeled: the choice between DEM and TIN ultimately depends on the nature of the topographic surface being modeled and the erosion processes which formed it. Research is needed on the significance of accuracy versus storage and processing efficiency in choosing between different data models, and on the role of processes of spatial differentiation in this equation. Center research in this area has focussed on developing efficient methods of constructing TINs, and comparisons between the methods of DEM generation used in public agencies. David Theobald completed his Master's thesis on modeling hydrology using TINs in March 1990, and two papers are being prepared from this work, one on the theory and one on practical issues. Mark Kumler, a PhD student at Santa Barbara, worked at USGS Menlo Park in the summer of 1989 on the potential of TIN models for terrain modeling, and will continue this work as a PhD dissertation.

1.1.7 Aggregation, disaggregation and modifiable areal units

Several of the papers presented at the specialist meeting dealt with the impact of reporting zones on the results of socio-economic modeling, and the more general question of the impact of the spatial data model itself. There is a need for further development of techniques which exploit the capabilities of GIS for investigating and controlling data model effects. This includes the effects of reporting zone aggregation and disaggregation, as well as issues of interpolation between incompatible zones. As part of the initiative effort we have developed a general framework for transferring socioeconomic data from one set of spatial objects to another (reported by Deichmann, Goodchild and Anselin at the North American RSA meetings, 1989; Conference on Advanced Computing in the

Social Sciences, Williamsburg, VA, 1990). The problem is seen as one of estimating one or more underlying density surfaces under various assumptions.

Work by Amrhein at the Buffalo site of NCGIA has been focussed on the analysis of aggregation effects in migration modeling. Although social processes are best defined at the individual level, most data are available only at aggregate levels in order to protect privacy. Amrhein examines the extent to which scale is important in the calibration of migration models, and finds a surprising degree of independence. Also at Buffalo, Rogerson has examined the effects of aggregation in the time domain on migration modeling, arguing that migration data collected for periods of differing lengths will yield inconsistent population forecasts.

One possible strategy for dealing with problems of scale and resolution in spatial data is to search for approaches which are comparatively free of such effects. For example, Tobler argued at the specialist meeting that the best models of spatial phenomena would be ones which were independent of scale, and Fotheringham's presentation was concerned with ranges of scale over which certain phenomena (urban density in this case) showed a form of invariance. This area overlaps strongly with Initiative 3, which is concerned with the representation of objects at multiple scales. Tobler will present a paper on design criteria for a resel-based processing system at Zurich in July, 1990. Fractals provide one theoretical framework for dealing with scale effects, and work at the Buffalo site by Stewart Fotheringham, in collaboration with Michael Batty and Paul Longley of the University of Wales, has explored fractal models of uncertainty in socioeconomic data.

At Santa Barbara, Luc Anselin and Serge Rey continued their work on statistical tests for spatial dependence in linear regression models, with support from NCGIA. The simulation work completed during the summer of 1989 has been submitted to *Geographical Analysis*. The results illustrate the power of Lagrange multiplier tests to distinguish between substantive spatial dependence, and autocorrelation of errors. Also at Santa Barbara, Paul Slater has continued his research on spatial interaction models in continuous space, and their extension to time. A paper based on this work has been submitted for review.

1.2 Annotated list of publications

1.2.1 Articles in refereed journals

Amrhein, C.G. and R. Flowerdew. The effect of data aggregation on a Poisson regression model of Canadian migration. *Environment and Planning A* (to appear).

It is generally assumed that the manner in which data describing a spatial process are organized into mutually exclusive and collectively exhaustive areal units affects analytical results. The research reported in this paper finds a surprising absence of aggregation effects when interaction models are fitted to Canadian migration data, and advances several alternative explanations. (Also presented at Canadian Regional Science Association meetings, June 1989; North American Regional Science Association meetings, November 1989)

Anselin, L. and S. Rey. Properties of tests for spatial dependence in linear regression models. *Geographical Analysis* (under consideration).

The paper compares the properties of two tests for spatial dependence, based on a large number of Monte Carlo simulations on a regular lattice. The results provide an indication of the sample sizes for which the asymptotic properties of the tests can be considered to hold. They also illustrate the power of the tests to distinguish between substantive spatial dependence, and spatial dependence as a nuisance (error autocorrelation).

Arbia, G. and R.P. Haining. Error propagation through map operations. *Technometrics* (under consideration).

In image processing and GIS, a new map is constructed by carrying out a sequence of operations on a set of stored source maps. These operations typically include addition, ratioing and overlaying of two or more maps. But each source map may

contain error. The paper investigates the effects of different types of source map error on the error structure of the resulting map and shows how these effects also depend on the spatial structure of the true source map.

Batty, M., P.A. Longley and A.S. Fotheringham. Urban growth and form: scaling, fractal geometry and diffusion-limited aggregation. *Environment and Planning A* 21:1447-72 (1989).

The authors propose a model of growth and form in which the processes of growth are intimately linked to the resulting geometry of the system. The model generates highly ramified tree-like clusters of particles, or populations, with evident self-similarity about a fixed point. The authors suggest that the model provides a suitable baseline for the development of models of urban structure and density which manifest similar scaling properties.

Carson, Y.M. and R. Batta. Locating an ambulance on the Amherst campus of the State University of New York at Buffalo. *Interfaces* (to appear).

The paper applies the stochastic network approach described in the Berman *et al.* citation below to model uncertainty in travel times on the road network of the SUNY campus, and uses this to optimize locations for mobile units.

Fotheringham, A.S.. Guest editorial: What's the fuss about fractals? *Environment and Planning A* (to appear).

This short paper reviews the potential of fractals as a rich source of scale-free models of urban form, and argues for a link between fractal generating processes and the operation of certain probabilistic aspects of urban systems.

Fotheringham, A.S., M. Batty and P.A. Longley. Diffusion-limited aggregation and the fractal nature of urban growth. *Papers of the Regional Science Association* 67:55-69 (1989).

This paper introduces the mechanism of diffusion-limited aggregation (DLA) as a new basis for understanding urban growth. Through DLA, urban form is related to the processes of rural-to-urban migration and contiguous growth. Despite being based on very simple principles, DLA simulations are shown to have properties found in most urban areas, such as negative density gradients and ordered chaotic structures. Urban density gradients can occur independently of the generally accepted reasons for their presence.

Goodchild, M.F. Geographical data modeling. *Computers and Geosciences* (under consideration).

Data modeling is defined as the process of discretizing spatial variation, but is often confused with issues of data structure, and driven by available software rather than by a concern for accurate representation. The paper reviews the alternative data models commonly available in spatial databases, and assesses them from the perspective of accurate representation of geographical reality. Extensions are discussed, particularly for three dimensions and time dependence. (Presented at Inter Agency Committee on Geomatics seminar, Ottawa, March 1990; GIS Design Models Conference, Leicester, March 1990).

Goodchild, M.F. and Yang Shiren. A hierarchical spatial data structure for global geographic information systems. *Computer Vision, Graphics and Image Processing* (under consideration).

Hierarchical spatial data structures such as the quadtree offer distinct advantages of data compression and fast access, but are difficult to adapt to the globe. Following Dutton, we propose to project the globe onto an octahedron, and then recursively subdivide each of its triangular faces into four triangles. We provide procedures for addressing the hierarchy, and for computing addresses in the hierarchical structure from latitude and longitude, and vice versa. At any level in the hierarchy the finite elements are all triangles, but are only approximately equal in area and shape; we provide methods for computing area, and for finding the addresses of neighboring triangles. (Presented at Advanced Computing in the Social Sciences Conference, Williamsburg, VA, April 1990; University of Victoria, Simon Fraser University, February 1990).

Gopalan, R., K.S. Kolluri, R. Batta and M.H. Karwan. Modeling equity of risk in the transportation of hazardous materials. *Operations Research* (to appear).

The authors develop and analyze a model to generate an equitable set of routes for hazardous material shipments. The objective is to determine a set of routes which while minimizing total risk of travel, will also spread the risk equitably among the zones of the geographical region in which the transportation network is embedded. The findings indicated that one can achieve a high degree of equity by modestly increasing the total risk, and by embarking on different routes so as to evenly spread the risk among the zones.

Gopalan, R., R. Batta and M.H. Karwan. The equity constrained shortest path problem. *Computers and Operations Research* (to appear).

The article examines the problem of finding a shortest path on a network subject to equity constraints. Suggested applications include the routing of a vehicle carrying hazardous materials, routing a police car through a city, and planning a politician's campaign tour.

Lindner-Dutton, L., R. Batta and M.H. Karwan. Equitable sequencing of routes in the transportation of hazardous materials. *Transportation Science* (under consideration).

The authors consider the problem of equitably sequencing a set of hazardous materials shipments. The objective function is to minimize the sum of maximum differences in risk that exist between any two zones, where the sum is taken over the trips made.

Mark, D.M. and F. Csillag. The nature of boundaries on area-class maps. *Cartographica* 21:65-78 (1989).

Most research on cartographic line generalization has concentrated on linear features such as coastlines, rivers and roads, but these methods may not be appropriate for generalization of other types of lines. The paper presents a model of boundaries between classes on area-class maps, such as soil maps. Appropriate generalization may involve the construction of surfaces representing probabilities of class membership.

Rogerson, P.A. Buffon's Needle and the estimation of migration distances. *Mathematical Population Studies* (to appear).

The paper suggests a procedure for estimating migration distances from data on the proportion of migrants crossing regional boundaries. The method makes use of Buffon's Needle, a problem in geometrical probability from the 18th Century. An application to migration distances in the US is given.

Rogerson, P.A. Migration analysis using data with time intervals of differing widths. *Papers of the Regional Science Association* (to appear).

Migration data collected for periods of differing lengths will yield inconsistent population forecasts and alternative interpretations of mobility levels and migration patterns. Examples are given to illustrate the effects of migration interval choice. In addition to the level of mobility, the geographic pattern of migration flows is also affected by the choice of interval width.

Sivakumar, R.A. and R. Batta. Modeling wind effects in the transportation of liquified-gas hazardous materials. *Transportation Science* (under consideration).

Complementary to the problem of locating a site for the storage of hazardous material is the problem of transporting the substance to the chosen destination in the safest possible manner. One factor that has a significant bearing on the threat posed by an obnoxious vehicle carrying liquified-gas is the prevailing wind condition at the time of accidental release. The paper develops models that quantify the risk posed by the vehicle at different points on the network, and the covariance of risks. These are input to a decision model to determine the safest path from a set of available paths.

Slater, P.B. Path-dependent spatial interaction models for estimating migration histories at arbitrarily fine space-time intervals. *Geographical Analysis* (under consideration).

Models are proposed for taking either a recorded internal migration table for the geographic units of a nation and two times (start and end) or simply the population distributions over the units at the two times, and estimating the probability distributions over the possible migration histories of individuals. The probability distribution is sought which reproduces the recorded population distributions and the naturally weighted convex combination of these distributions at the intermediate points in time.

Viswanathan, V. and R. Batta. Demand point approximation to the planar, Euclidean distance, p-median problem with and without polygonal barriers to travel. *Operations Research* (under consideration).

This research on demand point approximations is important for three reasons: urban planners frequently make the simplifying assumption of restricting facility locations to demand points; reducing the number of candidate locations for facilities enhances our ability to solve larger problems; and demand point approximations lead to feasible heuristic solutions to actual problems. The paper develops both heuristic and exact methods for solving the problem of locating p facilities to serve a population dispersed over demand points, when travel across certain specified areas is prohibited.

1.2.3 Articles, chapters and monographs in other refereed outlets

Berman, O., S.S. Chiu, R.C. Larson, A.R. Odoni and R. Batta. Location of mobile units in a stochastic environment. In R.L. Francis and P.B. Mirchandani, editors, *Discrete Location Theory*. Wiley, New York (to appear).

Much previous work on finding locations for mobile units has assumed that the characteristics of the street network are known and fixed, and optimized for such parameters as travel time. Unfortunately congestion levels on street networks vary dramatically, leading to severe sub-optimality. Uncertainty is represented by allowing the network to adopt several discrete states, and finding probabilistic optima.

1.2.4 Articles in other outlets

Amrhein, C.G. and P. Schut. Data quality standards and Geographic Information Systems. *Proceedings: GIS for the 90s* (Ottawa, March 1990) (to appear). (Also presented at the Western Regional Science Association meetings, February 1990)

The rapid expansion of GIS applications has spawned a parallel increase in demands for information regarding the quality of spatial databases, particularly from government data collection agencies. But while positional accuracy is important, it is not a sufficient measure of accuracy for all applications. This paper reviews a range of accuracy measures for digital cartographic products in the context of one agency, Statistics Canada.

Anselin, L. What is special about spatial data? Alternative perspectives on spatial data analysis. Technical Report 89-4, National Center for Geographic Information and Analysis, Santa Barbara, CA.

This background paper prepared for the Initiative 1 Specialist Meeting describes the unique statistical characteristics of spatial data, and some of the techniques developed for dealing with them. Also published by IMAGE (Anselin, L. In D.A. Griffith, editor, *Spatial Statistics, Past, Present and Future*. Monograph Series, Institute for Mathematical Geography, Ann Arbor, Michigan).

Goodchild, M.F. Accuracy of spatial databases: Initiative 1 Specialist Meeting Report. Technical Report 89-1, National Center for Geographic Information and Analysis, Santa Barbara, CA.

This report on the I1 Specialist Meeting contains a summary of the discussion and abstracts of the paper presentations.

Goodchild, M.F. and B. Klinkenberg. Statistics of channel networks on fractional Brownian surfaces. In N. Lam, editor, *Fractals in Geography* (to appear).

The random topology model of channel networks is inappropriate as a null hypothesis in testing for the presence of geological controls, because packing and surface continuity impose additional constraints of a geometric nature. A range of surface conditions are simulated using fractional Brownian processes, and drainage networks extracted under a number of assignment rules. The paper raises a number of questions concerning interpretation of significance tests using the random model.

Goodchild, M.F. and S. Gopal. *Accuracy of Spatial Databases*. Taylor and Francis, New York (1989), 290pp.

This book is a collection of 23 of the papers presented at the Specialist Meeting for I1 and subsequently revised and edited. It contains several contributions from NCGIA personnel (Santa Barbara - Goodchild, Theobald, Tobler, Slater, Kennedy; Buffalo - Batta, Fotheringham, Amrhein), plus introductions to each section by the editors.

Goodchild, M.F. and Wang Min-hua. Modeling errors for remotely sensed data input to GIS. *Proceedings, AutoCarto 9* (Baltimore, MD) pp. 530-7.

Different views of spatial resolution and accuracy present a major obstacle to the integration of remote sensing and GIS. Accuracy in remote sensing is modeled using probabilities of class membership in each pixel; in vector-based GIS it is modeled using concepts such as the epsilon band. The problem of linking the two views of accuracy reduces to one of

realizing a stochastic process which must satisfy conditions of prior and posterior probabilities, and spatial dependence. The authors propose two suitable methods, one storage intensive and the other computationally intensive.

Goodchild, M.F. and Yang Shiren. A hierarchical spatial data structure for global geographic information systems. Technical Report 89-5, National Center for Geographic Information and Analysis, Santa Barbara CA (1989).

(see section 1.2.1 above)

Pullar, D. Comparative study of algorithms for reporting geometrical intersections. *Proceedings, Fourth International Symposium on Spatial Data Handling*, Zurich, July 1990 (to appear).

Two broad approaches to solving the geometrical intersection problem can be distinguished based on how adaptive the solutions are to the input data. "Theoretical" algorithms are only dependent upon the number of line segments, whereas the behavior of "pragmatic" algorithms depends on the statistical character of the underlying data. This paper compares the performance of theoretical and pragmatic solutions to the segment intersection problem using both algorithmic analysis and empirical tests.

Theobald, D.M. Automated delineation of hydromorphological features on a triangular irregular network-based digital elevation model. Master's thesis, Department of Geography, University of California, Santa Barbara (1990).

Information about terrain is especially useful in hydrological and geomorphological modeling. Automated techniques to represent terrain surfaces and delineate drainage networks have recently been developed, making it feasible to describe the terrain quantitatively. Not only do different data structures have different characteristics, but the manner in which flow is modeled with these data structures has a large impact on the quality of the simulation. The thesis examines hydrological modeling on TIN surfaces, and artifacts which arise under different TIN generation strategies.

Tobler, W.R. Design criteria for a resel processing system. *Proceedings, Fourth International Symposium on Spatial Data Handling* (Zurich, July 1990) (to appear).

There now exist several micro-computer processing systems incorporating algorithms and display techniques which are appropriate for the class of objects known as images. Zooming, histogram equalization, contrast stretching, edge enhancement, filtering and ratioing are common options used in these systems. The author proposes a system for the equivalent processing of general resolution elements (resels). Such a system requires a high resolution display and a new repertoire of processing algorithms.

Veregin, H. Accuracy of spatial databases: annotated bibliography. Technical Report 89-9, National Center for Geographic Information and Analysis, Santa Barbara, CA (1989).

Veregin's bibliography includes over 250 references on accuracy, error and uncertainty in spatial data. Each includes a short paragraph of annotation. The bibliography is intended as a tool for supporting research both inside and outside the Center.

Veregin, H. Taxonomy of error in spatial databases. Technical Report 89-12, National Center for Geographic Information and Analysis, Santa Barbara, CA (1989).

The taxonomy is intended as a user's guide to current knowledge of error problems in GIS, and is keyed to Veregin's bibliography. Its 90 pages include discussion of error source identification; detection and measurement; propagation modeling; strategies for error management; and strategies for error reduction.

2. Initiative 2: Languages of Spatial Relations

2.1 Narrative summary of progress to Date

Research Initiative 2 has as its broad aim the identification of elements of a fundamental theory of spatial relations. Initiative 2 chooses natural language and mathematics as the media for investigating this problem; later initiatives might approach the same objective through studies of vision, of spatial behavior, or other topics. Research is being conducted at all three NCGIA sites, although the bulk of the NCGIA resources, and of results, are at Buffalo and Maine. Although there is continuous interaction between the sites, in broad terms the work of the Maine group concentrates on the formalization of geometric concepts as they are used in GIS, while the Buffalo group focuses on the more cognitive aspects of spatial relations.

We currently are working on clarifying the notions of "concepts of space", as for example in Euclidean geometry, "geometric data models", as for instance the topological data model, and "geometric data structures" as they are used in an implementation. We use algebraic specification methods to formally describe the concepts at each level and expect to be able to construct mappings from the conceptual through the data model to the implementation level. So far, we have studied isolated cases, focussing on topological relations where we have achieved a systematic organization of the different existing relations. Work in progress tries to formally define a method of qualitative (in contrast to the regular quantitative) spatial reasoning, which is closer to the cognitive structures human beings seem to use.

Formalization of the concepts and methods used to deal with spatial data is important as it links the cognitive realm with implementation in an actual GIS. This becomes very apparent at the user interface, where human users communicate with the GIS system. For example, the spatial conditions that exist in a spatial query language must be formally defined and methods for their tests programmed.

In order to establish links between different spatial concepts expressed as algebras we will need to define morphisms (i.e., mappings between objects and operations from one algebra to those of another one) and we are currently studying the mathematical tools which will allow us to extend and combine algebras. We make extensive use of the results from software engineering studies, especially those relating to object-oriented techniques.

The following pages describe research accomplishments to date, other tangible products of the Initiative, and plans for the remainder of the project. It is intended that few NCGIA resources will be devoted to this Initiative after August 1990, although some of the research will continue beyond that date; one major theme within I-2 has developed into a new Research Initiative, number 13, "User Interfaces for GIS."

2.1.1 Computational models of locative expressions

One research theme at Buffalo has studied spatial prepositions and other spatial-relational terms in a cross-linguistic framework. This has had both theoretical and applied sides, and expects to contribute to cognitive science as well as to GIA. We feel that the NCGIA should not be totally inward-looking, but should also attempt to contribute significantly to nearby fields.

The theoretical side of this has been an examination of spatial prepositions. Annette Herskovits had previously analyzed the role of prepositions in locative expressions in English, and identified about 35 basic "use types" for the prepositions "in", "on", and "at" (Herskovits, A., 1987, *Spatial Prepositions in English*, Cambridge University Press). The thrust of our Initiative 2 work on this has been to test these use types in a cross-linguistic framework. A cross-linguistically valid list of use types for spatial prepositions (or equivalent grammatical elements) is essential to a computational model of spatial relations. Work has concentrated in particular on the relation involving lateral contiguity, as implied in "Cleveland is *on* Lake Erie". Preliminary results have been incorporated into a co-authored paper to be presented in Zurich. An extended paper on the "*on* the lake" relation is being prepared by Mark for submission to

Cognitive Science later in 1990.

