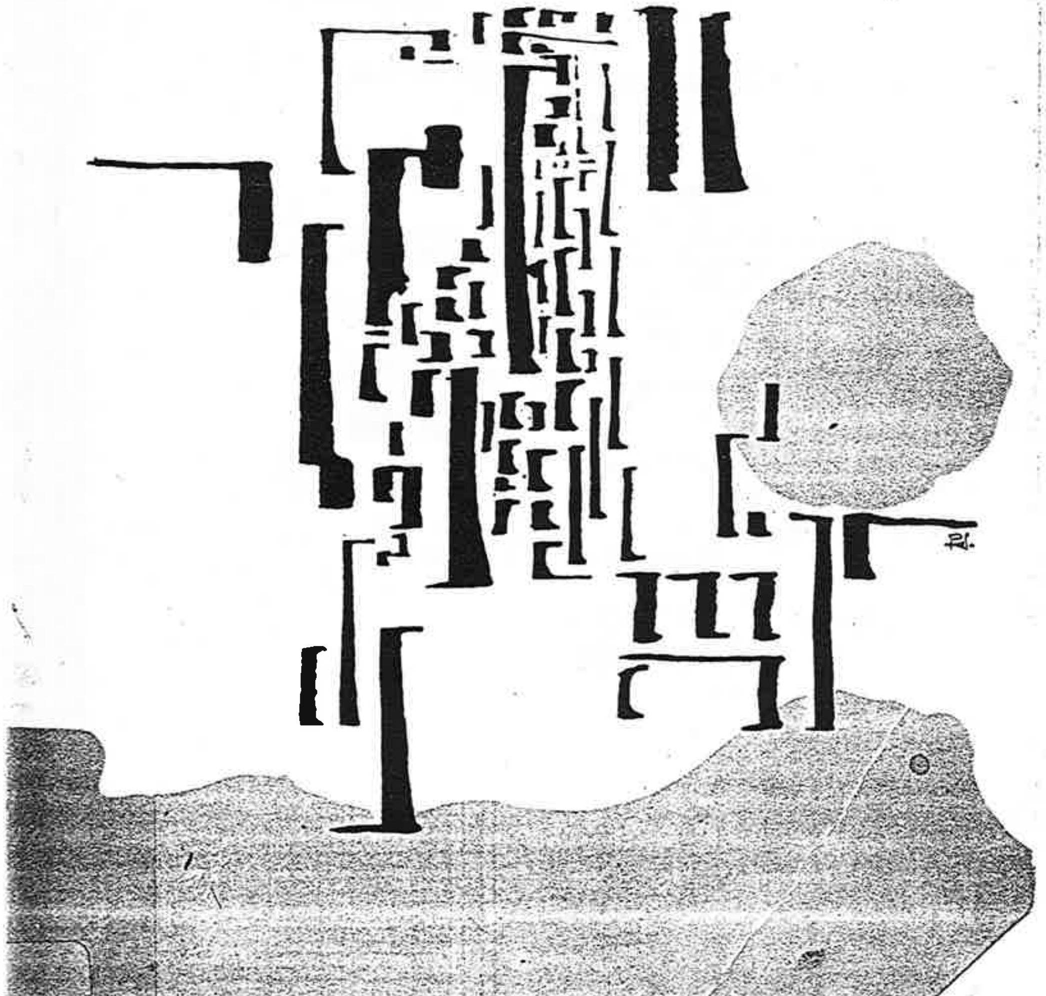
ENVIRONMENT INFORMATION SYSTEMS

EDITED BY R.F. TOMLINSON



THE PROCEEDINGS OF THE UNESCO/IGU FIRST SYMPOSIUM ON GEOGRAPHICAL INFORMATION SYSTEMS, OCTOBER 1970, OTTAWA, CANADA. A PUBLICATION OF THE INTERNATIONAL GEOGRAPHICAL UNION COMMISSION ON GEOGRAPHIC DATA SENSING AND PROCESSING.

ENVIRONMENT INFORMATION SYSTEMS

Edited by R.F. Tomlinson

The Proceedings of the UNESCO/IGU
First Symposium on Geographical
Information Systems
Ottawa, September 1970

A Publication of the International
Geographical Union Commission on
Geographical Data Sensing and Processing

Library of Congress
Catalogue Card Number
76 - 155485

TABLE OF CONTENTS

TABLE OF CONTENTS

TABLE OF CONTENTS

TABLE OF CONTENTS

FIRST SYMPOSIUM ON GEOGRAPHICAL INFORMATION SYSTEMS

OTTAWA, CANADA

September 28th - October 2nd, 1970

SUMMARY OF PROCEEDINGS

The First Symposium on Geographical Information Systems was held in Ottawa, Canada, between September 28 and October 2, 1970. The meeting was held under the joint auspices of the Natural Resources Division of UNESCO (with the co-operation of the Canadian National Commission) and the Commission on Geographical Data Sensing and Processing of the International Geographical Union.

The object of the First Meeting was to explore the scope for environmental data handling in digitized form and how it may be used to shorten the period between the collection of the data and their utilization.

The essential difference between most data and those describing the environment of the surface of the earth is that the latter frequently have a location identifier as part of the data element. For the purposes of the Symposium, "geographical data" were considered as being those which allow this location value to be manipulated in concert with the remaining value of the data. Throughout the Symposium the terms "geographical data" and "environmental data" were used synonymously as were the terms "geographical information system" and "environment information system".

The emphasis of the First Meeting was on the acceptance, storage, manipulation, retrieval and display of geographical data. The participants reported on and clarified the state of development of the various systems and capabilities of manipulating geographical data. As the larger systems that handle such geographical data contain a variety of capabilities that can be at various stages of development and as there is not in general any one commonly accepted set of categories of these various stages of development, it was thought useful prior to the Symposium to define the following categories which were used during the Symposium and in the following proceedings:

Category	Criteria
1) IMPLEMENTED	Being used on a routine basis in operational systems that have either been documented or can be visited and seen.
2) DEVELOPED	Conceptual work finished. Where applicable, computer code written and error free to a minimum workable level. Economically acceptable to users. Not yet documented or in routine use.
3) FEASIBLE	Conceptual work finished. Where applicable, computer code being written, but not checked for errors at all. Cost of use not determined in detail, but can be estimated.
4) EXPERIMENTAL	Conceptual stage being worked on. Various approaches being tested. Final methods not determined.
5) PLANNED	Manipulation capabilities that are required and are funded for research and which are logical extensions of current manipulation capabilities.
6) REQUIRED OR POTENTIAL	Anything else other than the above.

The First Meeting was one of invited participants representing nations who had experience in developing geographical information systems. A total of forty-eight persons participated from nine countries. These were: Canada, England, France, West Germany, Israel, Scotland, Sweden, U.S.A., and the U.S.S.R. A full list of participants is appended to this report.

The format of the Meeting departed somewhat from the usual conference format and the difference is reflected in the main text that follows this summary. The first day of the Meeting was used to introduce the participants to one another, to present the spectrum of approaches to environmental data handling and to identify the sources of expertise that were to be referred to during the remainder of the week. Symposium participants gave brief outlines explaining

the fundamentals of their work. Summary notes on these outlines are given in Section I of this book and are classified and commented upon in the report of the discussion that follows. The daytime sessions on the second, third and fourth days were used for the examination of the phases of input; use and manipulation; and display of geographical data respectively. One paper only was given each day. Rather than being a report on a specific piece of work, this paper was an index paper of the subjects under discussion, giving a statement of what was currently thought to be possible or impossible with reference to the work of participants in the audience who could explain the processes or difficulties. During and/or following the paper and under the direction of the Chairman-of-the-Day, the participants commented on all aspects of the paper relating to experiences they had had in the areas under discussion, pointing out the weaknesses or merits of a particular approach and providing graphic or other illustration of the topic under consideration. The discussion was particularly directed to ascertain the state of development of the capabilities under discussion.