During the summer of 1989, Susan Haller, a graduate student in Computer Science at Buffalo, conducted research on locative expressions; her work on the project was supported by the NCGIA, and supervised by Stuart Shapiro and Jeanette Neal (of the CALSPAN/UB Research Foundation). Haller added the generation of locative phrases in response to "where is" questions in the CUBRICON Multi-Media Interface, a DARPA-supported project monitored by the US Air Force. Haller's generation program chooses a reference ("ground") object based on the discourse context, what is currently displayed on a map, and the importance of various displayed objects, and generates a locative phrase relating the figure object to the ground object. The contents of the locative phrase depend on the distance of the figure to the ground and on the current scale of the map. Major findings are summarized in papers by Haller and Mark.

Further work on this topic will concentrate on linking the theoretical work to mathematical and formal research being conducted at the Maine site.

2.1.2 Driving directions and narrative theory

Researchers at Buffalo and Santa Barbara have instituted a long-term study of informal driving directions. Mark collected 31 sets of driving directions in the United States (mostly in Buffalo) by approaching strangers in public places, asking for directions to some other public places (taping the response), and later transcribing them. In Valencia, Spain, Gould collected 22 similar sets of directions (in Spanish). Also, in Los Angeles, Sucharita Gopal collected sets of verbal directions for walking to destinations on the UCLA campus.

A major thrust of the work at Buffalo is to relate the wayfinding directions to more general theories of narrative, including how people understand stories, and how computer programs can be designed to do the same. This may lead to procedures for understanding the spatial aspects of stories in a more general context, to computer integration of geographic texts into GIS, and also could lead to improved verbal directions from vehicle navigation-aid systems (see below). We have found that the driving directions collected so far exhibit many of the properties of stories and other forms of narrative, including deixis, fictive motion, and metaphor. Preliminary results were presented by Mark at the Association of American Geographers (AAG) meeting in Baltimore, March, 1989. More detailed results are being presented in a session of four papers at the 1990 AAG meeting at Toronto, in a paper to be presented in Zurich by Freundsuh and others, and in publications by Mark and Gould.

2.1.3 Acquisition and representation of spatial knowledge

A major background theme contributing to the linguistic aspects of spatial relations is the study of spatial knowledge acquisition. Scott Freundsuh (Buffalo) continued his doctoral dissertation research on models of spatial knowledge and on its acquisition. This work is an out-growth of Freundsuh's masters thesis on children's map-use and navigation abilities. He is currently conducting an experiment which explores how "regularity" of the environment affects both the acquisition of spatial knowledge, and the resulting accuracy of this knowledge in adults.

2.1.4 Vehicle navigation aid systems

One area of application for models of spatial knowledge and of the production of spatial language is in the context of provision of real-time navigation assistance for vehicle drivers. Three papers on this topic, written by members of the Initiative 2 research group at Buffalo, appeared in the proceedings of the IEEE-sponsored *Vehicle Navigation and Information Systems '89* conference in Toronto, Ontario, September 1989. Earlier, Gould presented a paper on this topic at the AAG East Lakes Division Meeting in Akron, Ohio, in October 1988; an expanded and revised version of that paper subsequently has been accepted by the *East Lakes Geographer*. Results of the research on verbal directions for wayfinding will eventually be applied to this problem. Mark is also co-authoring (with H. Claussen, University of Hannover) a review paper on this topic for a forthcoming volume edited by J.C. Muller, to be published for the International Cartographic Association by Elsevier.

2.1.5 User interfaces for GIS

An applied theme in the work at Buffalo was to examine the possible role of cross-linguistic differences in spatial language with regard to user interfaces. Initially, the study emphasized Spanish-English differences. A paper on this work, co-authored by Mark, Gould, and Joan Nunes, a Doctoral student from Universidad Autonoma de Barcelona (Spain) who studied at Buffalo between March and June 1989, was presented by Andrew Frank at the Second Latin-American Conference on Geographic Information Systems, in Merida, Venezuela, September 1989, when Mark and Gould's travel plans were cancelled due to Hurricane "Hugo". That paper mixed a linguistic analysis of differences in locative expression with a discussion of cross-linguistic and cross-cultural issues in GIS

user interfaces, and has been reprinted as an NCGIA Report. Gould received a National Science Foundation "Dissertation Improvement Grant" to continue this research with human-subjects testing in Barcelona (Spain) and Quito (Ecuador) during 1990-91.

Gould organized a special session on user interfaces for the GIS/LIS '89 meeting in Orlando, Florida, December 1989. Papers by Gould and by Mark are listed below. Stuart Shapiro (Chairman, Department of Computer Science, Buffalo) also presented a paper, co-authored with Jeanette Neal, Susan Haller, and others, in this session, but an error led to its omission from the *Proceedings*.

Finally, one of the most significant outcomes from Initiative 2 has been the establishment of a new Research Initiative, number 13, "User Interfaces for GIS", to be co-led by Mark and Frank. This is the first new Initiative to be adopted since the NCGIA was awarded by NSF. Initiative 13 will address human-computer interaction methods and related issues in the design and implementation of user interfaces for GISs and other geographical software packages. The Specialist Meeting is tentatively planned for November of 1990.

2.1.6 Metaphors for user interfaces

Humans use different means to convey and perceive spatial information than they use for non-spatial information. Traditional "typed" query languages for user-machine interaction have been shown to be cumbersome in a spatial environment where much data is represented graphically. At Maine, Jeff Jackson (graduate student), Werner Kuhn and Max Egenhofer are investigating appropriate techniques to communicate spatial information. These investigations are based upon the methodologies for user interface design developed in Egenhofer's and Kuhn's dissertations. Kuhn and Jackson have been working on the application of metaphors in user interfaces, e.g. the zoom and pan paradigm to view geographical and abstract information spaces. Jackson is finishing his master's thesis on this topic. He and Kuhn have presented their work at a surveyor's and at a computer science conference. They will write an article for a journal. The major results so far have been a formal definition of user interface metaphors and a new understanding of pan and zoom operations, based on fundamental properties of human vision.

Talks and presentations on this topic included: 1) Kuhn presented user interface issues at the International Seminar on Land Information Management, Land Tenure and Cadastral Systems (University of Maine); 2) Frank presented a paper on user interfaces at the International Cartographic Association's meeting in Budapest; 3) Kuhn presented a paper "Human Interaction with GIS - Research at the National Center for Geographic Information and Analysis" at the Annual Meeting of the New England / St. Lawrence Valley Association of American Geographers, Boston University; and 4) Kuhn presented a paper on metaphors in user interfaces at the Annual ACSM/ASPRS Convention in Denver, CO. In addition, Kuhn and Jackson gave an interactive poster presentation on formalizing user interface metaphors at the ACM CHI'90 Conference in Seattle, WA.

2.1.7 Visualization of spatial information

Mike White (former graduate student) has worked with Kuhn on the visualization and manipulation of spatial relations and constraints. Kuhn continues this and investigates abstraction mechanisms (classification, generalization and aggregation) as a means to represent metrical relations (e.g. on distances, angles and areas) graphically. This work is based on the Geometric Constraint Calculus from Kuhn's dissertation. A paper for the International Symposium on Spatial Data Handling (Zurich 1990), a presentation at the American Association of Geographer's meeting and a journal article are in preparation. Po-Siu Hsu (a PhD student) is investigating measures of information density for map displays. Frank, Beard, Hsu, and Kuhn are writing a journal article on this topic.

2.1.8 Algebras of spaces and morphisms between spaces

During the I-2 Specialist Meeting, a conceptual framework for spatial relationships was identified. Within this framework, there is an algebra for each space describing its properties and morphisms mapping from one space into another. Herring, Frank and Egenhofer will investigate this conceptual framework. The first algebra has been defined with Egenhofer's theory of topological relationships. A second algebra will be defined for order relationships based upon the recent dissertation by Wolfgang Kainz, University of Vienna, who has previously collaborated with the University of Maine. Some spatial relationships exist in both algebras, and we will investigate what the mappings between the two spaces are. Interim results will be presented in a paper at the 4th International Symposium on Spatial Data Handling at Zurich and a paper for the *International Journal of GIS* will report on the final findings.

2.1.9 Formal definitions of topological relationships

In Egenhofer's recently-completed dissertation, an approach to define topological relationships has been proposed. This method is based upon fundamental principles of algebraic topology. During the summer of 1989, Robert Franzosa from the Department of

Mathematics assisted Egenhofer in revising the mathematical section so that the theory can be published in a mathematical journal.

In cooperation with John Herring (a participant at the I2 Specialist Meeting from Intergraph Corporation), Egenhofer will investigate possible extensions of the theory described above. So far, the theory applies for n -dimensional objects in n -dimensional space. Investigations will focus on the extension to n -dimensional objects in m -dimensional spaces ($m > n$). The results of this investigation will be published in the *International Journal of GIS*. Also, a joint paper for the 4th International Symposium on Spatial Data Handling in Zurich is in print.

2.1.10 Metrical relationships

While metrical relationships between 0-dimensional objects in spaces of dimension > 0 are well-known, there has been little investigation of metrical relationships between 1- and 2-dimensional extended objects. Metrical relationships are relationships which describe the distance or direction between two objects. It has been demonstrated that representations, such as the center of gravity, are inadequate for objects of concave shapes and are not appropriate for reasoning about metrical relationships. Frank and Egenhofer will study metrical relationships within a framework of 1-dimensional intervals. This framework can be applied both to distance relationships and, in a variation with cyclic intervals, for direction relationships. The results will be published in two journal articles, the first one describing the theory and the conceptual framework; and the second article demonstrating how the theory may be applied to query languages and for reasoning about metrical relationships.

2.1.11 Qualitative reasoning about space

Most current methods of spatial reasoning are quantitative, i.e. are based on measurable (metric) properties of space and the configuration of objects about which one reasons. For example, simple vector algebra allows for the deduction of distance and direction from A to C given the distance and direction from A to B and from B to C. More elaborate (quantitative) systems can deal with situations where the directions are expressed in different reference frames (e.g. in front of the home) or when the quantities are not exactly known. There have also been approaches to translate qualitative terms (near, far, south, etc.) into fuzzy set quantities and then to calculate with them. Frank has started to explore what power purely qualitative systems have. He has constructed a set of rules for reasoning about distances and directions in a hierarchically structured world of named places. Places are grouped into a hierarchy of clusters (simulating "regions").

Directions are expressed as one of eight cardinal directions and distances are ordered values. The structure is built from information about what places belong to the same "region" and the direction from a place to the center of the region. Rules for inferring distances and directions between arbitrary places are formalized. It uses only symbolic reasoning and does not rely on coordinates. A draft paper with most of the formalism has been written. We want to explore how closely this system corresponds to cognitive structures and how much it deviates from traditional Euclidean reasoning about distances and directions.

This approach opens an interesting field for study. The system is relatively generic and considerable variations are possible in order to approximate better some observed behavior. We also assume that quite different qualitative spatial reasoning systems can be constructed, compared and possibly combined. We can compare these qualitative spatial reasoning systems with the methods humans use. Of interest is to establish situations where humans draw conclusions that the formal system does not permit, and situations where humans reach different conclusions.

2.1.12 Query languages

Interactive query languages are the user's tools to request data from a database. Conventional database query languages, such as SQL, lack the particular functionality to address the retrieval and representation of spatial data. In Egenhofer's PhD thesis, a comprehensive collection of spatial query language components have been presented. These are: the treatment of complex objects and corresponding (geometric) operations; the graphical representation of query results; interaction with graphically-displayed objects via direct manipulation; the combination of multiple query results in a single rendering; the description of the graphical representation in terms of colors, patterns, and symbols; and the definition of spatial context to be added to the user query. An extension of SQL, including these spatial features as well as the design of a human interface to such a spatial query language, has been proposed. Interim results were presented at geography and refereed computer science conferences and final results have been submitted to refereed journals.

2.1.13 Other activities

The major International Conference at the end of Initiative 2 will be a *NATO Advanced Study Institute* entitled "Cognitive and Linguistic Aspects of Geographic Space", which will be held in Las Navas del Marques, Spain, July 8-20, 1990. The ASI received a grant of 2 million Belgian Francs (about \$54,000) from NATO, and also has partial support from the NCGIA. Mark is the Director of the ASI, and Frank is the co-Director. A major product of the ASI will be a book, to be edited by Mark and Frank, which will contain

about 12 major review papers and a set of rigorously-reviewed research papers.

Immediately after the ASI, Initiative 2 participants will present at least five papers at the Fourth International Symposium on Spatial Data Handling in Zurich. Papers reporting on results of I2 have been accepted from Egenhofer and Herring ("A Mathematical Framework for the Definition of Topological Relationships"); Herring, Egenhofer, and Frank ("Using Category Theory to Model GIS Applications"); Kuhn ("Editing Spatial Relations"); and Hsu ("An Analysis of Spatial Structure Data in Landscape for Geographic Information Systems").

We have compiled a "Working Bibliography on Languages of Spatial Relations". In 1991, Longman's will publish "Geographical Information Systems: Principles and Applications", to be edited by David Maguire, David Rhind, and Michael Goodchild. At least two chapters of the handbook will be I-2 related and written by NCGIA personnel. David Mark and Andrew Frank will co-author a chapter on language issues for GIS, and Max Egenhofer and John Herring are writing a chapter, "High Level Data Structures".

2.2 Annotated list of publications

2.2.1 Articles in refereed journals

Egenhofer, M.J. and A.U. Frank. User interface for spatial information systems. *Geologisches Jahrbuch* (under consideration).

Computerized information systems that are tailored to spatial data handling play an emerging role as sources of spatially related information. By managing spatial data in computer systems, new methods are developed to share and exchange spatial information. Human understanding of spatial data and human means of communicating with spatial information systems are the focus of this investigation. The paper investigates the dualism between operations to manipulate the graphical representation of spatial objects and how these representations are observed.

Egenhofer, M. and R. Franzosa. Point-set topological spatial relations. *American Mathematical Monthly* (under consideration).

A novel theory of topological spatial relations between sets has been developed in which the relations are defined in terms of the intersections of the boundaries and interiors of two sets. By considering empty and non-empty as the values of the intersections, a total of sixteen topological spatial relations are described, each of which can be realized in \mathbb{R}^2 . This set is reduced to eight relations if the point sets are restricted to the homeomorphic images of a 2-disk. It is shown that these relations correspond to the set-theoretic notions of equality, disjointness, and containment in the interior.

Frank, A., M. Egenhofer and W. Kuhn. Computational topology: data structures and algorithms. *Cartography and GIS* (under consideration).

The finiteness of computers and their number systems seriously impedes implementations of Euclidean geometry, the most common geometric data model in current commercial Geographic Information Systems. To overcome the dependencies on coordinates, a topological data model is proposed based on the recognition of incidences. Operations to be performed on an irregular triangulation are selected from a finite set of operations on simplicial complexes. Algorithms for neighborhood queries and creations of new complexes are presented and results of the implementation of a prototype on top of an object-oriented database management system are reported.

Egenhofer, M. and D. Pullar. A formal description of topological relations between one-dimensional spatial objects. *Computer Vision, Graphics and Image Processing* (under consideration).

The topological relationships are a specific subset of the large variety of spatial relationships. They are characterized by the property to be preserved under topological transformations, such as translation, rotation, and scaling. A model of topological

relations is presented which is based upon fundamental concepts of algebraic topology in combination with set theory. Binary topological relationships may be defined in terms of the boundaries and interiors of the two objects to be compared. A formalism is developed which identifies 16 potential relationships. An earlier version of this paper appeared in proceedings form (Egenhofer, M. A formal definition of binary topological relationships. In W. Litwin and H. Shek, editors, *Lecture Notes in Computer Science Vol. 367* (3rd International Conference on Foundations of Data Organization and Algorithms (FODO), Paris, France). Springer-Verlag, New York (1989) pp. 457-472.)

Egenhofer, M. and A.U. Frank. LOBSTER: combining AI and database techniques for GIS. *Photogrammetric Engineering and Remote Sensing* (Special Issue on Knowledge-Based Expert Systems) (to appear).

The powerful logic-based concept of Prolog has been integrated with a database suitable for spatial data handling to form a database query language that is more flexible and powerful than the currently used SQL. This experimental implementation, called LOBSTER, allowed researchers to explore a number of areas of a GIS. Examples from object-oriented modeling, geomorphology, and query optimization show the application of such a language. Problems encountered during the application of LOBSTER include the absence of consistency checking during input of rules and facts, and the lack of appropriate techniques to detect cyclic rule definitions.

Freundschuh, S.M. Can young children use maps to navigate? *Cartographica* (under consideration).

This paper explores preliterate children's abilities to use a map for navigation. Using a map with a route indicated on it, 33 children attempted navigation of an obstacle course. It was found that some 4 year olds, and most 5 year olds were able to use the map to navigate the obstacle course. No formal conclusions were drawn concerning the 6 year olds. It was demonstrated that all of the children who participated in this study were capable of extracting the necessary information from a map.

Gould, M.D. Exploring cognitive tests as a means to categorize drivers by navigation ability. *The East Lakes Geographer* 24 (to appear).

Although it has been shown that drivers possess differing levels of ability at navigating large-scale environments, it is not yet known which basic cognitive abilities are direct contributors to navigation. This paper suggests that drivers may be tested for individual differences in key basic cognitive abilities and that navigation ability categories can be inferred from the results. A goal of driver categorization is to better serve individual drivers with navigation assistance such as computerized wayfinding systems or new methods of direction-giving.

Mark, D.M. and M.D. Gould. Wayfinding as discourse: A comparison of verbal directions in English and Spanish. *Language and Communication* (under consideration).

Verbal driving directions were solicited and tape-recorded from 31 English-speaking subjects in North America and from 22 Spanish-speaking subjects in Spain. The transcripts were analyzed within the context of principles of deixis from narrative theory and of cognitive models of spatial language. Both English- and Spanish-speakers provide verbal driving directions utilizing many of the cognitive schemata and linguistic-coding rules discussed recently in the context of narrative comprehension. Verbal driving directions may easily be collected in a natural environment, allowing a unique opportunity to contrast attributes of natural language with results obtained by research in narrative comprehension.

2.2.2 Articles in refereed conference proceedings

Egenhofer, M. and A.U. Frank. Towards a spatial query language: user interface considerations. In D. DeWitt and F. Bauchilar, editors, *14th International Conference on Very Large Data Bases* (Long Beach, CA (1988) pp. 124-133.

Current database query languages, which are well-suited to treat alphanumeric data, do not reflect the properties of spatial data. In order to visualize the power needed for a spatial query language, a user interface is designed which gives specific consideration to the coexistence of representations as graphical renderings, such as maps, and lexical listings, such as tables. The different properties of spatial and non-spatial data give rise to lexical formulations of queries in combination with references to graphical objects or areas on maps. An earlier version of this paper will appear in *Proceedings, International Cartographic Association* (Budapest, August 1989).

2.2.3 Articles, chapters and monographs in other refereed outlets

Frank, A.U. and D.M. Mark. Language issues for GIS. In D.J. Maguire, M.F. Goodchild and D.W. Rhind, editors, *Geographical Information Systems: Principles and Applications* (to appear).

This chapter reviews the major issues involving language and GIS, and thus provides a summary of Initiative 2 for a broad audience of GIS users and researchers. After an introduction, the chapter first discusses cognitive science, especially cognitive linguistics, and then describes the mathematical representation of geographic space in GIS. Query languages, especially with respect to spatial data, are reviewed. Issues discussed include natural language queries and commands, including a discourse model of human-computer interaction; input of geographic data to GIS in text form; and natural language output.

2.2.4 Articles in other outlets

Barrera, R. and J. Vasquez-Gomez. A shortest path method for hierarchical terrain models. *Proceedings, AutoCarto 9* (Baltimore, Maryland) (1989) pp. 156-63.

An algorithm is presented for obtaining the shortest path between two points on a terrain represented by a triangular-faced polyhedron. The terrain model is hierarchical, i.e. it has several level of precision, the representation at each level refining the previous one. The proposed algorithm consists of two phases. In the initial phase, terrain representations at increasing precision levels are searched for regions where no optimal paths can trespass; these regions are not to be considered any further. In the final phase, a standard shortest path algorithm is applied on the remaining areas.

Buyong, T. and W. Kuhn. Local network adjustment for a measurement-based multipurpose cadastre. *FIG, XIX International Congress Proceedings* (Helsinki, Finland) (to appear).

In a measurement-based system, it is impractical to adjust all measurements in the database every time coordinate values of some points are required. Instead, the adjustments include only those neighboring measurements that significantly influence the results of the adjustment of the desired area. The local adjustment bears the same results as the global adjustment for most cadastral purposes.

Buyong, T. and W. Kuhn. Local adjustment for cadastral measurement databases. *Proceedings, 1990 ACSM/ASPRS Annual Convention* (Denver, CO).

Early parcel-based land information systems were created mainly for land taxation and land management. Today's computerized systems however, serve multiple purposes with a large variation of functionalities. Some of the functions

demand more up-to-date metric information. This scenario has resulted in the emergence of the concept of a cadastral measurement database. The demand for fast response and the high frequency of query make adjustment of the complete measurements in the database impractical. This paper proposes that only measurements in the locality of the query area need to be adjusted.

Buyong, T. Utility mapping systems based on measurements. *Proceedings, URISA '89* (Boston, MA) Vol. 2, pp. 222-230.

Many utility companies have taken significant steps towards implementing computerized utility mapping systems in a multipurpose cadastral system setup. The cadastral base map needs frequent upgrades due to ongoing improvements and revisions. Current implementations of coordinate-based utility mapping systems do not permit these upgrades to be reflected in the utility mapping system easily. A computerized mapping system based on measurements is proposed in which the changes in the base map are automatically propagated in the utility mapping system. This approach guarantees correct registration of utility maps and base maps.

Buyong, T. and A.U. Frank. Measurement based multipurpose cadastre. *Technical Papers, ASPRS/ACSM Annual Convention* (Baltimore, MD) Vol. 5, pp. 58-66 (1989).

Many local level governments have taken significant steps towards implementing a multipurpose cadastre. Most of them follow the National Research Council study which recommends starting a multipurpose cadastral system with a good network of geodetic control points. This approach, however, has several major problems. The establishment of geodetic network is very costly and time consuming and often cannot be provided by local level governments. We propose a multipurpose cadastral system based on measurements where the implementation does not require the immediate completion of a geodetic control network.

Egenhofer, M.J. and J. Herring. A mathematical framework for the definition of topological relationships. *4th International Symposium on Spatial Data Handling* (Zurich, July 1990). (to appear).

Frank, A.U.. Artificial intelligence in GIS applications. Workbook for U.S. Army Symposium on Artificial Intelligence Research for Exploitation of the Battlefield Environment, November 1988.

Frank, A.U., D. Hudson and V. Robinson. The future GIS - an expert system?. In B. Thompson and D. Kopec, editors, *Artificial Intelligence & Intelligent Tutoring Systems* (University of Maine Spring Symposium (1989) pp. 82-100.

The essential topics in Geographic Information Systems (GIS), spatial and temporal reasoning, are difficult and require powerful tools and a formal systems approach. Future GIS will include tools and methods developed in artificial intelligence research and will resemble today's expert systems. An expert system is defined here to be a computer-based mechanism that produces results comparable to those obtainable from a human expert in some field. Moreover, intelligent user interfaces and reasoning about data quality can also benefit from the power of expert systems.

Freundschuh, S.M., D.M. Mark, S. Gopal, M. Gould and H. Couclelis. Verbal directions for wayfinding: implications for Geographic Information and Analysis Systems. *Proceedings, Fourth International Symposium on Spatial Data Handling* (Zurich, July 1990) (to appear).

This paper presents and compares the results of four studies of verbal directions for wayfinding, which provide natural, easily accessible methods for the investigation of several issues relating to cognitive maps and spatial problem-solving. These

wayfinding directions were analyzed for properties such as: numbers and kinds of landmarks; presence of metrical information such as distance and travel times; deictic and reference frame terms; "style" of directions; and total duration of the direction giving act. These properties were tabulated against factors such as sex, written vs. spoken directions, and the language spoken.

Freundschuh, S.M. Does anybody really want or need vehicle navigation aids? *Proceedings, VNIS'89* (IEEE Conference on Vehicle Navigation and Information Systems, Toronto, September 12-14, 1989), 439-442.

This paper reviews current research on spatial knowledge acquisition in computer science, psychology, and geography, and suggests how these models can enhance VNA development. In addition, the paper discusses current research in cognitive psychology and environmental behavior concerning the presentation of driving instructions to drivers of automobiles. The development of effective VNAs is not a problem to be solved by only one discipline, but rather requires a multi-disciplinary endeavor. In many cases, research findings in one discipline do support research findings in other disciplines. However, caution should be taken when interpreting these results.

Freundschuh, S.M., M.D. Gould and D.M. Mark. Issues in vehicle navigation and information systems. Technical Report 89-15, National Center for Geographic Information and Analysis, Santa Barbara, CA.

This NCGIA report collects and reprints the three papers by NCGIA researchers that were presented at *VNIS'89*, the IEEE-sponsored interdisciplinary conference on Vehicle Navigation and Information Systems, held in Toronto, Ontario, September 12-14, 1989. The papers are listed separately in this section, and their contents are described.

Gould, M.D. Considering individual cognitive ability in the provision of usable navigation assistance. *Proceedings, VNIS'89* (IEEE Conference on Vehicle Navigation and Information Systems, Toronto, September 12-14, 1989), 443-7.

The design of vehicle navigation aids (VNAs) has emphasized hardware innovations, and software which is optimized to this hardware. User interfaces to VNAs generally adopt a single mode of presentation of navigation information. This assumes a homogeneous user community and ignores individual variation regarding drivers' levels of spatial cognition, their attention, and other problem solving abilities. Future VNAs should be designed by devoting more attention to the manner in which people structure, recall, and utilize spatial information naturally. Literature on spatial knowledge, individual differences, and human subject categorization is reviewed, and recommendations for the design of future VNAs are made.

Gould, M.D. Human factors research and its value to GIS user interface design. *Proceedings, GIS/LIS'89*, Orlando, Florida, 541-550.

User interfaces for both CAD/CAM and database management systems have been optimized to the systems' respective problem domains, partly due to attention to preceding research in human factors. Geographic information systems, however, do not possess user interfaces optimized for the geographic problem solving domain. Greater attention to innovative human factors research and focus upon cognitive science, rather than upon hardware/software aspects of human-computer interaction, will lead to the first generation of spatially-oriented user interfaces for GIS. Directions for future research are suggested.

Gould, M.D. and M. McGranaghan Metaphor in Geographic Information Systems. *Proceedings, Fourth International Symposium on Spatial Data Handling* (Zurich, July 1990) (to appear).

Developers of geographic information systems are beginning to adopt graphical, direct-manipulation user interfaces, usually modeled after the "desktop" metaphor. While the desktop is perhaps an optimal metaphor for general office automation, it is lacking as an organizing metaphor for GIS. The cartographic map, though an integral part of the GIS interface, is also lacking as an organizing metaphor. Definitions of metaphor are reviewed, to provide insight on its value in a GIS. The role of nesting various levels of metaphor is suggested as a means to organize the diverse system administration, geographic analysis, and cartographic display needs of GIS.

Haller, S.M. and D.M. Mark. Knowledge representation for understanding geographical locatives. *Proceedings, Fourth International Symposium on Spatial Data Handling* (Zurich, July 1990) (to appear).

This paper addresses the generation of natural language to express spatial relations. It is indirectly concerned with understanding such locative expressions, and with what their structure reveals about human conceptualizations of space. The main focus is the knowledge representations that are needed to support locative phrase generation. The multiple representation of, say, a city as sometimes a point and sometimes a polygon is modelled in a semantic network as two (or more) intensional objects related to the same extensional ("real world") object. The knowledge representation is then examined in a cross-linguistic perspective, and a detailed example is given.

Haller, S.M. Spatial relations and locative phrase generation in a map context. Technical Report 89-14, National Center for Geographic Information and Analysis, Santa Barbara, CA.

This paper explores the generation of natural language locative phrases with respect to maps and in response to "Where is?" queries by users. Although alternative orientation strategies are possible, we restrict our problem by assuming that the correct ground (reference point) used for locating a figure (place being located) is some sort of stationary landmark of sufficient size and importance that is visible on the map. Hence, our problem involves choosing a suitable landmark for the ground, building a knowledge representation to express the relationship of the figure to the ground, and generating natural language to express that relationship.

Heiler, S. and A.U. Frank *et al.*. The integration of heterogeneous computing environments. In E.N. Fong and A.H. Goldfine, editors, *Information Management Directions: The Integration Challenge*. NIST Special Publication 500-167 (1989) pp. 123-134.

Herring, J., M.J. Egenhofer and A.U. Frank. Using category theory to model GIS applications. *Proceedings, 4th International Symposium on Spatial Data Handling* (Zurich, July 1990) (to appear).

Hsu, P. An analysis of spatial structure data in landscape for GIS. *Proceedings, Fourth International Symposium on Spatial Data Handling* (Zurich, 1990) (to appear).

The spatial structure of the landscape is a configuration of the physical open space of a given site with components such as structure, identity, and meaning. It is generally the organization of spaces where topographic characteristics, vegetation mass, built forms, spatial relationships, spatial organizations, and ordering principles interact. This research presents findings on the elements of spatial structure that are most significant to designers and planners. A model of representation is proposed to further understand the perceptual space as part of a spatial structure.

Jackson, J. Developing an effective human interface for GIS using metaphors. *Proceedings, 1990 ACSM/ASPRS Annual Convention* (Denver, CO).

The emphasis in the design of GIS has historically been placed on architecture, database management, and data structures. Unfortunately, little attention has been paid to the development of a system which will effectively manage the interaction between the user and the software. This paper describes the advantages of using metaphors to make unfamiliar computer processes correspond to more common situations.

Kuhn, W. From constructing towards editing geometry. *Proceedings, 1990 ACSM/ASPRS Annual Convention* (Denver,CO).

With the rapid dissemination of GIS/LIS technology to a wide variety of user communities, human-computer interfaces are becoming more and more important in this area. Existing user interfaces of GIS and LIS emphasize the retrieval and display (i.e. output) of data over the acquisition and editing (i.e. input). The development of powerful and easy to use tools for the latter has been hindered by the lack of a theoretically sound basis. In order to overcome this impediment, the author has developed a language to express general geometric relations in the plane.

Kuhn, W. Interaction with spatial information systems: from constructing towards editing geometry (in German). ETH Zurich, Institute of Geodesy and Photogrammetry, Mitteilungen No. 45.

Kuhn, W. and A.U. Frank. Human interaction with GIS/LIS: editing geometric models. *FIG, XIX International Congress Proceedings* (Helsinki) (to appear).

The acquisition and interactive manipulation of geometric data are fundamental tasks in the use of geographic and land information systems (GIS/LIS). Most systems offer methods for these tasks which are derived from manual geometric constructions. They tend to have restricted and complex user interfaces. The paper explains these shortcomings and proposes an alternative approach, based on the idea of sketching and declaring geomatic constraints.

Kuhn, W. Editing spatial relations. *Proceedings, Fourth International Symposium on Spatial Data Handling* (Zurich, July 1990) (to appear).

Kuhn, W. and J. Jackson. Specifying metaphors algebraically. Poster accepted for CHI'90, ACM SIGCHI Conference on Human Factors and Computing Systems, Seattle, WA.

Mark, D.M. Cognitive and linguistic aspects of geographic space: report on a workshop. Miscellaneous Report, National Center for Geographic Information and Analysis, Santa Barbara, CA.

This is a report on a workshop entitled "Cognitive and Linguistic Aspects of Geographic Space", held in Buffalo, New York, June 11-12, 1988. The workshop brought together about 20 researchers from geography, linguistics, cognitive science, and engineering, to discuss topics of mutual interest. The Workshop was supported by a "Conferences in the Disciplines" grant from SUNY at Buffalo, by the Faculty of Social Sciences (SUNY/Buffalo), and by the Graduate Research Initiative in Cognitive and Linguistic Sciences (SUNY/Buffalo). Although held before the NCGIA grant was awarded, the workshop represented, in many ways, the start of Initiative 2.

Mark, D.M. (1989) A conceptual model for vehicle navigation systems. *Proceedings, VNIS'89* (IEEE Conference on Vehicle Navigation and Information Systems, Toronto, September 12-14, 1989), 448-53.

This paper defines a number of terms and concepts related to vehicle navigation, and presents a model which can be used in designing and evaluating components of the navigation system. Support of human navigation has long been a vital function of maps. Recent technological developments have led to new forms of navigation aids, not all of which include maps. Cognitive science studies the workings of the mind, and provides a useful theoretical basis for examining the navigation process and its relation to spatial learning. This provides a conceptual framework for an evaluation of maps and map-alternatives for road navigation.

Mark, D.M. Cognitive image-schemata for geographic information: relations to user views and GIS interfaces. *Proceedings, GIS/LIS'89*, Orlando, Florida, 551-60.

Image-schemata are idealized conceptual models for human perception and cognition. Many such schemata are spatial, and some are geographic. Users interact with Geographic Information Systems (GISs) in order to learn about, or make decisions about the world. This paper contends that optimal user interfaces for GIS will be based on image-schemata for geographic and other spatial phenomena. The concept of user views also relates to this schema-based approach. An early focus on users and interfaces is important in system design, especially for systems to be used by people from different disciplines, cultures, and languages.

Mark, D.M. Languages of spatial relations: researchable questions and NCGIA research agenda. Technical Report 89-2A, National Center for Geographic Information and Analysis, Santa Barbara, CA.

One of the objectives of the Initiative 2 Specialist Meeting was to compile a list of "researchable questions" related to the topic. This Report lists the original set of 65 "Researchable Questions" identified during that meeting. Many of the questions are of a scope similar to a Master's degree thesis; certainly, graduate students looking for thesis, project, or dissertation topics are encouraged to peruse this list and to consider addressing one or more of these questions. The second part of the report contains specific research projects to be conducted at each of the three NCGIA sites contained in the second part of this report.

Mark, D.M. and A.U. Frank. Concepts of space and spatial language. *Proceedings, AutoCarto 9* (Ninth International Symposium on Computer-Assisted Cartography, Baltimore, MD, 1989), 538-56.

Development of a comprehensive model of spatial relations is important to improved geographic information and analysis systems, and also to cognitive science and behavioral geography. This paper first reviews concepts of space. A critical distinction is between small-scale spaces, whose geometry can be directly perceived, and large-scale space, which can be perceived only in relatively small parts. Fundamental terms for spatial relations often are based on concepts from small-scale space, and are metaphorically extended to large-scale (geographic) space. Reference frames, which form an important basis both for spatial language and for spatial reasoning, are discussed.

Mark, D.M., A.U. Frank, M.J. Egenhofer, S.M. Freundschuh, M. McGranaghan and R.M. White. Languages of spatial relations: Initiative 2 specialist meeting report. Technical Report 89-2, National Center for Geographic Information and Analysis, Santa Barbara, CA.

GIS progress is impeded in many cases by poorly-designed or inappropriate user interfaces and query languages. Also, some types of geographic data are in text form, and their entry into a GIS requires either language analysis or a great deal of human effort. Cognitive science provides a framework for relating geographic language, and spatial relations and concepts, to GIS. This report is essentially a synthesized transcript of what was said during the Specialist Meeting for NCGIA Research Initiative 2. That meeting brought together geographers, cognitive linguists, engineers, computer scientists, and others. Working groups identified a research agenda for the topic.

Mark, D.M., M.D. Gould, S.M. Freundsuh, M.J. Egenhofer, W. Kuhn, M. McGranaghan and S. Svorou. Working bibliography on "Languages of Spatial Relations." Technical Report 89-10, National Center for Geographic Information and Analysis, Santa Barbara, CA.

This report contains over 500 references on "Languages of Spatial Relations" and related topics. The core focuses on linguistic studies of how natural languages represent and express objects and relations in geographic space, and on GIS data structures. The "related topics" include selected or key papers on human spatial cognition and learning, and on human-computer interfaces. The bibliography is intended to be the first edition of a "working bibliography" of this, the topic of NCGIA Research Initiative 2.

Mark, D.M., M.D. Gould and J. Nunes. Spatial language and geographic information systems: cross-linguistic issues. *Proceedings, 2nd Latin American Conference on Applications of Geographic Information Systems* (Merida, Venezuela), 105-30. (This paper was reprinted as NCGIA Technical Report 90-1).

Most existing GISs were designed by English or German speakers. Since languages impose structure on the cognition and perception of space, GIS data models, query languages, and user interfaces probably contain artifacts of the language spoken by their designers. Natural language studies are important because GISs of the future will have to handle natural language in a number of situations. This paper expands on general principles of cognitive linguistics, with emphasis on cross linguistic issues. Then, it reviews the primitive geographic relations represented in most Indo-European languages by prepositions, concentrating on differences between English and Spanish.

Neal, J.M., C.Y. Thileman, Z. Dobes, S.M. Haller and S.C. Shapiro. Natural language with integrated deictic and graphic gestures. *Proceedings of the DARPA Speech and Natural Language Workshop*, Morgan Kaufmann, Inc., Los Altos, CA, (to appear).

People frequently and effectively integrate deictic (pointing) gestures with speech (natural language, NL) when conducting human-to-human dialogue. Similar multi-modal communication can facilitate human interaction with computers. As part of the CUBRICON project, we are developing NL processing technology that incorporates deictic and graphic gestures with simultaneous coordinated NL for both user inputs and system-generated outputs. The domain for the research is tactical Air Force mission planning, and involves geographic data and problems, and the generation of locative expressions.

Rapaport, W.J., E.M. Segal, S.C. Shapiro, D.A. Zubin, G.A. Bruder, J.F. Duchan and D.M. Mark. Cognitive and computer systems for understanding narrative text. Technical Report 89-07, Department of Computer Science, University at Buffalo.

This report describes a long-term project to develop a computational theory of how humans understand narrative text. The theory will be informed by joint research from the viewpoints of linguistics, cognitive psychology, language acquisition, literary theory, geography, philosophy, and artificial intelligence. The report describes the knowledge representation and natural language processing issues involved in the computational implementation of the theory. It includes a section by Mark on how readers might develop an understanding of the geographical space in which a story occurs. This report is a longer version of a research proposal to NSF.

Steiner, D., M.J. Egenhofer and A.U. Frank. Cartographic symbols: an object-oriented approach. *Proceedings ICA* (Budapest, 1989) (to appear).

An object-oriented approach cartographic output package has been proposed as a solution to the shortcomings of existing procedural graphics packages for the display of query responses to GIS. This paper discusses the design of the portions of this package which deal with symbology and the mapping of objects to the display. Spatial objects can be represented by a variety of different symbols depending on the context of the map. The object-oriented approach used in the design of this cartographic package facilitates this variation of symbology. An earlier version of this paper appeared as Steiner, D., M.J. Egenhofer and A.U. Frank. An object-oriented cartographic output package. *Technical Papers, ASPRS/ACM Annual Convention* (Baltimore,MD) Vol. 5 (1989) pp. 104-113.

3. Initiative 3: Multiple Representations

3.1 Narrative summary of progress to date

The scope of concern in the use of the term "Multiple Representations" refers to changes in geometric and topological structure of a digital object that may occur with the changing resolution at which that object is encoded for computer storage, analysis and depiction. Cartographic and digital data arguably provide only a sample of the geographic features they are intended to represent. As discrete approximations of a continuous reality, each cartographic object may capture at most a subset of its geographic counterpart. The resolution at which features are captured by digital encoding methods will often bias the amount and types of details about the features that are encoded. The implications for geographic information and analysis formed the basis for prioritizing a research agenda.

In the context of GIS and geographical analysis, this is an important problem because the nature of geographic data renders changes in its appearance and underlying structure depending on the resolution at which it is digitally encoded. Topographic information and statistical phenomena with a spatial component both tend to exhibit characteristics of scale-dependence, and it renders more difficult the process of automating the mapping sciences as well as the analysis of geographical pattern. The expense and tedium of generating unique database information for every desired scale of representation continues to challenge data production, to inhibit research efforts, and to limit the reliability of many GIS applications involving spatial decision support.

Products of the initiative continue to develop since the Specialist Meeting in February 1989, and include research projects and publications, a bibliography, and generation of a multi-scale database for research and teaching. Details will be discussed briefly for each in turn. A comprehensive discussion of the prioritized research agenda and products may be found in the Specialist Meeting report.

It is important to realize that this provides a starting point for research on multiple representations research, an agenda of research priorities to be discussed, criticized, and expanded by interested researchers in many disciplines. Topics discussed and explored at the Specialist Meeting and in subsequent research are of course colored according to special interests and knowledge of the particular researchers. The formal endpoint of the Initiative (currently scheduled for August, 1990) is not intended as the terminus of the research, but rather as a point at which the establishment and refinement of a research agenda will provide impetus for the rest of the research community.

At the Specialist Meeting, attention focused on research questions that can be addressed now given the current level of technology and software engineering, and given the current state of knowledge. Two sets of issues surfaced for immediate research attention. It is important to understand that these are not the only two areas, but rather the two broad topics felt by the Specialist Meeting participants to be of highest priority and most readily pursued at present.

3.1.1 Database issues

Database issues must be addressed to accommodate multiple representations in a single or multiple version data management strategy. Of central importance is the need to organize multiple topological and metrical versions for efficient access, and implementation of linkages between multiple representations. New strategies for maintenance of materialized database views is another important research topic. Rules to preserve consistency between database views generated at different levels of resolution must be refined as well. Projects developing from these topics include investigation of the following questions:

- What representations should be included in a system?