Each day a Rapporteur, or group of rapporteurs, provided an analysis of the day's discussion. This was written each night and distributed to the participants the following morning as a working document. The participants then commented on the working documents during the Symposium.

The Symposium was thus a process of both presentation and examination of current ideas and concepts, the results of both of these functions being presented in the main text of this report. The process relied totally on the participants and, particularly, on the Chairman-of-the-Day, the Index Speakers and the Rapporteurs. These were:

DAY ONE	CHAIRMAN	Dean Edson	U.S. Geological Survey
Monday	SPEAKERS	System Outlines by Participants	
September 28	RAPPORTEURS	Warren Schmidt	C.I.A., Washington
		Tom Waugh	Edinburgh University
DAY TWO	CHAIRMAN	Warren Grabau	U.S. Army Engineers
Tuesday	SPEAKER	Ray Boyle	Univ. of Saskatchewan
September 29	RAPPORTEURS	David Sinton	Harvard University
		Alan Schmidt	Harvard University

DAY THREE Wednesday September 30	CHAIRMAN SPEAKER RAPPORTEURS	Duane Marble Peter Kingston Carl Steinitz David Sinton Alan Schmidt	Uppsala University IBM, Halifax Harvard University Harvard University Harvard University
DAY FOUR Thursday October 1	CHAIRMAN SPEAKER RAPPORTEUR	Bill Aumen Waldo Tobler Victor LaGarde	U.S. Army - TOPOCOM University of Michigan U.S. Army Engineers

On the final day, Friday, October 2, the rapporteurs presented their working reports to the Symposium for comment and discussion.

Evening sessions on the second and third day were held as informal discussions to talk through the less structured areas of concern to the participants. These areas included the compatibility between various levels of environmental data systems and between data in different areas and, also, the security and confidential nature of environmental data. Invited guests with technical expertise relevant to the discussion, including lawyers and Members of Parliament, participated in these informal evening discussions. As they were structured for an informal exchange of views, the evening sessions were deliberately not reported upon and are not included in this report. The concepts developed, however, may be found reflected in certain aspects of the daytime discussions.

As indicated earlier, the overall objective of the Symposium was to explore the scope of environmental data handling. During the course of the meeting, systems with a wide range of capabilities were identified and it may be worthwhile in this summary of the proceedings to present a tentative framework, or general systems theory, within which the relative capabilities of the systems can be understood and, by extension, can be related to user requirements to handle location specific data.

The three basic parameters of a geographical information system are: a) the character of the location identifier; b) the volume of pieces of data carried in concert with any one set of location identifiers; and c) the manipulation facilities that the system possesses. These can conveniently be thought of as three axes of a cube (Fig. I).

From the proceedings of the Symposium, it was possible to recognize various incremental values along these axes and to use them as identifiers of system character. As more information becomes available about the various systems, it will be possible to regroup