- How can multiple representations be linked to maintain consistency?
- How can multiple representations be organized for more efficient queries?
- How are multiple representations stored in a database?

Data models of multiple representations

Current geographic information systems use only a single model or representation of the world and thus support only a single level of detail. GIS users need to organize multiple representations at different levels of detail and different map scales. While known multi-resolution metrical representations like strip trees are based on the primitive notion of topology used for geographical spaghetti, multiple geometrical representations are based on a complete notion of topology. Ongoing research by Bruegger, Frank and Kuhn at the Maine site includes work on integration of metrical and topological database information, as well as formulating hierarchies between topological cells.

Linkages between multiple representations

Presently multiple representations are stored in our collection of maps. Representations of the coastline of Maine, for example, exist in the many historical renditions and current formats and scales of maps produced by different organizations. Other than sharing an approximate geographic location and a human cognitive association, these representations have no logical and explicit connections. Kate Beard has been exploring methods to explicitly and logically connect different representations of the same geography. Objects in different representations can be assumed to share some metric, topological or attribute information. Current work at Maine by Beard is focused on the formalization of operators which can yield or connect representations with different levels of detail. A research proposal has been submitted to USGS to continue this research effort.

Maintenance of materialized views

Materialized database views are derived from a database and stored in anticipation of providing fast response to future requests for that data. As a view is a derivation from the base data and generally must maintain consistency with it, one or more views over the same base data are a form of multiple representation. Ongoing dissertation research at Maine by Doug Hudson concerns re-establishing the consistency of the views after a change has occurred to the base data, by a technique called "autonomous update computability". Views generated by this technique may be materialized at remote sites, and subsequently updated by sending just the set of base data changes to the remote site. Portions of the view may be autonomously updatable even when the entire view is not and some computational advantage may be extracted from this.

Other research questions

A related set of topics focuses on database queries. At the Specialist Meeting, particular interest was expressed about the stability of query processing in multi-scale database searches. How predictable is the variation in response to a database query made at different levels of resolution? Is the variation more readily predictable for particular data domains, or particular levels of resolution? What range of variation can be considered acceptable for specific types of spatial decision support? To what extent does the architecture of the database impact upon the stability of query processing?

A final topic relates to the determination of the scale of a given database. Depending upon the status of updates, database information should be derived from the amalgamation of scales from all data sources.

The scale of a database is perhaps not meaningfully described by the simple numeric of a Representative Fraction (RF), for example as with USGS "1:24,000 hydrography" (one example from many cited). Asking the scale of a database is in fact making a query about the lineage of the database. The Proposed Standard for Digital Cartographic Data addresses lineage as a data quality issue, but discussions during the Specialist Meeting indicate the issue should be more broadly defined to encompass generation and archiving of geographic information.

3.1.2 Spatial modeling issues

The effects of scale and zoning are being studied in the context of the Modifiable Areal Unit Problem, using multivariate exploration of Buffalo census tract data. Stewart Fotheringham is leading this work at Buffalo, with particular interest in the sensitivity of parameter estimates to changes in the scale at which data are reported. The intention is to discover consistent findings with regard to scale-dependent sensitivities. To date most of the empirical work has been completed and the initial draft of a publishable manuscript

is in preparation. As a side benefit of the work, block group boundaries for the Buffalo metropolitan area have been input to ARC/INFO for subsequent investigation of the urban density gradient over multiple levels of resolution.

A related issue concerns the scale at which various types of geographic process (for example, diffusion and aggregation) are likely to impact upon the structure of geographic features (for example, urban form and growth). At the Buffalo site, Fotheringham is applying fractal and chaotic models to the study of geographic process. Skepticism about the utility of fractal models has been discussed in an editorial by Fotheringham. Michael Batty, who is working with Fotheringham and will be moving to Buffalo from the University of Wales in late 1990, has compared fractal models with Dielectric Breakdown models in the context of modeling urban form.

3.1.3 Generalization issues

Attention in this realm focused upon the nature and display of cartographic data. Generalization for analytical purposes must also drive the research agenda. A major thrust of DeLotto's dissertation, completed at Buffalo, is that attention must be extended to representation and analysis of geographic terrain and higher dimensional objects, and not limited to the current primary focus on line simplification. Assessment of error in generalization continues to be an additional important research priority.

Clarifying the concept of resolution

The need to derive a more comprehensive and flexible definition of resolution forms an important research direction. Variations in feature resolution that may occur within a single data set (e.g., US county boundaries) must be explored to improve the efficiency and accuracy with which they are utilized for statistical and cartographic applications. Simplification algorithms must be designed that self-adjust according to the local resolution of objects. Resolution of enumerated data is a topic of some interest to researchers in spatial statistics and spatial analysis, where it is commonly referred to as the "Modifiable Areal Unit" problem. Research in this area was discussed in Section 3.1.2.

Digital feature description

A second item placed on the research agenda involves formalization of digital feature description and categorization models. Buttenfield argues in her paper for the upcoming NCGIA-sponsored Syracuse symposium that digital models must accommodate the complexity of compound and hierarchical objects, and be sensitive to those aspects of geometry that may change with scale. The work of Jasinski, published as a Center Technical Report, shows that measures of feature geometry and complexity must be evaluated in terms of robustness and redundancy. In a paper based on work under this Initiative, and published in *Cartographica*, McMaster studies conversion between digital models in terms of error propagation, associating particular models with specific generalization operations. Refined understanding of the properties and operations associated with a model or class of digital models (for example, terrain models) will improve decisions about which generalization algorithms will be appropriate for a particular model.

This research issue is not limited to simplification of map features, as Walters argues in a book chapter based on work under this Initiative. Problems of consistency in attributes and feature codes affect the transfer of data between federal agencies, and attest to the need for consistent attribute assignments in digital databases. Research on the logic of feature code assignments was ascribed a high priority at the Specialist Meeting. Categorization of digital data may be improved by attending to cognitive aspects of feature categories, in addition to the current emphasis on accuracy and (perceptual) recognizability so prevalent in database and cartographic research. There are implications for the logic of spatial query languages as well, and herein lies a link between Initiatives 2 and 3.

Formalizing rules for map generalization

A third set of research topics involves rules for map generalization that must be defined explicitly to automate the cartographic process. Mark argues in his paper for the Syracuse symposium that rules for setting and modifying tolerance values in feature or object simplification cannot be developed without improving techniques for encoding and manipulating those features and objects. A project funded by USGS (Buttenfield, PI) is currently addressing tolerance modification for linear features at Buffalo; results of the first phase of research are just complete and will be reported by Buttenfield at the Syracuse symposium.

Rules for symbolic transformation of map information must also be derived and evaluated. For example, the decision to collapse an urban area down to a point is determined in part by scale, map purpose, and by dimensionality of the object at its source scale. Other symbolic operations such as coarsening of a categorical coverage, displacement of features during map reduction, etc. are understood only intuitively. Issues of conflict resolution must be defined and prioritized. A comprehensive typology of generalization operations must be formalized, and this marks an area of generalization research already being undertaken outside the NCGIA.

A symposium to address substantive issues prerequisite to the development of a knowledge base for cartographic generalization is planned for April 1990, to be jointly sponsored by NCGIA and Syracuse University. The symposium will be hosted by Robert

McMaster (Syracuse) and Barbara Buttenfield (Buffalo). Participants have been invited from universities in North America, Europe and New Zealand and from the private sector (Intergraph, PRIME-Wild, and The Analytical Sciences Corporation). Each participant will present a paper focusing on a specific issue (for example, conflict resolution, or preserving the balance between artificial intelligence and amplified intelligence) that must be resolved in order to implement a knowledge base for automated generalization.

NCGIA Participants (papers abstracted below):

M. K. Beard, Maine
B.P. Buttenfield, Buffalo
W.A. Mackaness, Buffalo
D.M. Mark, Buffalo

Other Academic Participants:

Marc P. Armstrong, University of Iowa
Herbert Freeman, Rutgers University
Mark S. Monmonier, Syracuse University
Jean-Claude Muller, ITC, Enschede, The Netherlands
Bradford Nickerson, University of New Brunswick, Canada
Tim S. Nyerges, University of Washington

Federal Agency Representative

Keven Roth, National Mapping Division, USGS

Private Vendors:

Gail Langran, Intergraph Corp., Seattle, Washington
Stuart Shea, The Analytical Sciences Corporation, Reston, Virginia
Robert Weibel, PRIME, Zurich, Switzerland

The symposium agenda has been presented to several book publishers to solicit interest in the publication of papers, with the intention of publishing a proceedings volume by the end of 1990. Additionally, two special paper sessions are planned for the Association of American Geographers Meetings (Toronto, April, 1990). Both sessions will include presentations condensed from the rule base symposium discussed above.

3.1.4 Bibliography

Incorporated into the Initiative bibliography published by the Center as a Technical Report are citations from journal articles, monographs and books, and technical reports on research in computer science, cartography, and spatial modeling. The common thread is the exploration of methods by which to generate and utilize a single detailed database to produce and analyze graphical representations at many different scales and resolutions. Implications for map representations include problems of appropriate simplification and symbolization of both features and attributes, as well as problems of maintaining consistency throughout database update.

The bibliography is intended as a tool for the general community, to provide an overview of research in a topical area that has to date not been identified in its own right. It became clear as the literature search was undertaken that the volume of research that has been reported is being catalogued under categories that are quite diverse, and little if any centralization to the topic now referred to as "Multiple Representations" has penetrated keyword generation to date. Publication of this bibliography should reduce redundancy of research efforts and alert individuals who do not commonly interact about the communality of efforts to understand the nature of things that change with scale, encourage interchange between researchers in many disciplines, and maintain currency in a field that is to date only crudely defined.

Information sources for this bibliography include the usual bibliographic tracing through published research reports and general reference. Online searches were performed through computerized bibliographic data bases and CD ROM archives of the Science Citation Index, INSPEC (Electrical and Electronic Engineering Abstracts), GEOBASE (GeoAbstracts) and GEOREF (Bibliography

and Index of Geology). Several government abstracts and archives have been searched, including Government Reports Announcements and Index (GRA&I), Scientific and Technical Aerospace Reports (STAR), Energy Research Abstracts (ERA) and Monthly Catalog of US Government Publications.

3.1.5 Multi-agency, multi-scale database

The discussions for the data base focused in part on the need to create a standard of reference upon which to benchmark generalization algorithms and database search and query operations. Without a standard data set, it remains difficult to compare the results of various algorithms and to improve our understanding of the nature of digital representations that may change substantially depending upon the resolution at which they are collected, archived, and mapped. A standard data set based on federal products that could be placed in the public domain would provide such a standard. Additionally, this data would provide a useful teaching aid, to demonstrate the use of federal agency data sets, and to explore the transferability of data from one agency format to another. Finally, it would provide researchers and students alike with the chance to work with multiple digital views of the same area, produced according to the various digital specifications of participating US federal agencies.

The NCGIA was a logical agent to represent the academic perspective in the effort, as part of its mandate is to remove impediments to pursuit of research and education in GIS, and to support similar research and educational efforts at other academic institutions through its outreach program. Academic researchers do not always have resources to receive and process multiple digital data sets from federal data producers. Federal agencies do not often archive small data sets that can be compared with other data sets from other agencies.

Five sets of data are planned, containing data from several agencies. Data to be included in the first set (Lee County, in southern Florida) will be provided by US Census, USGS, and NOS. As time and funding permit, other data sets will follow. Each data set is intended to comprise 10 - 20 floppy diskettes and will have a small ASCII instruction file explaining how to unpack and load the data set. Data documentation will be a part of the packed data on the disks. NCGIA will disseminate the data sets in the public domain on a cost recovery basis, in both DOS and Macintosh formats. Dissemination is currently planned to begin in early 1990, pending final release of TIGER files. Further details on the database project will be presented in an NCGIA Technical Report, and a journal article will be submitted to disseminate news on the existence and availability of the data set to the widest possible community.

3.2 Annotated list of publications

3.2.1. Articles in refereed journals

Batty, M. Generating urban forms from diffusive growth. *Environment and Planning A* (to appear).

A formal model of diffusion-limited aggregation is presented and used to generate a continuum of urban forms from linear to concentric. A brief derivation of the continuum is presented, and related to potential theory. The forms produced by the model are subject to fractional power laws which relate occupancies of sites and densities to distances. Fractal dimensions can be derived from these fractional powers and conventional and fast methods of estimation are introduced.

Batty, M., P.A. Longley and A.S. Fotheringham. Urban growth and form: scaling, fractal geometry and diffusion-limited aggregation. *Environment and Planning A* 21:1447-72 (1989).

(see section 1.2.1 above)

Beard, M.K. Review and recommendations for automated generalization. *International Journal of Geographical Information Systems* (under consideration).

The history of automated generalization now spans approximately twenty-five years, but solutions developed to date have not yielded satisfactory results. It seems timely to evaluate reasons for this limited success. Limitations of early computer technology were partially responsible, but other factors are hampering progress. This paper investigates impediments, including difficulties in formalizing the process, oversimplification of the process, the cartographic mindset, and hardware and software limitations. Recommendations for future progress are presented.

Bruegger, B.P. and W. Kuhn. Multiple topological representations. *International Journal of Geographic Information Systems* (under consideration).

Topological representations are considered an important part of a modern GIS. However, they support only a single level of detail. This paper proposes a system of multiple topological representations (MTR) which includes several representations at different levels of topological detail. The representations are linked by hierarchical relations between topological cells. Topological reasoning at small scales can be considerably improved when using MTR. This extends the range of scales at which a GIS can be used.

Buttenfield, B.P. Scale-dependence and self-similarity in cartographic lines. *Cartographica* 26(1):79-100 (1989).

This paper provides a typology of two classes of geometry for cartographic lines, based on the well-known Richardson plots. The first class contains objects having self-similar geometry, features whose structural characteristics are replicated either precisely or statistically with changes in scale. Self-similar features are currently described by fractal models, which some argue are appropriate for all cartographic objects. The fallacy of this statement is demonstrated. The second class of features is fully scale-dependent, and contains cartographic objects whose geometry varies distinctly with changing scale. Both models are described and applied to examples of digital line features, to demonstrate their worth in encoding and preserving particular types of cartographic detail during automatic generalization.

Fotheringham, A.S. Guest editorial: What's the fuss about fractals? *Environment and Planning A* (to appear).

(see section 1.2.1 above)

Fotheringham, A.S., M. Batty and P.A. Longley. Diffusion-limited aggregation and the fractal nature of urban growth. *Papers of the Regional Science Association* 67:55-69 (1989).

(see section 1.2.1 above)

McMaster, R.B. The integration of simplification and smoothing algorithms in line generalization. *Cartographica* 26(1):101-21 (1989).

The cartographic generalization of vector data in digital format involves six distinct processes, including simplification, smoothing, enhancement, displacement, merging, and omission. This paper proposes a conceptual model for the geometric interaction of these components. It is proposed that sequential processing of line data based on this model will improve the quality of the original digital information, minimize database storage and database error, and produce aesthetically acceptable generalizations at greatly reduced scales.

3.2.4 Articles in other outlets

Batty, M. Cities as fractals: simulating growth and form. In R.A. Earnshaw, editor, *Fractals and Chaos*. Springer-Verlag, New York (to appear).

The morphology of cities bears an uncanny resemblance to dendritic clusters of particles which have been recently simulated as fractal growth processes. The analogy is explored, presenting deterministic and stochastic models of fractal growth, and suggesting how these models might form an appropriate baseline for urban models. In particular, the diffusion-limited

aggregation (DLA) models and dielectric breakdown (DLM) models are outlined. Comparisons are made between simulated clusters and the form of medium-sized towns. Using the DBM formulation to illustrate constraints and distortions, the model is applied to the town of Cardiff, Wales.

Beard, M.K. Design criteria for automated generalization. *Proceedings International Cartographic Association Conference* (Budapest, August 1989) (to appear).

The history of automated generalization spans approximately 25 years, but solutions developed to date have not yielded satisfactory results. It seems timely to evaluate reasons for this limited success. Although limitations of early computer technology are partially responsible, other factors are hampering progress, including difficulties in formalizing the process, the graphic map and the physical mindset, and persistent hardware and software problems.

Beard, M.K. Constraint-based transformations in map generalization. *Proceedings of the Symposium Towards a Rule Base for Map Generalization* (Syracuse, New York, April 1990) (to appear).

Generalization has traditionally been practiced as an individual artistic skill and as such it lacks a universal and consistent definition. Although some general procedural steps can be identified, the diversity of cartographers' approaches to the task has defied formalization and hindered automation. Formalism requires specification of some well-defined and unambiguous rules, but anticipating rules to cover every possible condition is not practical. This paper explores the specification of generalization as a set of constraint-based transformations, and develops in the following sequence: (1) specification of an exhaustive set of operations to abstract the spatial and nonspatial aspects of a representation; (2) description of the essential information needed to support each operation in the set; (3) development of a model to support the information as defined and facilitate operations in the set; and (4) design of an interface allowing users to specify conditions (constraints) which must remain invariant during generalization.

Bruegger, B.P. and A.U. Frank Hierarchical cells complexes. *Papers and Proceedings of Applied Geography Conferences* (Binghamton, NY) (1989) pp. 238-40.

Multipurpose GIS offer considerable cost savings due to the sharing of data acquisition and maintenance efforts among several user groups. The paper demonstrates how a single topological representation impedes the implementation of such a system by causing response times to be unacceptably high when relatively large spatial objects are involved. In order to solve this problem, additional representations with reduced detail have to offer higher level views of the spatial objects. Lattice structures under spatial inclusion and hierarchies over topological cells are shown as examples for such higher level representations.

Bruegger, B.P. and A.U. Frank. Hierarchies over topological data structures. *Proceedings, ASPRS/ACSM Annual Convention* (Baltimore, MD) vol.4 (1989), pp. 137-45.

The paper describes a system of multiple, ordered topological representations connected by hierarchical relationships between topological cells. The lowest level representation organizes all spatial objects in detail, while the higher order representations contain only large and important objects represented with less detail. This structure makes the implementation of multipurpose GIS possible, where very small (e.g. parcels) and very large spatial objects (e.g. nations) co-exist. The response time to queries is independent of the size of the objects in contrast to present topological data structures. The spatial objects can be displayed in several resolutions. (A similar paper was presented at GIS/LIS '89 in Orlando, FL).

Bruegger, B.P. and A.U. Frank. Hierarchical extensions of topological data structures. *Proceedings, 17th Annual International Surveyors Conference (FIG)* (Helsinki, Finland) (to appear).

Humans maintain multiple mental models to reason about the world. They use models appropriate to given situations in order to avoid getting lost in detail. GIS normally feature only a single representation for their reasoning process in a given domain. This paper presents multiple topological representations designed to improve reasoning efficiency of GIS.

Buttenfield, B.P. Rules for automating feature distinctions. *Proceedings of the Symposium Towards a Rule Base for Map Generalization* (Syracuse, NY, April 1990) (to appear).

Automatic methods are presented to improve cartographic line simplification, and specifically to determine locations in a coordinate file where the geometry of a line feature changes, to enable automatic tolerance value modification. Tolerance values must be modified periodically to insure both accuracy and recognizability of graphic details on a generalized map. At present, decisions about where to adjust tolerance values are made manually, and form an expensive bottleneck to map production for government and commercial organizations. Any automatic solution requires formalized description of the amount and type of details that occur along the extent of the digital file. This is complicated by the need to accommodate scale-dependence in the formal descriptions.

Buttenfield, B.P. and J.S. DeLotto. Multiple Representations: Report on the Specialist Meeting. Technical Report 89-3, National Center for Geographic Information and Analysis, Santa Barbara, CA.

Buttenfield, B.P., J.S. DeLotto and J.V. McKinney. Multiple representations: a bibliography. Technical Report 89-11, National Center for Geographic Information and Analysis, Santa Barbara, CA.

DeLotto, J.S. Elevation matrices as a data source for automated terrain classification. *Proceedings, International Cartographic Association Conference* (Budapest) (to appear).

Parametric descriptions of terrain geometry have been proposed by a number of authors as a path to automating terrain classification, including measures of slope angle, aspect, hypsometry, and local convexity. Gridded topographic data (DTMs) remain the most widely used format for digital terrain, but are limited as a source for quantitative interpretation of terrain. Because sampling for production of DTMs does not vary with complexity of terrain, DTMs often contain inadequate representations of geomorphically significant features. A number of parameter extraction techniques are compared to demonstrate this limitation.

DeLotto, J.S. The role of scale in automated terrain classification. Master's thesis, Department of Geography, University at Buffalo (1989).