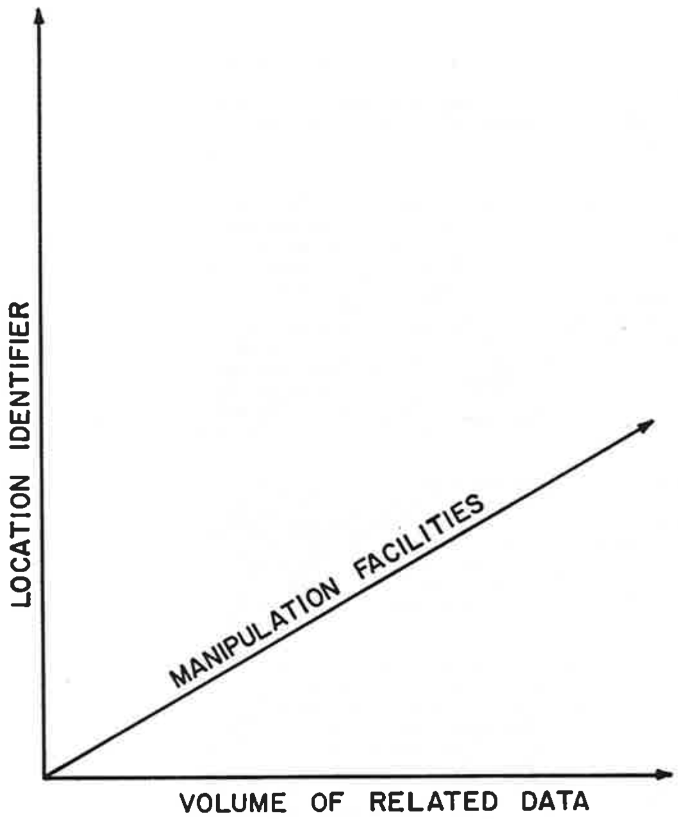


FIG. I

and refine the values. At this time, they only represent a useful way to think of the steps in the progression along each axis.

Location Identifier

There are essentially three categories of methods for identifying location (Fig. II). The simplest of these is the external index. The location identifier in this category that can be applied to individual data elements, or to groups of elements, can be generally descriptive; examples are names of administrative areas, street addresses, postal zones or names of arbitrarily placed grid units. The key characteristic of this category of location identification is that the position of the data element cannot be determined without reference to a master index, external to the system, that shows the boundaries of the named places. It is probably reasonable to say that, of non-mapped data having any geographical indexing today, the overwhelming proportion is by this descriptive method.

The second category is the allocation of a co-ordinate value as a location identifier or index to an otherwise unspecified area. (Lines are considered to be areas which have no significant dimension on one axis.) A typical application of this category is the use of centroid co-ordinate allocation of values to irregularly shaped administrative areas. This level represents the first at which the relative position of the data elements are known to the systems.

The third and most specific category is where the boundary of the data is known implicitly or explicitly to the system. The lowest order within this category is the use of an arbitrary grid where the co-ordinate position of the data element implies a boundary known to the system even though that boundary is not explicitly carried in the system. Explicit boundaries form the higher orders. Lines described by widely spaced nodal points can be thought of as simple examples of explicit boundaries. The hierarchy extends upwards as the number of points used to describe a line and the resolution of the image increase.

Volume of Related Data

The second parameter (Fig. III), that of the volume of pieces of related data that can be carried at any one location within one set of location identifiers, is the second axis of the record plane (the first axis being that of the location identifier).