Recent efforts to integrate techniques of automated terrain classification into geographic information systems are hindered by inadequate application of existing knowledge in fields of numerical taxonomy, image processing, and image understanding. The thesis reviews principles of terrain classification and geomorphometry, identifies problems that may occur when classification is performed from gridded elevation data, and proposes techniques to accommodate the scale dependent nature of topography.

Jasinski, M.J. A comparison of complexity measures for cartographic lines. Technical Report 90-1, National Center for Geographic Information and Analysis, Santa Barbara, CA. (Master's Thesis, Department of Geography, University at Buffalo).

Existing geometric measures commonly applied to cartographic line features are examined for statistical redundancy. The goal is to determine a quantifiable definition of "line complexity", or specifically to discover the components of a quantifiable definition. It is hypothesized that components include aspects of distance, angularity, and density of detail, and that redundant measures are being applied by various researchers. A test data set comprised of both cultural and naturally-occurring line features is measured and compared statistically to evaluate the geometric measures.

Mackness, W.A. Application and evaluation of map generalization techniques. *Proceedings of the Symposium Towards a Rule Base for Map Generalization* (Syracuse, NY, April 1990) (to appear).

To date, generalization techniques have been researched in isolation from one another, and appraised using mathematical techniques. It is becoming increasingly apparent that the combination and degree of application of these techniques are dependent on one another, and need to reflect the geographical character of features, the map type, and user knowledge, and also to encompass the objectives of aesthetics. This paper attempts to summarize these qualitative attributes using an "attribute-rose diagram". Attributes typically used to describe a map are represented by spokes. The distance up the spoke is proportional to the strength with which that adjective describes the map. The problem thus becomes one of pattern recognition, namely associating generalization techniques according to the shape of the rose.

Mark, D.M. An object-oriented approach to the implementation of phenomenon-based cartographic generalization. *Proceedings of the Symposium Towards a Rule Base for Map Generalization* (Syracuse, NY, April 1990) (to appear).

This paper proposes that support of phenomenon-based ("geographical") generalization requires three high-level classes of objects. One class would be made up of geographic objects, features of the real world such as shorelines, settlements, roads, lakes, etc. Another class would include cartographic objects, such as the symbols and features that are shown on topographic maps. The third class includes data-structure objects, such as points, pixels, line segments, images, etc. If the classes are properly linked, then generalization rules may be applied in one domain, and the effects propagate through to the others.

Mark, D.M. Conceptual basis for geographic line generalization. *Proceedings, AUTO-CARTO 9* (Baltimore MD, April 1989) pp. 68-77.

Line generalization is an important part of any automated map-making effort. Generalization is sometimes performed to reduce data volume while preserving positional accuracy. However, geographic generalization aims to preserve the recognizability of geographic features of the real world, and their relations. This essay discusses geographic generalization at a conceptual level.

Walters, D.K.W. Discrete contour descriptors. In J.C. Simon, editor, *From Pixels to Features*. North Holland Elsevier (1989) pp. 287-298.

A set of discrete contour descriptors is proposed. The descriptors are able to describe discrete contours, such as those found in digitized maps, using parameters which are independent over the similarity transformations of scaling, rotation and translation. As the descriptors are based on discrete geometry, they do not have the implementation problems of descriptors based on continuous geometric concepts. In addition, the descriptors are perceptually valid: contours which humans perceive to be similar have similar descriptors. The descriptors are much simpler than spline-based or related descriptions, and thus can lead to more compact representations that maintain all of the perceptually relevant information.

Weibel, R. and J.S. DeLotto. Automated terrain classification for GIS modeling. *Proceedings, GIS/LIS '88* (San Antonio, TX) vol. 2 (1989), pp. 618-27.

Automated terrain classification involves the partitioning of an area into homogeneous topographic regions through quantitative interpretation of a digital terrain model. By evaluating geometric parameters extracted from raw elevation values of a DTM, it is possible to characterize the nature of topography in either general terms, i.e., roughness, or in terms of a specific application, i.e., hydrology or geomorphology. A recent interest in terrain classification can provide meaningful input to modeling and GIS processing tasks.

4. Initiative 4: The Use and Value of Geographic Information

4.1 Narrative summary of progress to date

The Initiative 4 research objectives were originally defined, in part, as:

1. Identify primary and subsequent users of spatial information, and determine the value of such information;
2. Develop and test models of the decision-making process regarding land use, focusing on the role of geographic information; and
3. Evaluate the direct and indirect benefits of GIA/GIS.

The work plan which evolved from the Specialist Meeting held in May of 1989 consisted of three parts:

1. Develop a taxonomy on the use and value of geographic information;
2. Investigate a method for assessing the value of using geographic information; and
3. Investigate the factors affecting the adoption and diffusion of geographic information technologies.

The overall objective of the work plan is to develop fundamental integrated frameworks and tools for evaluating the use and value of geographic information. The general form of the research in developing these frameworks is to conduct, in an iterative manner, investigations into the actual use of geographic information in selected situations. The vehicles used are in-depth case studies, limited surveys, and finally broadly based surveys of geographic information users, including both users and non-users of geographic information systems.

The following pages describe research accomplishments to date, other tangible products of the initiative, and plans for the remainder of the project. Research is being conducted on this Initiative at the Maine and Buffalo sites. It is intended that few NCGIA resources will be devoted to this initiative after December 1990, although derivative research is likely to continue after that date and to be incorporated into Initiative 9 on Institutions Sharing Geographic Information.

4.1.1 Taxonomy of geographic information use

The objective of the taxonomy is to provide a framework for the study of the use of geographic information and a framework for identifying the benefits of such use. The definitions of a taxonomy (or taxonomies) have been summarized by Nancy Obermeyer in a paper submitted to the *International Journal of GIS*. Useful taxonomies classify objects on the basis of one or more characteristics (or attributes) for the purpose of identification and description in unambiguous terms. Taxonomies can be useful in the systematic study of objects provided the classifications allow each instance of the object under study to be observed and measured in a manner which facilitates comparison between observations and allows generalization of a set of observations to recognize significant patterns.

A taxonomy should reflect the purposes or objectives of the activity for which it is created. In this instance the objective is to characterize the use of geographic information in decision-making. A second objective is to identify and quantify the value, or benefits of the use of geographic information to one or more parties (preferably in a price-based form).

One taxonomy developed by the research is defined in terms of the set of questions for which answers are sought to further our understanding of the use of geographic information and the value (or benefits) derived from such use. Each question represents a single "view" of the use of geographic information. These "views" reflect not only what we would like to know at the present time, but also recognize the current state of development of systematic procedures for the use of geographic information (GIS) and the constraints imposed by the current level of development. The taxonomy represented through these questions is not representative of

any complete taxonomy of geographic information use (or GIS). The Obermeyer paper argues that other taxonomies (or views) designed for different objectives may be constructed and would be valid for a broader understanding of the role of GIS and geographic information in a societal context. The "views" presented here are restricted to the purpose of the investigation and constrained by anticipated information availability from experts, users, and others currently involved with GIS.

One relevant set of questions regarding the use of geographic information is as follows:

1. What are the uses of geographic information (type of task: inventory, process control, browse or non-systematic search, analysis for structured problems, or spatial decision support system)?
2. What are the primary functional capabilities required (display, query, measurement, map overlay, calculations, statistical analysis, modeling)?
3. Why is geographical information useful?
4. Who is the user (analyst, middle manager, decision-maker, general public, organized public opinion group)?
5. How are spatial data handling procedures organized (tasks, applications, modules)?
6. What is the form or format of the geographic information product (map, other graphic, text)?
7. What is the level of use (policy, management, operations)?
8. What is the decision-making process stage (problem identification, development of alternatives, selection of solution, implementation programming, monitoring of results)?
9. What is the type of organization using the information (public, private, not-for-profit)?
10. What is the application area (e.g., local government, forestry, etc.)?
11. What is the role of the organization in the handling of geographic information (user, supplier, combination)?
12. What quantitative models are used to produce the needed information (this is a further sub-division of modeling in 1 above, and would include entries such as location/allocation, gravity-based models, entropy models, etc.)?

Answers to the above set of questions form an integrated initial taxonomy for the purpose of investigating the use and value of geographic information. That is, answers to a comprehensive set of standard questions may provide a rational and highly useful means of classifying geographic information uses and users.

In contrast to this approach to taxonomy development is an approach which draws from taxonomic work in the natural sciences and geographic theory for its foundations. A preliminary taxonomy of geographic information and its uses has been developed from such foundations and has become the basis of an iterative process that is seeking the input of geographic information experts and surveying users to achieve consensus on an appropriate taxonomy.

The two taxonomy approaches being pursued are supportive of each other and each is being further developed through survey methods. When fully investigated they may eventually be combined into a single comprehensive framework or may continue to exist as independent but complementary frameworks.

Preliminary results and a preliminary taxonomy were presented by Obermeyer at the GIS/LIS '89 meeting in Orlando. In addition, two companion papers presenting different views and approaches to structuring a taxonomy of geographic information uses have been submitted by Obermeyer and Calkins to the *International Journal of GIS*. A proposal for further development has been submitted to several federal agencies.

4.1.2 The use and value of geographic information

It has been recognized that a realistic determination of the value of geographic information can only be accomplished through a better understanding of the use of the information and the factors which influence the ability of decision-makers to use the information. A paper by Holly Dickinson submitted to the *International Journal of GIS* shows that principles of valuation of information in other

disciplines are based on either establishing a monetary value for the information or defining benefits as "intangible," meaning an adequate measure cannot be defined and/or measured. Prior studies of the value of geographic information have identified benefits as tangible or intangible. The tangible benefits are mostly direct cost or staff time savings while the intangible benefits are listed as better information, better planning and improved decision-making.

Significant benefits from the use of geographic information are contained within the intangible category. The objective of attaching a monetary value to these types of benefits can only be reached through a better understanding of actual use patterns in decision-making contexts and identification of the factors facilitating or inhibiting such use. Thus the research plan for this Initiative first focuses on the task of describing and understanding, in detail, the use of geographic information in decision-making.

Various methods for describing information use in a formalized model have been reviewed by Benwell, Sallis and Firms in unpublished work. Their preliminary conclusion is that most popular methods for describing information flows have one or more drawbacks (such methods as flow charts, data flow diagrams, CPM or PERT diagrams, etc.). Benwell has suggested that a variant of Petri nets (colored Petri nets) might be suitable for the purpose of modeling the use of geographic information. Formal modeling of use would then provide a framework to attach value measures to the geographic information. This notion is currently being tested at Buffalo using two in-depth case studies of GIS: the use of geographic information by the Washington State Department of Natural Resources (forestry application), and a local government geographic information system (Town of Amherst, New York) involving eight departments of town government (planning, building, assessment, recreation, engineering, highway, police and fire dispatch). In each of these case studies, selected applications using geographic information are being modeled using the Petri net framework. Successful modeling of selected applications will be followed by: 1) attaching measures of cost and value to the Petri net solutions to determine if reasonable measures can be estimated for the value of geographic information; and 2) testing with other applications where available documentation exists (extensive documentation of local government applications has been contributed by Environmental Systems Research Institute of Redlands, California for this study).

A journal article in anticipation of the value research for the initiative was published by Dickinson and Calkins (1988) in the *International Journal of GIS* with a comment piece following in a later volume. Foundation principles and ongoing work were described in a paper at the GIS/LIS '89 meeting in Orlando and an annotated bibliography on the value of information has been prepared and published by the NCGIA. A journal article reviewing the methods for establishing the value of information from the fields of economics, management science and information science has been submitted by Dickinson to the *International Journal of GIS*.

4.1.3 Adoption and diffusion of geographic information technologies

The primary objective of this research is to formally identify and verify factors and processes influencing GIS users, both public and private, to adopt geographic information innovations, i.e., innovations such as digital data sources, standards, and special purpose software. From the identified factors and processes the goal has been to develop an analytically-based diffusion model that can predict whether a geographic information innovation in its current state of development is likely to be widely adopted. If the likelihood is non-adoption, the model should identify those technical, economic, institutional and social variables which are the most critical in impeding the adoption process. The research team expects that critical variables for adoption/non-adoption will also emerge as significant components of the Petri nets representing geographic information use in existing GIS.

Two models for describing diffusion of graphic information innovations, content and process models, are being used in this research. Data is being collected from a limited number of local government agency case studies to identify those diffusion model variables and processes which are significant in local level GIS use environments. A follow-up cross sectional survey of a larger body of local GIS users, both public and private, will allow statistical testing of the derived models.

Initial results of theoretical investigations were presented by Gary Jeffress, a PhD candidate at Maine, at GIS/LIS '89 in Orlando as well as by Harlan Onsrud at the same conference. Case study and survey results will not be published until the research work plans are largely complete although a foundation article titled "Diffusing geographic information innovations" will soon be submitted by Onsrud to the *International Journal of GIS*.

In addition to the on-going case studies, funding proposals have been submitted to several federal agencies to study the diffusion of federally produced digital data, software and standards. Two additional proposals to study the diffusion of infrastructure management systems have been submitted to AM/FM International and the Engineering Directorate of NSF.

4.1.4 Other activities

The research has led to two PhD theses which are well underway; "The use and value of land ownership information" by Gary Jeffress, University of Maine, and "Petri net models for examining the use of geographic information in decision making: a method to

assess costs and benefits" by Holly Dickinson, SUNY-Buffalo.

An international research seminar on "Land Information Management, Land Tenure, and Cadastral Systems" was held on July 27-29, 1989 at the University of Maine. The seminar was cosponsored by NCGIA and the Atlantic Institute and many of the presentations directly related to Initiative 4. The seminar featured presentations by invited speakers from the World Bank and universities representing several nations.

A session consisting of four consecutive papers on various research aspects of the use and value of geographic information was presented by NCGIA researchers at GIS/LIS '89 in Orlando, Florida.

4.2 Annotated list of publications

4.2.1 Articles in refereed journals

Calkins, H. A taxonomy for surveying the use and value of geographic information. *International Journal of GIS* (under consideration).

Dickinson, H. Methods for establishing the value of information: application to the value of geographic information. *International Journal of GIS* (under consideration).

Frank, A.U. The geographic information system and its use for valuation. *Property Tax Journal* 8:85-98 (1989).

In order to evaluate and assess real estate property information properly, it is necessary to describe the property. The major part of this information is necessarily space-related, that is, related to the land it describes. The uses of GIS could be classified as planning or operational; these uses have very different requirements for the quality of the data in the system. The two primary architectures for GIS, file and database oriented, can be linked to these use classes. Systems for valuation purposes should be database oriented.

Hintz, R.J. and H.J. Onsrud. Upgrading real property boundary information in a GIS. *Journal of Urban and Regional Information Systems* (to appear).

One difficult issue facing GIS developers in the US today is the current inability to create spatially accurate, legally supportive and operationally efficient land ownership data bases. Solutions providing strong legal foundations for GIS are not simple. This paper describes technology for establishing a measurement based management system at the local government level. Sophisticated surveying computations, least squares analysis, statistical techniques, and blunder detection methods have now been largely automated. Tools which were available previously to only highly specialized surveying experts are now potentially useable by surveying technicians. Through use of these powerful tools, maintenance over time is readily achievable for cadastral measurements in a GIS database.

Jeffress, G. The use of GIS in economic policy analysis. *Northeast Journal of Business and Economics* (under consideration).

Analysis of spatially derived data can be useful in isolating the locational effects of policies and programs. Economic effects due to location can be identified and presented in graphical form. Policy scientists make extensive use of spatial data; they can more effectively communicate this data through geographical analysis and presentation. Thematic maps showing the outcome of policy research and analysis are more readily understood than tabular and statistical presentation of spatially referenced economic data. This paper describes the nature of GIS and its role in economic policy analysis.

Jeffress, G., R. Hintz and H. Onsrud. An automated measurement management system: a useful tool for surveyors and digital cadastral data base managers. *The Australian Surveyor* (to appear).

The increased use of digital data recorders by the surveying profession is linked to the creation of digital cadastral data bases (DCDB). The management of digital measurements, whether made in the field or on a digitizing tablet, is fundamental to the production of the spatial information contained in both survey plans and digital maps. This paper describes the application of a measurement based management system for use by the surveyor or the manager of a DCDB. Sophisticated survey computations, least squares analysis, constraints by variance, statistical techniques, and blunder detection have now been conveniently packaged into an automated tool box.

Obermeyer, N.J. Regional equity in turbulent times: the experience of the Regional Transportation Authority of Northeastern Illinois. *Applied Geography* (to appear).

This paper examines regional equity within the context of a changing demographic environment and shifting political power structures and relationships. Taking as an example the Regional Transportation Authority (RTA) of northeastern Illinois (which includes Chicago and its suburbs), the paper analyses the sources of revenue and the distribution of public transportation services and their impacts on regional equity and the perception of regional equity among the groups involved. The evidence from this case suggests that regional equity, while frequently cited as a public policy goal, is not objectively quantifiable. The resolution of specific regional equity problems is achieved not through the development and refinement of mathematical calculations, but rather through negotiations among the parties involved.

Obermeyer, N.J. Bureaucratic factors in the adoption of GIS by public organizations: preliminary evidence from public administrators and planners. *Computers, Environment and Urban Systems* (to appear).

In this paper, the author argues that because of organizational reliance on "standard operating procedures" on the one hand, and professional training and socialization on the other, public organizations tend to favor the status quo, while public administrators screen out the geographical components of professional tasks and rely on nongeographic information systems. This argument is based on a theoretical understanding of bureaucracy, organizational decision-making, and the search for information used by organizations in the decision-making process. Support for this argument is based on a preliminary study of the use of geographic information by public administrators and planners, as evidenced by the presence of maps, within the professional publications of the American Society for Public Administration and the American Planning Association and the American Institute of Certified Planners.

Obermeyer, N. The potential for institutional information sharing in GIS: an organizational-managerial perspective. *International Journal of GIS* (under consideration).

As development of geographic information systems progresses, and technical problems are overcome, new problems arise. This paper explores the organizational and managerial roots of difficulties in sharing data bases. Historically, a combination of organizational structure and operations has resulted in the fragmentation of work projects into individual tasks. The resulting deskilling of the work force, and growing powerlessness among manual and clerical workers have been acknowledged as negative outcomes of such fragmentation. This paper identifies three means by which alliances may be formed and information may be shared: appeals to professionalism, coercion, and bargaining. The author proposes a theory based on the relative power of the participants to predict which of these three strategies will be used.

Obermeyer, N.J. A systematic approach to the taxonomy of geographic information. *International Journal of GIS* (under consideration).

This paper represents an effort to lay groundwork for the development of a taxonomy of geographic information and its uses that draws from taxonomic work in the natural sciences. Beginning with a discussion of some of the pitfalls of taxonomy development, the author proposes a framework for a preliminary taxonomy of geographic information and its uses. This preliminary taxonomy will become the basis of an iterative process that will seek the input of geographic information experts and survey users of geographic information.

Onsrud, H. The Land Tenure System of the United States. *Forum: Zeitschrift des Bunds der Öffentlich Bestellten Vermessungsingenieure* (Jan. 1989) pp. 23-28.

The purpose of this paper is to briefly describe the land conveyancing system used in the US. In addition to its operational aspects, several shortcomings of the system are described. Although the US conveyancing system at first observation appears to be far inferior to the land registration system used in the Federal Republic of Germany, the US system supports certain policies which are considered important within the US social framework. It is submitted that recent technological advances may be moving US jurisdictions towards a conveyancing system which will have the same overall effect as a land registration system.

4.2.3 Articles, chapters, and monographs in other refereed outlets

Calkins, H. GIS and public policy. In D. Maguire, D.W. Rhind and M.F. Goodchild, editors, *Geographical Information Systems: Principles and Applications*. Longmans (to appear).

Obermeyer, N.J. Bureaucrats, clients and geography. Research Paper No. 216, University of Chicago Geography Research Papers, 135pp.

This study explores the relationship between public organizations, their clients, and policy decisions relating to the location of major facilities. Most research on the geographical siting of facilities assumes that decision-makers are guided by rational and predictable factors, whereas work on organizational behavior points to the importance of seemingly irrational influences, especially noneconomic constraints. In examining one such constraint, the influence of client groups on a national regulatory agency, this study seeks a better understanding of the varied nature of administrative decision-making.

4.2.4 Articles in other outlets

Beard K. Designing GIS to control the misuse of spatial information. *Proceedings, URISA '89* (Boston, MA) Vol. 4, pp. 245-55.

Commonly recognized map errors include those associated with data collection (source error) and the processing of data for map compilation (process error). Another error component, use error, is defined and added to the typology. This paper argues that without attention to use error, large investments to reduce source and process error may be wasted. Traditional representation of spatial information on paper maps has limited our ability to control this form of error in any significant way. While the misuse of maps cannot be entirely avoided, computer technology offers a possibility for limiting the opportunities for misuse. This idea is explored by examining ways in which maps are misused, and from this exploration, formulating geographic information system design strategies that may counteract the potential for misusing spatial information. (This paper also appears in *Proceedings, AutoCarto 9* (Baltimore, MD) pp. 808-17 (1989)).