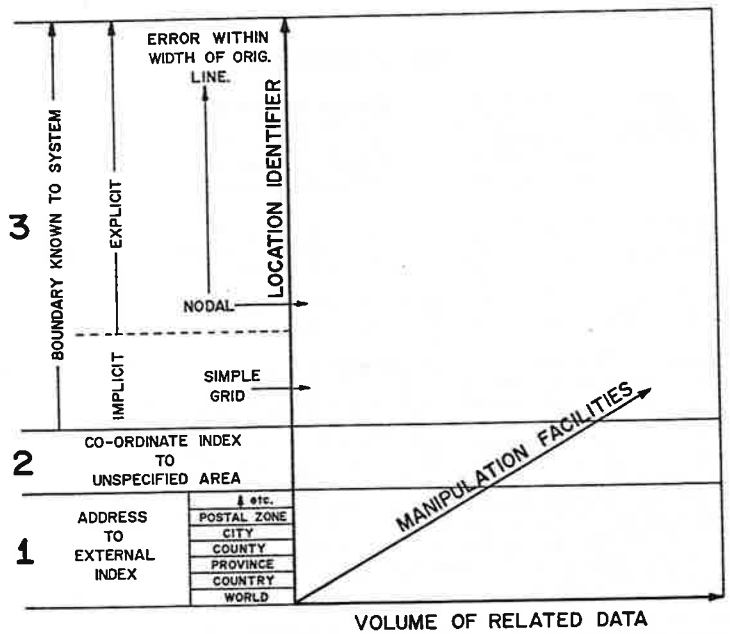


FIG. II

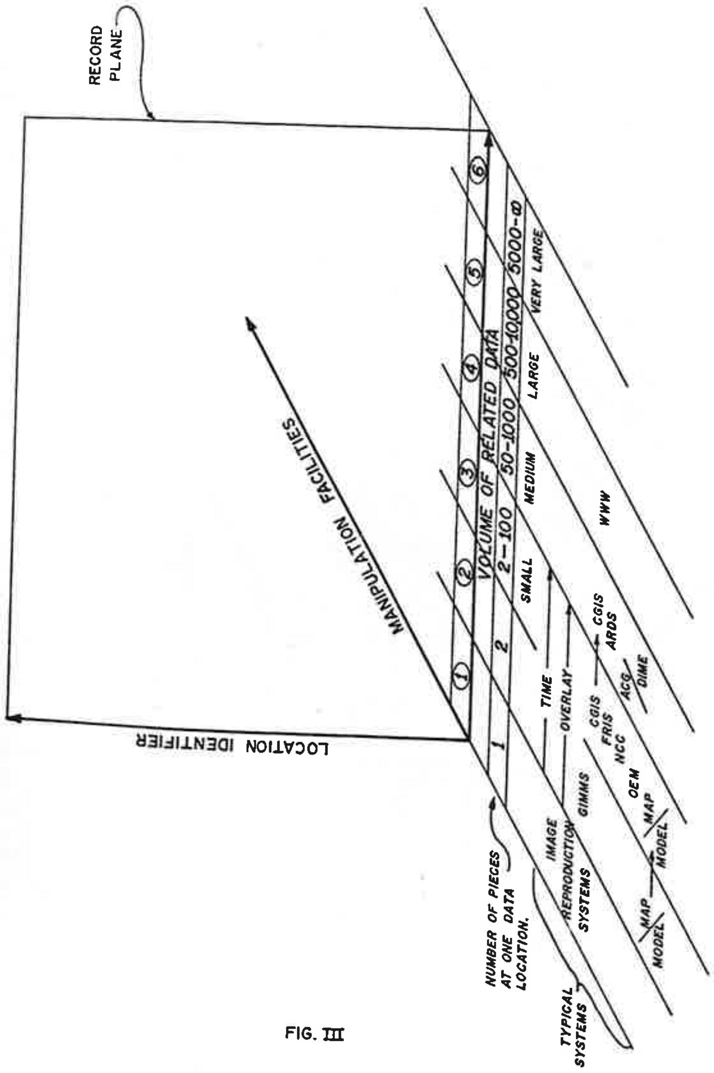


FIG. III

SUMMARY OF PROCEEDINGS

During the Symposium, reference was variously made to binary bits of storage (representing the most efficient form of compaction and, hence, the true value of data volume that has to be handled by any system), and pieces of data, these latter being the recognizable elements of real world data. The relationship between the bits and pieces can be expressed as the sum of the number of the classifications or groupings used (or at least the next highest binary number that will hold that number of classifications) for each category or piece of data or data element, i.e.

$$\sum_{\text{Category} = 1}^{\text{Category} = n} \text{Classification Groups} = \text{magnitude of bits of data}$$

With the relationship expressed, pieces will be used as the units to subdivide the volume of related data axis, as the relationship between pieces of data and the content of actual systems is considerably easier to understand. At the lower end of the axis there are two clear subdivisions. Further along the axis the picture is less clear, both as to significant groupings and as to the exact data content of any one system. In this part of the axis, overlapping ranges have been used to distinguish small, medium, large and very large volumes of related data.