Dansby, B. and H. Onsrud. Geographic Information Systems - a map for the future. *Probate and Property*, Section of Real Property, Probate, and Trust Law, American Bar Association, (1989) pp. 20-27.

As members of the real estate conveyancing community, lawyers have an important role to play in developing standards and practices needed to improve land records. To be of most use, the data collected by local governments needs to be referenced to individual land parcels. The information on individual parcel boundaries in a GIS, referred to as the cadastral layer, forms a bridge between legal ownership rights in land and the physical location of those rights on the earth. Thus, the cadastral layer is often referred to as the legal foundation for any widely useful GIS. Parcel boundaries, and therefore the cadastral layer, are legally defined by the conveyancing documents recorded in connection with real estate transactions. Because of the primitive state of conveyancing records systems in almost all jurisdictions in the US, the GIS currently being implemented are at risk of having inadequate cadastral foundations.

Dickinson, H. Bibliography on the economics of information. Technical Report 89-8, National Center for Geographic Information and Analysis, Santa Barbara, CA.

Epstein, E. and T. Duchesneau. The impact of information science on conflict. *Proceedings, URISA '89* (Boston, MA).

Geographic and land information system technology is often adopted and used for the efficient generation and distribution of traditional spatial information products and services. These actions are predictable given the responsibilities and burdens on public officials and private citizens who make growth and resource management decisions. These decisions are often contentious, with the controversy extending to the data and information used in the decisionmaking, and lead to new perceptions about the nature and utility of the system and its products. These perceptions are developed by the parties to the growth and resource management struggles, and provide important direction for system designers and builders.

Jeffress, G.A. and H. Onsrud. Does LIS/GIS have a role in economic growth? *Technical Papers, ASPRS/ACSM Annual Convention* (Baltimore, MD) vol. 4 (1989) pp. 285-92.

Much has already been documented on the conceptual models, efficiencies and technical aspects of LIS/GIS; little has been said of the economic consequences of LIS/GIS information. This paper looks at the concept of economic growth and the problems of understanding and quantifying the benefits of LIS/GIS information. The benefits and costs of this information are discussed along with the concept of how LIS/GIS information products contribute to economic growth. Two different cases of benefit/cost study of spatially related information products are discussed.

Jeffress, G.A. and D.C. Conway. GIS innovation diffusion. *Technical Papers, GIS/LIS '89* (Orlando, FL).

In 1989 it is estimated there will be a \$600 million investment in GIS hardware and software in North America. This figure is expected to climb to \$1 billion in 1990 and to continue at this level for the foreseeable future. This paper looks at the introductory process and diffusion of GIS technology. How do organizations choose to invest in GIS? What factors prevent the introduction of GIS innovation within organizations? The investment in and commitment to introduce GIS technology can be expensive. Aspects of the decision process which lead to investment in GIS are presented and discussed.

Moreno, R.J. and H.J. Onsrud. Legally supportable cadastral information system. *Proceedings, ACSM/ASPRS Conference* (Denver, CO).

Research has begun towards the development of an automated land information system that will support spatial integrity to the extent that surveying measurements allow and will survive most legal challenges to the quantitative and qualitative validity of the cadastral information contained in the system. This paper describes three levels of applications in which a legally supportable cadastral information system (LSCIS) could be particularly useful. Properties necessary for a useful LSCIS, a brief overview of research developments, and a planned research project are outlined.

Obermeyer, N. A systematic approach to the taxonomy of geographic information use. *Technical Papers, GIS/LIS '89* (Orlando, FL).

This paper represents an effort to lay groundwork for the development of a taxonomy of the use of geographic information. Beginning with a discussion of some of the pitfalls of taxonomy development, the author proposes a framework for the taxonomy of geographic information use. An iterative process is suggested that would seek the input of geographic information experts and survey users of geographic information to achieve consensus on an appropriate taxonomy.

Onsrud, H. Understanding the uses and assessing the value of geographic information. *Technical Papers, GIS/LIS '89* (Orlando, FL).

This article describes a process for identifying and addressing critical use and value research issues. The process was derived from discussions which took place at a specialist meeting hosted by the NCGIA. Also described are two initial research thrusts currently being carried out utilizing the process.

Onsrud, H. Legal and liability issues in publicly accessible Land Information Systems. *Technical Papers, GIS/LIS '89* (Orlando, FL).

This paper focuses on current practices and rule of law which individuals contributing data to a publicly accessible land information system or using data or products from such a system should be aware of.

Onsrud, H. The cadastral mapping challenge for surveyors in the US. *FIG, XIX International Congress Proceedings* (Helsinki).

This paper describes social policies supported by the land tenure system of the US and summarizes the general status of local government cadastral maps. An approach for incorporating surveying measurements into the cadastral maps of the US is presented.

Onsrud, H. Liability concerns for surveyors contributing data to public Land Information Systems. *ACSM Bulletin* (to appear).

LIS now being used do not visually communicate to the user the likely accuracy of the data, the original source of the data, nor much else about the entities displayed. The user is able to "zoom in" precisely on any entity on the screen and receives the false impression that this precision also represents high precision on the ground. If a database is used by anyone other than those operating it, there will always be liability exposure. No LIS will ever be complete for all conceivable purposes of the system nor will the accuracy of data meet the required needs of all potential users. To ensure the greatest use of publicly maintained LIS at the local level, the database manager, those delivering data to the system, and users of the system should all be protected from the consequences of misuse and errors in the system.

Onsrud, H. and B. Dansby. GIS and the legal community. *ACSM Bulletin* (1989) pp. 30-6.

As members of the real estate conveyancing community, lawyers have an important role to play in developing standards and practices needed to improve land records. Society is placing unprecedented demands on government and the private sector to provide land information and analysis of that information. The legal community is being called upon to supply the critical component of spatially accurate, legally supportive ownership information. A GIS that includes deeds or title records will be beneficial to lawyers and the conveyancing community. Such a system has the potential of remote access to title documents, remote filing of title documents and computer-assisted title examinations.

Onsrud, H. and R. Hintz. Upgrading boundary information in a GIS using an automated survey measurement management system. *Technical Papers, ACSM Annual Convention* (Baltimore, MD) Vol. 4 (1989) pp. 275-84.

Onsrud, H. and A. Frank. Surveying, mapping, and LIS education in the United States. *Geo-Information-Systeme* Vol. 2, (1989) pp. 20-2.

This paper briefly describes some current concerns of the surveying and mapping academic community in the United States and how that community is responding to those concerns. Initiatives such as the ILI Centers of Excellence in Land Information Studies program and the NCGIA are described. The likely impacts of such initiatives on research and education in the US are also described.

Onsrud, H., H.W. Calkins and N.J. Obermeyer. Use and value of geographic information: Initiative Four Specialist Meeting summary report. Technical Report 89-6, National Center for Geographic Information and Analysis, Santa Barbara, CA.

Onsrud, H., H.W. Calkins and N.J. Obermeyer. Use and value of geographic information: Initiative Four Specialist Meeting report and proceedings. Technical Report 89-7, National Center for Geographic Information and Analysis, Santa Barbara, CA.

5. Initiative 5: Architecture of Very Large GIS Databases

5.1 Narrative summary of progress to date

Initiative 5 is co-led by Terence Smith at Santa Barbara and Andrew Frank at Maine. Its subject matter consists of the set of problems posed by very large spatial databases (VLSDB), particularly those anticipated to become available in the 1990s. The EOS program of NASA, for example, will generate data at rates far beyond our current capabilities for processing and analysis. Clearly the effective processing, storage, manipulation and analysis of such datasets will require radically new approaches to data models, structures, algorithms and user interfaces.

The Specialist Meeting was held in July 1989 in Santa Barbara, using a somewhat different format. The first two days consisted of a Symposium on Design and Implementation of Large Spatial Databases, attended by over 170 participants. The papers presented at the meeting were fully refereed in advance, and have since been published in the Springer-Verlag Lecture Notes in Computer Science series. We plan to make this the first in a series of such conferences, and the second is now planned for the summer of 1991 in West Germany, to be organized by Oliver Gunther with Program Committee participation from NCGIA (Goodchild, Smith and Frank).

The final three days of the week consisted of a Workshop with 30 participants, organized as a series of discussions on the Initiative's research agenda. The results of the meeting have appeared as a Center Technical Report. Work is now under way at both Santa Barbara and Maine, and will likely continue for at least another 12 months.

The key problem for research relates to accessing data in VLSDB. While attempts will be made in the research effort to maintain some notion of generality with respect to the definition of a VLSDB, a model that underlies much of the research effort relates to the current view of EosDIS, namely a huge database of petabyte size consisting largely of image data of many different kinds, but also including other digitized and spatially-addressed data. In relation to access, difficulties arise because the spatial databases may be so large that users have no clear model of the structure and contents of the database. Accordingly, mechanisms must be provided for concisely and accurately specifying data sets of interest. Other difficulties arise because of requirements concerning the speed of data retrieval for users who require "synchronous" interaction with the system. In order to increase the speed of the retrieval process, the storage structures on which the data resides should permit concurrent access and processing.

The basic idea behind our research plan is that it should be possible to formulate a query to find a subset of the data in terms of a set of conditions or constraints that the data must satisfy; that it should be possible to parallelize the procedure that finds the data satisfying these conditions; and that it should be possible to employ physical storage mechanisms that permit parallel access in order to obtain adequately fast response to a query.

The following sections describe research currently under way in specific problem domains within this general framework.

5.1.1 Spatial algebra

During the Specialist Meeting in Santa Barbara a need for a better understanding of the geometric operations used in spatial databases became evident. In cooperation with the University of Bremen (FRG), the Maine site is working on an algebraic formalization of a map algebra. Dana Tomlin (Ohio State University) originally described a set of operations which can be applied to a "map layer" to produce a new one or to combine two map layers into another one. Other researchers have formalized somewhat similar concepts. We are attempting to formalize these ideas by a single method, namely using algebraic specification methods. We will describe the operations with the types of their arguments and give axioms that describe their behavior. Rigorous and formal definitions will allow us to compare the different proposals.

Algebras are mathematical instruments for describing the operations and properties of objects called sorts. A many-sorted algebra consists of a) a set of sorts, b) operations on these sorts and c) axioms that describe the properties of such operations. Many-sorted algebras are useful instruments for the description of spatial data and spatial properties. So far, those algebras have been developed for specific representations only, i.e., for raster and for vector data. A survey of spatial algebras is presently in progress at Maine; through this survey we expect to find methods for creating an algebra that is completely independent of the representation of spatial data. This work is closely related to Initiative 2.

In order to make major progress towards a systematic approach to handling very large databases, it is important to have a deep conceptual model of a VLSDB. Ideally, a spatial equivalent of the relational model would be of immense value. At Santa Barbara, Smith and his graduate students are developing such a model based on the cartographic idea of a "multimap" containing a set of different thematic layers. The basic task is to define a model that is simple and complete, in the sense of being able to represent arbitrary spatial objects, their interrelationships, and the processes that operate on them.

Researchers at both Maine and Santa Barbara are also working on a comparison among databases using an object-oriented data model, in anticipation that such database management systems would be more suitable for very large spatial databases. At Santa Barbara, Sun Guoqing is comparing various object-oriented data models, including EXODUS (University of Wisconsin) in view of the functional and definitional requirements of KBGIS (Smith's Knowledge-Based GIS). Scott Birko is taking a similar approach to extended relational models.

5.1.2 Managing changes in VLSDB

This area explores methods for handling modifications to a VLSDB. The research centers on ways to reduce the effort of modifying large spatial objects; these improvements are achievable by redesigning the architectural components of the DBMS. They are relevant in the context of "multiple representations" as different representations may be understood as different views of the same base data.

Research in this area at the Maine site involves four subprojects:

Propagation of changes

Geographical objects may be represented at different levels of abstraction. When a low-level object is modified, all corresponding high-level objects making reference to it should reflect those changes. The propagation of modifications may be done either in a lazy

manner, i.e., when the high-level object is accessed, or in an eager way, i.e., the moment the changes occur. We are exploring an approach based on a) materialization, i.e. the physical storage of objects at different levels of abstraction and on b) an eager propagation of changes. This approach minimizes the amount of redundant information needed for the update of objects at high levels of abstraction, and prescribes the type of information to be transmitted across the different levels of abstraction.

Transaction management

Standard concurrency-control algorithms consider two types of object operations only, read and write.

This subproject studies how knowledge of the semantics of operators can increase the concurrency in databases. We assume that this will result in solutions for the control of geometric change operations. A survey of literature has been finished, and work has been started on a) how to deduce commutativity of operations from the algebraic specification of an object and b) the use of state-dependent commutativity of operators for concurrency control.

Active databases

Active databases store rules for executing tasks whenever a condition is triggered. VLSDB can profit from active capabilities: active objects can transmit object changes across different levels of abstraction, guarantee data consistency, take corrective actions for improper operations, etc. A survey of current literature has been completed. The results found are applicable to the propagation of changes in a VLSDB.

Object oriented databases

A comparative study on different object-oriented databases is being undertaken in the context of change management. The results of these studies will help in the selection of an appropriate system for building prototypes.

5.1.3 A parametrized model of a physical VLSDB

One of the major impediments to research with VLSDB is that they are currently extremely expensive to implement, and must use a specific physical architecture. A valuable research function would be served by a model that simulates the performance of a large number of different physical configurations. Computer Science graduate students at UCSB are constructing a simulation model that uses a set of parametrized functions (e.g. disk access latencies).

5.1.4 A logic based query and data manipulation language

As queries to VLSDB are expected to become relatively complex, an important problem is the design of a family of declarative languages that will permit deductive query. These should combine the functions of query, host and data manipulation languages, possess the full expressive power of first order logic, and be based on the multimap model. Smith and his graduate students at Santa Barbara are currently constructing a family of languages using the spatial object language (SOL) of KBGIS, and the LDL developed at MCC. SOL and LDL share some simple features, such as set-valued terms, and LDL uses recursive Horn clause-form rules in a fundamental way. The goal is to make the language easy to use and suitable for representing problems that can be solved by constraint satisfaction techniques. It will contain a set of rules that will permit the recursive definition of new constraints from old, in order to increase the expressiveness of the system. The most challenging part of the research will be the representation of the content of the image data sought in terms of the objects contained in a variety of domains.

5.1.5 Constraint satisfaction procedures for dataset search

A natural approach to finding a dataset satisfying given constraints in a VLSDB may be formulated in terms of a constraint satisfaction problem. This approach has been examined by Sud Menon and Smith, who developed and tested procedures for solving such problems. It is possible, in the case of spatial databases, to take advantage of spatial constraints in the data (e.g. spatial proximity) in terms of spatial constraint propagation. In terms of VLSDB, the key problem is speed, since the procedures used are typically sequential. The most hopeful approach to obtaining speedup is by exploiting the parallelism inherent in the problems. Smith and his graduate students at Santa Barbara propose to develop and test parallel procedures that would speed up the process of finding complex subsets of data.

5.1.6 Metadata

Metadata is systematic, descriptive information about data content and organization, and its effective use will be critical in VLSDB. It is useful to define a metamodel, which includes the basic concepts in the data

model and the rules for defining, enforcing and modifying schemas and consistency rules for particular applications. Metadata can be viewed as an extension to the database schema, in which case the database is referred to as self-describing, or as a separate layer above the data layer.

Metadata can be broadly classified into dictionary metadata and directory metadata. Dictionary metadata deals with questions on the nature of the data and its meaning, whereas directory metadata deals with questions on the location of data and its access. In our opinion it is neither practical nor desirable to separate the two types in terms of implementation.

At Santa Barbara, Smith, David Lanter and their graduate students are looking at what metadata means for VLSDB. They will develop an architecture to store, retrieve and edit metadata, and to use it in support of browse, search and query operations.

5.1.7 Lineage data

The quality of synthesized spatially referenced data goes beyond the quality of data input, as validity is also based upon the logical consistency of manipulations used to derive new information. Documentation of data sources and the transformations applied to them is lineage information, and is an important special case of metadata. It serves as a communication mechanism between data suppliers and users. This research will make a significant contribution to the management of VLSDB by automating lineage information in a suite of software that will enable researchers to track errors in the spatial data handling process.

David Lanter and his graduate students at Santa Barbara are developing the Lineage Information Program (LIP), which will automatically track data sources through spectral and spatial transformations. Research will determine the lineage requirements of image processing and GIS technology, implement a testbed to support lineage-based algorithm development, and transfer the prototype to production environments.

5.1.8 Browse capability for image and digitized cartographic databases

Over the past three years, the Research Libraries Group (RLG), located in Stanford CA, has developed an external design for GRIN, a system for browsing geographically referenced materials in a catalog environment. The design, funded by a private foundation, provides a means for users to determine the existence and characteristics of data within VLSDB. The Santa Barbara site has signed an agreement with RLG for joint research and development effort, and has submitted a proposal to NASA for the first stage of prototype development. If successful, this application will lead to the development of the access workstation prototype, and to testing over RLG's existing network of close to 100 library sites.

5.1.9 Data compression for spatial data

Compression techniques are clearly important in VLSDB, particularly if order of magnitude compressions can be achieved without significant loss of information. At the Santa Barbara site, NCGIA is funding work by Allen Gersho and Smita Gupta in Electrical and Computer Engineering to investigate the appropriateness of existing vector compression methods for the special case of geographic images. Images of different spectral and spatial resolutions are being used in an effort to determine optimal coding schemes, and the potential for generic compression.

5.1.10 Tiling schemes

At Santa Barbara, Michael Goodchild has continued research on optimal tiling schemes for VLSDB, on the assumption that physical partitioning of databases will be necessary because of volume capacity constraints, or in the interests of access efficiency. Funding from DEC and USGS supported research on optimal tiling for optical WORM drives in 1988/9, and substantial theoretical work has been done on the topic by Goodchild, Richard Church and Yang Shiren. For data on the globe, the tesseral scheme described under Initiative 1 is clearly useful, and is being investigated as a potential tiling and indexing scheme for the Digital Chart of the World (a project of ESRI, funded by DMA).

Papers presented at AutoCarto 9 and the ICA in 1989 reported on some of the practical and theoretical results, and a major paper on the theory is in draft form. It analyzes a class of tiling schemes defined by digit interleaved keys, and shows that all such keys perform identically against an objective function based on linear dependence between access time and difference of key. A second paper has been prepared on an analysis of databases in which selective tiles are duplicated, which significantly outperform more conventional types.

5.2 Annotated list of publications

5.2.1 Articles in refereed journals

Egenhofer, M.J. and A.U. Frank. LOBSTER: combining AI and database techniques for GIS. *Photogrammetric Engineering and Remote Sensing*, Special Issue on Knowledge-Based Expert Systems.

(see section 2.2.1 above)

Ehlers, M. Multisensor image fusion techniques in remote sensing. *Photogrammetria* (under consideration).

Current and future remote sensing programs such as Landsat, SPOT, MOS, ERS, JERS, and the space platform's Earth Observing System (Eos) are based on a variety of imaging sensors that will provide timely and repetitive multisensor earth observation data on a global scale. Visible, infrared and microwave images of high spatial and spectral resolution will eventually be available for all parts of the earth. It is essential that efficient processing techniques be developed to cope with the large multisensor data volumes. This paper discusses data fusion techniques that have proved successful for synergistic merging of SPOT HRV, Landsat TM and SIR-B images. Examples are given for integrative rectification, enhancement of cartographic feature extraction and improvement of spatial resolution.

Ehlers, M., G. Edwards and Y. Bedard. Integration of remote sensing with GIS: a necessary evolution. *Photogrammetric Engineering and Remote Sensing* 55(11):1619-27 (1989).

GIS have developed from the need to combine attribute information about land with its cartographic representation in order to perform spatial analyses. Remote sensing and image analysis technology have been used in parallel to obtain information about the Earth at resolutions as small as a few tens of centimeters to planetary scales. The integration of such capable data acquisition and analysis technologies is becoming increasingly important for resource management. Some of the many preliminary efforts to this end which have been made over the past few years, both in the commercial world and in research circles, are discussed. Several key impediments to this integration are identified. This leads to the formulation of a three-stage process of integration which combines the strengths of both technologies. An earlier version of this paper appeared in proceedings form (Ehlers, M., K. Beard, and M. Jadcowski. The impact of high resolution satellite data on GIS. *Proceedings, IGIS Symposium* (Baltimore, MD) pp. 231-40 (1989).

Ehlers, M., M. Jadcowski, R. Howard and D. Brostuen. Application of SPOT data for regional growth analysis and local planning. *Photogrammetric Engineering and Remote Sensing* 56(2):175-80 (1990).

SPOT satellite data, combined with state-of-the-art image processing and GIS technology, are valuable tools for timely and accurate analyses of regional urban and suburban development. A combination of unsupervised classification and image interpretation techniques can be used for land-use/land-cover analysis and to determine regional growth patterns. With this procedure, accuracies for growth detection as high as 93 percent may be achieved and verified by a rigorous error analysis. Once incorporated in a GIS database, areas can be measured and the spatial distribution of growth patterns can be analyzed. Existing digitized map data or GIS layers such as zoning maps may then be overlaid and compared with the actual land-use/land-cover information.

Ehlers, M. Remote sensing and GIS: towards integrated spatial information processing. *IEEE Transactions on Geoscience and Remote Sensing* (to appear).

In recent years the demand for timely and accurate information has increased for a number of earth related disciplines, including urban and regional planning, natural resources management, agricultural studies, topographic mapping, and geological exploration. A much discussed approach to meeting these demands is the development of GIS, which allow efficient data storage and retrieval, incorporate processing and display functions, and permit the integration of digital remote sensing data. There are, however, major impediments for efficient integration of remotely sensed data with GIS, which are primarily based on their different concepts of space. An earlier version of this paper appeared in proceedings form (Ehlers, M. The potential of multisensor satellite remote sensing for GIS. *Technical Papers, ASPRS/ACSM Annual Convention* (Baltimore, MD) Vol. 4 (1989) pp. 40-5.), (Ehlers, M., and P.D. Haggerty. Interfacing remote sensing and GIS technologies for urban and regional planning. *Proceedings, URISA '89* (Boston, MA) Vol. 4. (1989) pp. 85-92.)

Gao, P. and T.R. Smith. Space efficient hierarchical structures: relatively addressed compact quadtrees for GIS. *Journal of Image and Vision Computing* (to appear).

Traditional pointer-based quadtree data structures are generally viewed as inferior to linear quadtrees when used in GIS. This paper presents an improved pointer-based quadtree called the relatively addressed compact quadtree. Storage requirements are comparable with those for linear quadtrees. A memory management scheme is also proposed for managing the pointer-based quadtree in secondary storage.

Menon, S. and T.R. Smith. A declarative spatial query processor for GIS. *Photogrammetric Engineering and Remote Sensing* 55(11):1593-600 (1989).

The design and implementation of a declarative GIS query processor capable of extracting the locations of complex objects from a spatial database is described. The processor extracts objects from the database in a single, automated step. Search is based on an efficient query processor, named forward constraint propagation, that integrates spatial constraint propagation, geometric search using hierarchical data structures, and an effective heuristic used in solving constraint satisfaction problems.

Smith, T.R., P. Gao and P. Gahinet. Asynchronous, iterative and parallel procedures for solving the weighted-region least cost path problem. *Geographical Analysis* 21:147-68 (1989).

Based on research completed before the NCGIA award, this paper defines a family of procedures for finding least cost paths, using local, asynchronous, iterative and parallel processes. Although the procedures are guaranteed to terminate, it has so far been impossible to prove that they always terminate in admissible paths in the case of a two-dimensional triangulation. Extensive simulations have shown convergence to admissible paths in all cases examined.

5.2.2 Articles in refereed conference proceedings

Barrera, R., D. Hudson and A.U. Frank. Materialized support for efficient view updates. *1990 Conference Proceedings, Very Large Databases Conference* (Sydney, Australia) (to appear).

Data base views are supported in order to decrease system response-time. For their efficient maintenance, it is important to provide views with update methods that will, after modifications have been made to base objects, propagate relevant changes to the views in an incremental and localized manner. View states are view expressions that are partially evaluated such that they result in maximally reduced intermediate data structures from which the full view can eventually be deduced and for each of which updates are autonomously computable. This paper is concerned with the formal description construction of minimal view-states. It establishes two conditions on data structures necessary and sufficient for their qualification as a view-state.

Egenhofer, M. and A.U. Frank. PANDA: an extensible DBMS supporting object-oriented software techniques. *Proceedings, Database Systems in Office, Engineering and Science* (Zurich, March 1989). In T. Haerder, editor, *Informatik-Fachberichte* Vol. 204, pp. 74-79.

The PANDA databases management system was designed for nonstandard applications which deal with spatial data. It supports an object-oriented program design with modularization, encapsulation, and reusability, and can be easily embedded into complex applications, such as spatial information systems or cartographic expert systems. Complex objects and their operations are defined. A layered structure on top of the programmer's interface provides object operations which include potentially complex consistency constraints.

Egenhofer, M.J., A.U. Frank and J.P. Jackson. A topological data model for spatial databases. In A. Buchmann, O. Gunther, T.R. Smith and Y.-F. Wang, editors, *Design and Implementation of Large Spatial Databases*. Lecture Notes in Computer Science 409, Springer-Verlag, New York, pp. 271-86 (1990).

There is a growing demand for engineering applications which need a sophisticated treatment of geometric properties. Implementations of Euclidean geometry, commonly used in current commercial GIS and CAD/CAM, are impeded by the finiteness of computers and their numbering systems. A spatial data model is proposed which is based upon the mathematical theory of simplices and simplicial complexes from combinatorial topology. It guarantees the preservation of topology under affine transformations. The implementation as a general spatial framework on top of an object-oriented DBMS is discussed.

Frank, A.U. and R. Barrera. The fieldtree: a data structure for GIS. In A. Buchmann, O. Gunther, T.R. Smith and Y.-F. Wang, editors, *Design and Implementation of Large Spatial Databases*. Lecture Notes in Computer Science 409, Springer-Verlag, New York, pp. 29-44 (1990).

Efficient access methods, such as indices, are indispensable for quick answers to database queries. This article describes the fieldtree, a data structure that provides one such access method. The fieldtree has been designed for applications where range queries are predominant and spatial nesting and overlapping of objects are common. Besides their hierarchical organization of space, fieldtrees are characterized by three other features: they subdivide space regularly, spatial objects are never fragmented, and semantic information can be used to assign the location of a certain object in the tree.

Goodchild, M.F. Invited talk: tiling large geographical databases. In A. Buchmann, O. Gunther, T.R. Smith and Y.-F. Wang, editors, *Design and Implementation of Large Spatial Databases*. Lecture Notes in Computer Science 409, Springer-Verlag, New York, pp. 137-46 (1990).

Geographical variation is infinitely complex, so the information coded in a spatial database can only approximate reality. The information will always be inadequate, in spatial resolution, thematic or geographical coverage. "Large" can be usefully defined as exceeding our current capacity to deliver. Traditional stores partition geographical data by theme and geographically. It is assumed that digital geographical databases will be largely archival, and will be similarly partitioned. A general model of a large archival store is presented. The properties of a generalized Morton key as a means of indexing tiles are analyzed, and its role in traditional systems of tile indexing is illustrated. For global databases, a tiling based on recursive subdivisions of the triangular faces of an octahedron using a rule of four is proposed. Earlier versions of this paper appeared in proceedings form (Goodchild, M.F. Optimal tiling for large cartographic databases. *Proceedings, AutoCarto 9* (Baltimore, MD) pp. 444-51 (1989)), (Goodchild, M.F. Optimal tiling for cartographic databases. *Proceedings, ICA* (Budapest, August 1989) (to appear)).

5.2.4 Articles in other outlets

Barrera, R. and J. Vasquez-Gomez. A shortest path method for hierarchical terrain models. *Proceedings, AutoCarto 9* (Baltimore, MD) pp. 156-163 (1989).

(see section 2.2.4)

Edwards, G., Y. Bedard and M. Ehlers. Advanced integration of remote sensing image analysis with GIS. *Proceedings, GIS National Conference* (Ottawa, Canada, 1990) (to appear).

Remote sensing image analysis systems and GIS show great promise for the integration of a wide variety of spatial information as a support to decision-making. Many preliminary efforts to unite them into one technology for information management have been made over the past few years, both in the commercial world and in academic research circles. In the first part of this paper, some of the successes achieved so far are discussed, along with their limitations. A characteristic of most successful highly integrated systems is that they are confined to very specific problems. In the second part of this paper, these successes are placed in the context of a vision of long-term integration. Several obstacles and concerns related to such integration are discussed.

Egenhofer, M. and A.U. Frank. Object-oriented modeling in GIS: inheritance and propagation. *Proceedings, AutoCarto 9* (Baltimore, MD) pp. 588-598.

The relational data model has proven to be too restrictive for applications with spatial data, such as GIS. The object-oriented approach seems to overcome some of the deficiencies. By incorporating the abstraction mechanisms of generalization and aggregation, the data model gets richer and more powerful than the relational model, and the application designer is given more and better tools to model complex situations. Two methods for the derivation of properties are introduced: (1) inheritance describing properties and methods of subclasses in is-a hierarchies, and (2) propagation deriving properties in part-of hierarchies. While inheritance acts in a top-down fashion along the generalization hierarchy, propagation can derive values from parts to the aggregates (bottom-up).

Egenhofer, M.J. and A.U. Frank. Object-oriented software engineering considerations for future GIS. *Proceedings, IGIS Symposium* (Baltimore, MD) (to appear).

Currently, experts in software engineering promote an object-oriented approach for the implementation of large software systems. This advanced technique, based upon the definition of object types in combination with the corresponding operations in a modular fashion, is characterized by clear coding which can be maintained easily. Powerful object-oriented abstraction mechanisms, such as classification, generalization, and aggregation, are the framework for an object-oriented model. Specific programming languages have been designed to allow programmers to implement software systems pursuing an object-oriented design. Future GIS will benefit from object-oriented abilities to concisely define complex problems.

Ehlers, M., K. Beard, P. Haggerty and B. Pearce. Synergistic integration of remote sensing, GIS/database technology and numerical modeling for marine GIS. *Proceedings, IGARRS '90* (College Park, MD).

GIS technology when coupled with synoptic data from remote sensing and results from numerical modeling will have particular utility in the area of marine research. Current GIS are designed to handle land information data (e.g. transportation and communications networks, land use/land cover classes, political boundaries) for user/uses concerned with the management of terrestrial areas. Work is presently underway at the University of Maine to design a marine GIS (MGIS) which would extend current technologies to handle marine data that may vary in time as well as depth. This paper presents a concept for MGIS research and preliminary results from a remote sensing database project for the Gulf of Maine.

Frank, A.U. Requirements for database management systems for large spatial databases. In R.E. Dahlberg, J.D. McLaughlin, B.J. Niemann Jr., editors, *Developments in Land Information Management*, Institute for Land Information, Washington (1989).

The paper presents a list of functions a database management system should provide for GIS, and some quantitative estimates are given on the size of a typical graphical retrieval. The main part of the paper shows a possible layered architecture for a Geo DBMS, including the physical storage level with clustering and buffer management, a layer to protect data, access methods using key values, spatial access (supported by physical clusters), data modelling concepts and corresponding operations, support for abstract data types and object-oriented programming, methods to deal with consistency constraints, and query languages. This architecture has been used for the construction of the PANDA database system.

Hsu, P. An analysis of spatial structure data in landscape for GIS. Proceedings, Fourth International Symposium on Spatial Data Handling (Zurich) (to appear).

(see section 2.2.4). An earlier version appeared in proceedings form (Hsu, P. Using high resolution digital images to describe the efficiency of urban spatial structure: an analysis of transportation network and its adjacent land use. *Proceedings, ISPRS Commission II/VII International Workshop* (Orono, ME).)

Stine, R.S. and D.P. Lanter. Considerations for archaeology database design. In K. Allen, S.W. Green and E. Zubrow, editors, *Interpreting Space: GIS and Archaeology*. Taylor and Francis, London (to appear).

This contributed chapter reviews GIS data models in the context of archaeological applications, stressing the implications of the need to store time-dependent data.

6. Initiative 6: Spatial Decision Support Systems

6.1 Narrative summary of progress to date

Initiative 6 extends into the area of decision-making, and examines the possible role of GIS and associated technologies in supporting the decision-making process. The Center decided to narrow the possible scope of the initiative somewhat, in two directions. First, we feel that we should adopt the definition of Decision Support Systems proposed by Geoffrion and others, by assuming that although a DSS is by definition oriented to the solution of ill- or poorly-structured problems, it nevertheless contains a number of well-structured models. This is in contrast to arguments put forward by Cowen and others that GIS and SDSS are in effect synonymous. Second, we feel that it is impossible to consider all possible applications of SDSS, and have narrowed the field to the general area of marketing, retailing, location theory and socioeconomic models. We feel that this will allow the research agenda to cover a broad range of possible applications, but to return to a narrower focus as appropriate.

The Initiative is led by Paul Densham (Buffalo) and Michael Goodchild (Santa Barbara). The Specialist Meeting was held in Santa Barbara March 14-18 1990, and brought together over 30 participants, primarily academics but with strong representation from the private sector. Unfortunately it proved difficult to identify suitable government agency participants because of the nature of the topic,

but many agency concerns were already familiar to the other participants.

The Initiative's potential research agenda was discussed in a presentation at GIS/LIS '89 which appears in the proceedings, and was refined substantially over the three working days of the meeting. Besides design issues, the meeting was particularly concerned about the extent to which work on SDSS could shed light on the decision-making process, or provide an experimental environment in which to study it. The results of the meeting will be appearing shortly as a Center Technical Report, and research will begin in earnest in July. A newsletter and mailing list will be established, at least for the duration of the Initiative, and Bruce Ralston (University of Tennessee) has volunteered to be editor.

Various publications relevant to Initiative 6 are listed below, including a paper presented by Helen Couclelis at the North American RSA meetings in November 1989.

6.2 Annotated list of publications

6.2.1 Articles in refereed journals

Armstrong, M.P. and P.J. Densham. Database organization alternatives for spatial decision support systems. *International Journal of GIS* (to appear).

Couclelis, H. Geographically informed planning: requirements for planning-relevant GIS. *Papers of the Regional Science Association* (under consideration).

The paper examines applications of GIS in planning. Current GIS technology implies a container-like view of space, whereas the higher functions of planning require a relative view. Unless current technology develops in this direction, its value to planning will be limited to low-level functions. Batty's view of planning as parallel distributed computation is found to be useful.

6.2.4 Articles in other outlets

Armstrong, M.P., P.J. Densham and D.A. Bennett. Object oriented locational analysis. *Proceedings, GIS/LIS '89* (Orlando, FL) pp. 717-26 (1989).

An object oriented approach to the analysis of spatial data is presented. The paper begins by describing object oriented programming, and then defines a structure of spatial and analytical objects within the problem domain of locating facilities. From these objects a set of object classes and inheritance structures is created. The spatial and analytical objects are represented using frames. A Smalltalk implementation of a heuristic location-allocation algorithm is also described.

Densham, P.J. and M.F. Goodchild. Spatial decision support systems: a research agenda. *Proceedings, GIS/LIS '89* (Orlando, FL) pp. 707-16 (1989).

Definitions of GIS often focus on the capture, storage, manipulation, analysis and display of spatial data, implying that GIS implicitly are designed to support spatial decision-making. For many spatial problems, however, GIS do not support decision-making effectively: analytic modeling capabilities are lacking and system designs are not flexible enough to accommodate variations in either the context or the process of spatial decision-making. The paper draws a distinction between GIS and spatial decision support systems in terms of system design, the types of problem to which each can be applied, and the decision-making process supported. Impediments to both design and implementation are identified, and a research agenda is outlined to address these problems.

7. Initiative 7: Visualization of the Quality of Spatial Information

7.1 Narrative summary of progress to date

The initiative will be led by Barbara Buttenfield (Buffalo) and Kate Beard (Maine) and is due to begin formally in late 1990, although planning has already begun. Most of the effort in recent months has focused on defining a scope for the initiative. Initial plans proved too narrowly focused on depiction of error, on graphical design, and on empirical evaluation of map symbols. Discussions with researchers in and out of NCGIA and reading over the winter have resulted in broadening the scope of the initiative, to incorporate issues of database management and the use of visualization to improve spatial modeling. Research publications to date represent efforts to consider visualization in GIS in this broader context. Planning, discussion and refinement will continue during the spring and summer.

Technology currently allows us to display large volumes of information quickly, and to process it. Information on the quality of data is also important for effective use of GIS data. It affects the credibility of data representation, and the reliability of decision-making and interpretations. Visual tools may help in model generation, and in steering simulations, and their incorporation into the research process will improve the quality of spatial analysis. Clearly, the quality of information will vary spatially, and visual tools for displaying error, statistical bias, and reliability can only improve and facilitate use of GIS. At present, those tools are either unavailable (in existing GIS packages) or not well-developed (error models and the process of visualization are only recently being addressed directly).

Ties to other research initiatives include development of measures for spatial accuracy and error (I1). Some of the research in that initiative focused on error propagation during buffering, overlay, and other GIS operations. In I7, mechanisms for management and depiction of error measures and indices of data quality will be a topic of interest. Management of data quality during database updates will require formalizing links between data variables and associated measures of data quality.

The impact of visualization on modeling is another relevant issue. Visualization can provide modeling capabilities that were previously unavailable. For example, a Master's project currently in progress at Buffalo is exploring the use of map animation for interpreting climatological patterns (temperature changes over the last century). Other questions are also relevant. What visual tools may be incorporated to generate models, or for sensitivity analysis? What impact will these modeling tools have on background computations (for example, in parallel processing)? And how will the ability to manipulate data quality as well as data modify or improve spatial decision-making (I6)?

The Specialist Meeting is scheduled to occur in March, 1991, immediately following the Baltimore AutoCarto 10 / ACSM / ASPRS conference. Several paper sessions are planned for presentation at the conference, to maximize visibility for the initiative, and to encourage involvement in the research by people outside the Center. Paper sessions are also being organized for the GIS/LIS conference in Anaheim in November 1990 to the same end. We are currently soliciting participation in these paper sessions from the general community. The Specialist Meeting will follow immediately after the Baltimore conference(s), someplace near Washington or Baltimore. Participation in these sessions will be by invitation. The format will involve discussion in small working groups. Position statements of 5-6 pages will be required of all participants about 6 weeks before the meeting in preparation. Participants will include researchers from statistics, computer science (with particular reference to the recent ViSC initiative), cartography, and geographical analysis. Federal agencies who have already expressed interest in active participation include EPA and the Soil Conservation Service (SCS). We hope to enlist industrial participation from fields to include medical imaging and diagnosis and telecommunications, as well as from the GIS vendors.

Initial efforts by Buttenfield and John Ganter have resulted in a preliminary version of a working research bibliography. At present, there are over 50 entries, and about 80% are annotated. Before online searches begin, the citations will be transferred to a relational database format, to provide capabilities for title, author and/or keyword search as the bibliography continues to develop. Bibliographic search will be pursued at other research sites, and the authorship list will likely grow along with the bibliography.

7.2 Annotated list of publications

7.2.1 Articles in refereed journals

Leonard, J.J. and B.P. Buttenfield. An equal value gray scale for laser printer output. *The American Cartographer* 16(2):97-109 (1989). To be reprinted in A.A. Wood, editor, *Design in Cartography*, American Congress on Surveying and Mapping.

Computer-generated line printer maps have been widely used to produce quick, inexpensive choropleth maps. In an attempt to make classes on choropleth maps more easily distinguishable, various equal value gray scales have been applied, with varying degrees of success. Their shortcomings have resulted from either technological restrictions or the application of an inappropriate grayscale. This research develops an equal value gray scale for use with 300 dpi laser printer technology, and evaluates the gray scale by psychophysical testing.

7.2.3 Articles, chapters and monographs in other refereed outlets

Buttenfield, B.P. and W.A. Mackaness. Visualisation. In D.J. Maguire, M.F. Goodchild and D.W. Rhind, editors, *GIS: Principles and Applications*. Longman, New York (to appear).

First, an argument is made for the rise of visualization as a discipline. The chapter then proceeds to trace its development through several related fields, comparing their definitions of its scope and research relevance. A conceptual framework is developed linking these disciplinary perspectives, and tying their research priorities to specific GIS issues. The impact of visualization on GIS is covered by citing applications of techniques to various phases of the data collection, analysis, and display process. In a summary section, research topics of particular relevance to progress in GIS are proposed.

MacEachren, A.E. with B.P. Buttenfield, J.C. Campbell and M.S. Monmonier. Visualization. In R.F. Abler, J.M. Olson and M.G. Marcus, editors, *Geography's Inner World* (to appear).

Geographers have always relied on concrete representation tools to facilitate visualization. Maps have been a dominant representation form in geography due to the geographer's concern with relative spatial location, but diagrams, graphs, images, etc. have also played a role. Recent interest within geography and other disciplines in "scientific visualization" has emphasized development of new and more powerful tools for visualizing. Geographers have much to contribute to development of scientific visualization because of their long tradition of visualizing the world through maps. Of concern in the present chapter are both the concrete (map) representations and their ability to help us define questions, develop hypotheses, and find answers.