The first subdivision is that which contains one piece of data per location. This represents an entity that stands by itself, such as a snow/no snow condition at a particular location or, more commonly, a road/no road condition at a particular location. This subdivision is the one containing the commonly recognized image holding bit-plane, found in image reproduction systems.

The second subdivision is that record which can contain two pieces of related data per location. This is a significant subdivision in that it is the first point in system content where pieces can be related, at one location. It is the first point at which time can be considered in that data string and is also the data level where the use of an overlay manipulation capability is first possible. Categories of systems within this subdivision include GIMMS (69), the Canadian Hydrographic System (6), and simple versions of the MAP/MODEL (3) system (though this latter is capable of extensions into the third subdivision).

The third subdivision contains data records with a small number of pieces of data at each location, typically 2 to 100. Into this subdivision come such systems as FRIS (8), NCC (54), OEM (42,43), and the current status of the CGIS (55), though this latter system has the capacity of being expanded into the fourth subdivision.

The fourth subdivision contains a medium number of pieces of data associated with each location identifier, typically 50 to 1,000. Included in this category are the data-intensive census type of records typical in the GRDSR System (18), ACG/DIME (1), and ARDS (9).

The fifth subdivision contains a large number of pieces of data associated with each location identifier, typically 500 to 10,000. This subdivision is currently characterized by data strings enlarged by the time multiplier. An example of this would be the information contained about one location on a ten-year series of weather maps, each produced every hour. The existing World-Wide Weather System has data records that are in this subdivision.

The sixth subdivision contains a very large number of pieces of data associated with each location identifier. The range is suggested as extending from 5,000 to infinity. No system at the Symposium had a data volume that fell into this category.

The six subdivisions described can be recognized as having grouped various types of systems together, and present a sequential order of magnitude of impact of the related data volume on the file structures that have to be incorporated in the various systems. In the diagram (Fig. III), the six subdivisions have been treated in a linear fashion. No attempt has been made to evaluate the present impact of related data volume on system design at this stage. It certainly may not be linear.

Manipulation Facilities

The third parameter of geographic information systems is that of the manipulation facilities incorporated in the system. Manipulation facilities are considered in terms of logical operations performed upon the data to produce a desired result. An operation is a change in data value, a comparison or a move of the data element. A hierarchy of manipulative facilities can be seen as those which incorporate increasing numbers of logical operations. These are laid out along the third axis of the cube (Fig. IV). There are three significant categories of system that can be identified: those without overlay capability; those with overlay capability; and those with a significantly developed system monitor

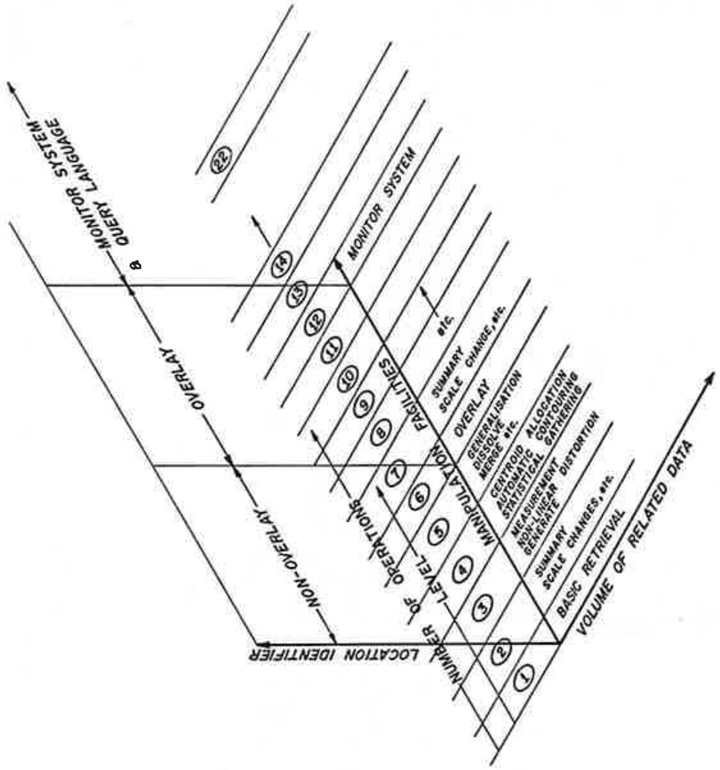


FIG. IV

and/or query language.