7.2.4 Articles in other outlets

Beard, M.K. Use error: the neglected error component. *Proceedings, AutoCarto 9* (Baltimore, MD) pp. 808-817 (1989).

(see section 4.2.4)

Buttenfield, B.P. Texture and value relations of computer-generated gray tones. *Proceedings, International Cartographic Association Conference* (Budapest, August 1989) (to appear).

A number of equal value gray scales have been developed to relate perceived value with physical measurement, including the recently published Leonard curve (42.5 dpi), the Munsell scale (>150 dpi), generated from very fine resolution painted chips, and the Williams scale (31 dpi). One of the more challenging aspects of understanding the perceptual gray spectrum relates to its sensitivity to modifications in texture, or visual coarseness. Determination of the relations between perception of value and texture will require comparison of the various gray scales in a common printing medium. Unfortunately, Williams died without ever publishing his reflectance values, and thus quantitative comparisons to his research results have not been

possible. This paper reports digital reconstruction of the Williams data set, and statistical comparison with the Leonard and Munsell curves.

Buttenfield, B.P. and J.H. Ganter. Visualization and GIS: what should we see, what might we miss? *Proceedings, 4th International Symposium on Spatial Data Handling* (Zurich, July 1990) (to appear).

The ease with which visualization tools may be integrated within GIS packages varies considerably depending on at least two issues, including the domain of the phenomena to be studied, and the purpose or intent of the user. This presents a substantial challenge to the system designer. The authors argue for a coordinated effort to formalize exploration of visualization with GIS. The paper suggests that GIS requirements for visualization include conceptual, technological, and evaluatory solutions, which may be seen to vary over three broad domains: inference, illustration, and decision-making. Each presents a challenge to the integration of appropriate visualization tools in GIS.

Ganter, J.H. Interactive graphics: linking the human to the model. *Proceedings, GIS/LIS '88* (San Antonio, TX), pp. 230-239.

Discovery and innovation, which have traditionally involved thinking visually and producing images, increasingly benefit from GIS and CAD. As new visualization technologies are implemented, it is particularly important to understand the human faculties which use pictures as tools in thinking. Science and engineering define problems, explain processes, and design solutions through observation, imagination and logic. This conceptual thought relies on a cognitive database of sensed verbal and non-verbal information retained, managed, and updated within the short and long-term human memories. Research suggests that the individual must actively manipulate a phenomenon under study, and its representations, to enhance and maintain this database, and to produce abstractions and generalizations.

Weibel, W.R. and B.P. Buttenfield. Map design for GIS. *Proceedings, GIS/LIS '88* (San Antonio, TX) vol. 1, pp. 350-59 (1988).

Maps are a major tool for decision-making with GIS. Current GIS software includes functions to create cartographic output automatically or interactively. However, none of the current turnkey systems include mechanisms to ensure the correct use of graphics functions. This may lead to poor use of graphics by untrained users. Poorly designed maps may convey false ideas about the facts represented by the data, and bias the decision-making process. This paper explores ways to improve the quality of GIS map products, and increase effectiveness of information transfer based on graphics. Alternatives presented include discussion of software enhancements, user interface design, and expert system technology.

8. Initiative 8: Expert Systems in Cartographic Design

8.1 Narrative summary of progress to date

At the November 1989 meetings of the NCGIA Scientific Policy Committee and Board of Directors, this initiative was delayed to an indefinite start date sometime after mid-1991. However, preliminary work on the Initiative continues, and one paper has been published, and another submitted. In addition, three papers related to Initiatives 3 and 8 are listed under Initiative 3. During the first half of 1990, William Macaness, NCGIA post-Doctoral research fellow at Buffalo, will coordinate a review of the goals, and terms of reference, for Initiative 8, and in consultation with Mark, Frank, and Buttenfield, will prepare recommendations on the nature and timing of Initiative 8 to the NCGIA in June 1990. Mark and Frank are the proposed co-Leaders of the Initiative.

8.2 Annotated list of publications

8.2.1 Articles in refereed journals

Wu, C.V. and B.P. Buttenfield. Reconsidering rules for point feature name placement. *Cartographica* (under consideration).

Existing guidelines for automated name placement are most often based on Imhof's and Yoeli's research. These guidelines of map name placement have not been verified empirically. This study addresses the validity of Yoeli's model concerning eight possible positions for labelling a name to a point feature and their priorities. Investigation is focused on how lettering is placed on road maps, and diversities in styles of name placement based on differing publishing houses. It is concluded that Yoeli's positional prioritization is not evident on these maps. The prioritization may be a matter of publisher's preference. Certain graphic constraints may confine these preferences into a unique solution crossing different maps. The objective of this research is to provide reliable and verifie guidelines for researchers developing algorithms for automated name placement.

8.2.3 Chapters in books

Buttenfield, B.P. and D.M. Mark. Expert systems in cartographic design. In D.R.F. Taylor, editor, *Microcomputers in Cartography*, International Cartographic Association (in press).

Expert systems are computer programs that represent expertise in the form of rules. They are best applied to fields which people learn by exemplification and apprenticeship; cartography is an example of such a field. This paper provides explanations of fundamental concepts from expert systems. Then, a model of the cartographic process is presented, followed by the overall design for a full cartographic expert system. Required functionalities are identified, some of the relations among the modules of such a system are described, and progress to date on the implementation of each phase of the process is assessed.

9. Initiative 9: Institutions Sharing Geographic Information

This initiative will begin following the end of Initiative 4.

10. Initiative 10: Temporal Relations in GIS

10.1 Narrative summary of progress to date

Although a start date for this initiative has not yet been set, in order to prepare properly a graduate student has been hired at Maine. We plan to establish a collection of background material to be used by the participants of the Specialist Meeting and to survey the literature and discussing its applicability for GIS. A literature review on temporal aspects has been concluded. This review covered three categories of references:

10.1.1 Temporal knowledge representation

This category covered a) models of time and b) logical formalisms used in temporal reasoning.

10.1.2 Temporal DBMSs

Included in this category were a) temporal versions of data models, b) temporal query languages, and c) physical aspects in the design of temporal DBMSs.

10.1.3 Database design using temporal concepts

This category covered specification methods for preserving static and dynamic consistency conditions.

11. Initiative 11: Space Time Models in GIS

No start date has been set for this initiative.

12. Initiative 12: GIS and Remote Sensing

12.1 Narrative summary of progress to date

This Initiative will be led by John Estes and Frank Davis at Santa Barbara, and will focus on all aspects of the relationship between the two technologies. The Specialist Meeting is currently planned for early 1991, effectively giving I12 higher priority than some of the lower numbered initiatives because of strong interest in this area, particularly in the context of global science. Work relevant to the initiative has already been cited in the areas of accuracy (I1) and large spatial databases (I5), which are two of the major issues which this initiative must confront.

13. Initiative 13: User Interfaces

13.1 Narrative summary of progress to date

An Initiative on user interfaces was approved by the Board of Directors at its December 1989 meeting in view of the strong interest in this area and its natural linkage to I2. A Specialist Meeting is planned in conjunction with the GIS/LIS '90 meetings in Anaheim CA in November.

Work is already under way on user interface issues in connection with I5 at Santa Barbara and Maine, and in connection with I2 at Buffalo and Maine, and the results of some of this work have already been cited in this report. At Maine, Werner Kuhn and Max Egenhofer have been selected to organize a 2-day workshop on "Visual Interfaces to Geometry" in conjunction with CHI'90, the annual conference of ACM's Special Interest Group on Computer-Human Interfaces. The objective of this interdisciplinary workshop is to explore and integrate advanced approaches to the visual representation and interactive manipulation of geometric information. Considerable expertise on various aspects of this problem has been accumulated in a number of fields such as Computer Science, Engineering, Cognitive Science and Linguistics, Geography and Cartography; however, little effort has been made to combine this expertise. Participants to this interdisciplinary workshop have been selected based on position papers submitted. In addition, three invited speakers (A. Herskovits, S.K. Chang, A. Frank) will present their views of geometry from linguistic, computational, and engineering points of view. Results will be published as a report in the *SIGCHI Newsletter*. Kuhn and Egenhofer have been also invited to serve as guest editors of a special issue on "Visual Interfaces to Geometry" of the journal *Visual Languages and Computing*. The workshop will also provide input for the upcoming specialist meeting.

Education

GIS is a comparatively new field, with origins which have been motivated more strongly by practical application than by intellectual curiosity. As the industry has grown and the potential of GIS has been realized more and more widely, an acute shortage of trained personnel has emerged. There are currently few academic programs in GIS, or programs providing substantial exposure to GIS. Frequently the only training available to users is that provided by the vendor.

Each site of NCGIA provides course sequences in GIS, and enrolls significant numbers of students from a variety of disciplines each year. At all three sites offerings have been expanded and better facilities are available. But however valuable these locally based efforts, the numbers of trained students produced will be insignificant in relation to the national demand. Projections of the GIS industry over the next decade lead to estimates of total employment in GIS applications on the order of 10^4 or even 10^5 in the US. For this reason the major effort of the Center in the past 18 months has gone into approaches which are likely to yield much greater numerical impact.

Resources have been devoted in two areas. First, we have developed a number of short courses. These are intended for use at conferences, or as simple tutorial manuals, or to be given in agencies or by vendors. They consist of workbooks and associated slides, and are sufficiently complete to be given by any one of a group of people, not necessarily associated with the Center. One of these, *Spatial Analysis using GIS*, has been offered on at least seven occasions since it was designed, in three cases by people with no direct Center affiliation. The workbooks for these courses have been published as Center miscellaneous publications, and widely distributed.

Over time we plan to extend the set of short courses available, specializing in topics of some intellectual depth, and including courses on the results of Center-related research. We see this as a way of increasing exposure of Center activities in depth, in a form of outreach, as well as of enhancing the set of topics available in conference workshops and short courses.

The "Core Curriculum Project" is currently the major educational initiative of the NCGIA. Since GIS is a new area of study for most universities and colleges, there exists a demand for comprehensive teaching materials covering the basic theory and concepts, as well as the technical and application issues. We see this as distinct from training in specific vendor products, which is a function much better performed by the vendors themselves. Because GIS is rooted in technology, it is easy to build courses around the tools and lose sight of the more conceptual basis of the field, and its theory. Yet an understanding of these areas is critical for people who will fill the managerial and analytical roles in GIS applications. Perhaps the greatest shortage at this time is of people who combine an understanding of the technology on the one hand, with its strengths and weaknesses, and the needs of real spatial problem-solving on the other.

The Core Curriculum is not intended to be a model program or set of courses. Instead it tries to respond to a perceived need for comprehensive materials to form the basis or core of a program, but which would be adapted and augmented to meet individual needs. The curriculum is modular at several levels, allowing the user to pick and choose topics, add local examples, stress particular areas or respond to gaps in student backgrounds.

During the Fall of 1988, the Center at Santa Barbara, with the help of academic and private sector experts from around the world, developed a comprehensive outline for a one-year sequence of courses in GIS. In the quarter system used at Santa Barbara there are three courses, but it is easy to restructure the contents into two to match a semester system. The outline was circulated widely, presented at several meetings, and iterated through a number of meetings. In draft form, it identified 75 one-hour lectures, grouped into modules of three or four, and into three courses of 25 lectures each. A brief outline was written for each lecture.

In late 1988 the Center invited selected experts from the GIS community to "write" the lectures. A model was distributed of about eight pages of notes, plus references for further reading, and exam and discussion questions. Response varied, ranging from very extensive essays on favorite topics to very rough outlines. Goodchild and Karen Kemp edited the responses into a consistent format in the summer of 1989. In draft form the materials consist of 75 sets of lecture notes with overhead and hand-out masters, slides, lists of suggested readings and videos, and twelve laboratory exercises complete with data sets on disk. It is contained in three 3-ring binders, one for each course, with about 350 pages of text each.

In early 1989 the Center began soliciting agreements to test the materials in classroom settings. An announcement appeared in several GIS newsletters and invitations were made verbally at meetings and in informal discussions. By the summer over 70 institutions had signed memoranda of understanding agreeing to teach part or all of the courses, and to provide feedback in the form of comments, responses to surveys, and suggestions. The test program, which will run until June 1990, includes educational institutions in the US, Canada, Britain, Australia, New Zealand, Hong Kong, and Taiwan. Academic departments represented by test sites include Geography, Environmental Studies, Engineering, Geology, Urban Studies, Surveying, Anthropology, Regional Planning, Forestry, Marine Affairs, and Landscape Architecture.

The evaluation phase has two main objectives. First, is the material comprehensive and does it provide students with useful knowledge and marketable skills? Second, how useful are the materials to instructors actually preparing and teaching these introductory GIS courses? The results of the evaluation will be of interest not only to the NCGIA, but also to the funding agencies, the users of the curriculum, and the education community in general.

Although the evaluation phase relies primarily on written surveys filled out by both students and instructors using the materials, an emphasis is being placed on direct contact with individual evaluators. Users Group meetings are taking place at GIS/LIS '89 in November at Orlando and at the Annual Conference of the Association of American Geographers in April 1990, at Toronto. By bringing these evaluators together in one place, we hope that some consensus will be reached on issues related to the teaching of introductory GIS in general and the role of the NCGIA Core Curriculum in particular. The project featured in a number of presentations and discussions at the GIS in Higher Education Conference at Ohio State University in June 1989.

Responses to the test version of the materials have been received from various sources. Copies were sent to several vendors, and we have encouraged participation in the form of lab materials adapted to particular vendor products, etc. Requests for copies have come from many individuals and groups beyond the set of test sites, and we have distributed several additional copies at cost, but have not encouraged further distribution because of the test nature of the materials.

The final result of the evaluation phase will be a thorough revision of the draft materials. Although the details of the production of this final version are not yet complete, the materials will be available for wide distribution by the end of July 1990, at a cost of approximately \$200. Publication information will be widely disseminated prior to release. This process has already begun with the distribution of a brochure and announcements to newsletters.

Although the final version is not yet available, it appears that the curriculum project has been successful in providing materials which stress the conceptual underpinnings and theoretical basis of GIS, as well as providing example applications. It has drawn attention to the Center in areas where the research program might never reach, such as teaching-oriented institutions. The concept of distributed teaching materials is not common, but we feel that the project fills a useful gap between software and textbook, and has value in allowing an institution to move rapidly to offer GIS courses, which was after all its intention.

Presentations about the project have been made at the GIS in Higher Education conference at Ohio State and at a GIS and Education conference in Leicester, UK in March 1990, as well as at several smaller regional meetings. A first paper on the project has been submitted to the *Journal of Geography in Higher Education*, a refereed journal. A workshop on teaching the curriculum will be offered at the Fourth International Symposium on Spatial Data Handling in Zurich in July, and will hopefully be the first such course on using the curriculum. A paper on the evaluation program will appear in the *Proceedings* of the same conference.

The test version of the project contained lab materials, including instructions, data and answers. The labs referred to specific software packages, since it is virtually impossible to develop generic lab materials. The choice of packages, IDRISI and pcARC/INFO, was based on a survey of the commonest packages among the test sites. However although this seemed appropriate for the test version, it is too limited for the final release. Instead we have decided to drop the labs from the 1990 release, and to devote the 1990/1 academic year to the development of a much more comprehensive lab collection, including a much wider selection of vendor products.

In future we plan to explore several other ways of supporting education in GIS. One might be a course based entirely on cases, along similar lines to the case courses of the Harvard Business School. This might be a useful way of addressing many of the practical, managerial issues of GIS application.

Another is a collection of case histories of the establishment of GIS lab facilities - funding, equipment, politics, etc. - which might aid instructors interested in constructing labs.

Outreach

The establishment of the Center brought an immediate and gratifying response from the GIS community and related interests. In its first year, the Center has had to find a balance between many competing demands in the area of outreach, namely:

1. Promulgation of research findings, and development of the Center's national and international role as a research institution;
2. Interaction with the burgeoning GIS industry in the US, to ensure that research results are available for utilization and that the Center's research agenda is responsive to industry's needs;
3. Education and training to help correct the current scarcity of courses and programs and the lack of trained personnel to fill increasing numbers of positions, due to the comparative newness of the field; and

4. General assistance to the GIS community (academics, consultants, governments and other users of GIS technology), specifically, the provision of advice deriving from the Center's expertise and national role.

Numerous visitors from the US and other countries have toured one or more of the Center's sites, and many presentations about the Center and its research and educational programs have been made by Center personnel, both on and off site (full lists were provided in the Annual Report, and will be updated annually). Contact has been established with similar organizations in other countries, such as the Regional Research Laboratories funded by the Economic and Social Research Council in the UK, and some formal agreements for research collaboration have been initiated. During the first year of operation, Center personnel participated in over 100 separate conferences, made over 30 conference presentations specifically about the NCGIA, and visited almost 100 other institutions. Visitors to NCGIA sites came from well over 100 other agencies and institutions.

In addition to the Initiative Specialist Meetings, NCGIA hosted or co-sponsored several conferences or workshops. These included:

- 1) The International Atlantic Institute Research Seminar on "Land Information Management, Land Tenure, and Cadastral Systems" (Maine; July 27-29, 1989; co-sponsored with the University of New Brunswick and Université Laval);
- 2) The Fifth Annual New York State Geographic Information Systems Meeting (Buffalo, October 16-17, 1989; co-sponsored with the New York State Chapter of URISA);
- 3) Northeast Area Remote Sensing System (NEARSS) Association Meeting (Maine, August 25, 1989; co-sponsored);
- 4) Advanced Workshop on Movement, Migration, and Transportation: Regional Science Perspectives (Santa Barbara, November 1989; co-sponsored with the International Geographical Union, Commission on Mathematical Models);
- 5) Second Annual Geographic Data Workshop (TIGER) (Santa Barbara, August 25, 1989).

A twice-yearly newsletter (June and December) is published, with a current circulation of over 1000; overview presentations on the work of the Center have been made at conferences; and stories on the Center have appeared in many magazines, newspapers, and journals and on local television.

The Research Plan of the Center was published as a refereed article in the *International Journal of GIS* in 1989, and an article by Will Craig comparing the Center research agenda with the results of discussions within URISA was published in the *URISA Journal*. Update articles on the Center have appeared in numerous newsletters, and a regular column appears in the bimonthly *GIS World*. Major reviews of Center progress have appeared in special issues of *ACSM Bulletin* and *Government Information Quarterly*, and a short article on the Center was published in *Environment and Planning A*. All of these, along with numerous presentations and sessions at conferences, have helped to draw attention to the existence of the Center and its programs.

Two series of publications have been initiated, one of Technical Reports from research initiatives, and the other of miscellaneous publications, including materials presented at NCGIA-sponsored workshops. The Santa Barbara site, which is the primary distribution center for these publications, is now receiving approximately 50 orders per week, and has standing orders from many libraries. Publications are distributed at cost, but limited numbers of copies are distributed free to the research community.

Annotated list of selected outreach publications

Articles in refereed journals

NCGIA. The research plan of the National Center for Geographic Information and Analysis. *International Journal of GIS* 3:117-36 (1989).

This paper encompasses the major portion of the "Research Plan" from our NCGIA proposal. The plan was published in order to disseminate our plans and goals to the broader GIS and GIA research communities.

Goodchild, M.F. GIS and basic research: the National Center for Geographic Information and Analysis. *Government Information Quarterly* (to appear).

A summary of the Center's activities and progress in its first year, to appear in a special issue of *GIQ* devoted to GIS.

Articles in other outlets

Goodchild, M.F. The National Center for Geographic Information and Analysis. *ACSM Bulletin* (December, 1989) pp. 35-40.

A review of the NCGIA after Year 1.

Goodchild, M.F. and D.W. Rhind. The US NCGIA and some differences between it and the RRLs. *AGI Yearbook* (to appear).

A comparison of the NCGIA with the UK Regional Research Laboratories.

Goodchild, M.F. The NCGIA. *Journal of Surveying and Land Information Systems* (to appear).

A contribution to the US report to FIG 1990.

Goodchild, M.F., K. Kemp, D.M. Mark and A.U. Frank. The U.S. National Center for Geographic Information and Analysis: an overview of the agendas for research and education. *Proceedings, Canadian National Conference on GIS: Challenge for the 1990s* (Ottawa, March 1989) pp. 918-925.

Mark, D.M., B.P. Battenfield, H.W. Calkins, M.F. Goodchild and H. Onsrud. Geographical research findings at the NCGIA during year 1: summary and implications for GIS. *Proceedings, Canadian National Conference on GIS: Challenge for the 1990s* (Ottawa, March 1990) (in press).

The NCGIA was established to conduct basic research on geographic information systems and geographical analysis. This paper reviews major results obtained on the first four major Research Initiatives of the NCGIA. Geography as a discipline is very important to the NCGIA's research agenda and to the development of GIA and the inclusion of spatial analysis in GIS. GIS itself is important for Geography, as a test-bed for the evaluation of geographic theory and as a link between theory and application.