Within the first category, the hierarchy of data manipulation can be laid out. The simplest (level 1) is the single operation of basic data retrieval. The second (level 2) is the result of two operations and these non-interactively can produce data summary, linear distortion elimination, classification change, selective search, scale change, and projection change. The non-interactive capability of carrying out more than one of these manipulations, obviously, establishes a sub-hierarchy within level 2, but as long as they are non-interactive, the system is still considered as being within level 2. The same sub-hierarchical grouping can be applied within other higher levels. SNAPS (65) and the World-Wide Weather System can be considered as examples of level 2 systems.

The third (level 3) is the result of three operations. These can result in the capabilities of measurement, non-linear distortion elimination and the ability to generate (circles,polygons,points). A system in level 3, or any other level, may not automatically include all the manipulation facilities contained in the previous levels, but the manipulations that it can carry out are of a more sophisticated nature. Typical of the systems within level 3 are NCC (54), AUTOMAP (48), OEM (42,43), ACG/DIME (1), and GRDSR (18) (the latter with its ability to do selective search given polygon specifications).

Level 4 is the result of four operations used in sequence, and these provide the manipulative facilities of centroid allocation (as measurement effectively has to be done first), automatic contouring and statistical gathering. SYMAP V (26) with its facility of automatic contouring may be considered at level 4.

The fifth (level 5) is the result of five operations used in sequence. This results in the facilities of generalization, dissolve, determination of inter-visibility, merge and simulation. Typical of this level are systems with the facility of hill shading or route determination.

At this point comes the second category incorporating the overlay function which takes six operations by itself. The preceding levels in the first category can be repeated as levels in the second category (with the addition of six existing operations on the front of each manipulation) providing levels 7 through 12.

Level 8 is thus a system with overlay capabilities that can

also carry out manipulations up to level 2, either before or after the overlay manipulation. Typical of the level 8 systems are: GIMMS (69) and the U.S. Army Corps of Engineers, Waterways Experimental Station System, as described at the Symposium.

Level 9 incorporates overlay capability within the facilities of level 3 and is typically represented by the MAP/MODEL (3) systems.

The third major category incorporates a significantly developed or implemented system monitor and/or a query language. This may exist with or without the overlay capability and, using a base of 12, provides levels 13 to 24. Its position on the axis thus need not be a linear progression from the overlay; it is placed there merely for convenience and, perhaps, to indicate a significantly high level of manipulative capability. The system of levels could be continued through this third category to indicate relative sophistication of manipulative capability. The only system known to have a full system monitor in conjunction with overlay and all lower level manipulative capability is the CGIS (55) which would have an approximate level of 22. The USGS (41) system now in the experimental phase and developments of the U.S. Engineers Waterways System anticipate this level of manipulative facility.

The parameters are combined in Figure V and the approximate location of some typical systems has been indicated within the framework. The diagram itself is not, and cannot be, thought of as a precise tool for relating system capabilities, though the conceptual framework that it provides allows some observations to be made.

The diagonal OS (see Fig. VI) essentially represents the potential capability of a file structure to handle information in any location specific information system. Progression along the diagonal from O to S is a progression of file structure design sophistication combined with machine data handling capability.

If the diagonal axis is divided arbitrarily between O and S into ten equal parts, the area of the cube passed through by the lower three units encompasses parameters that together may be considered as the area where current manual methods of storing and analyzing data are economic and satisfactory to the user.

In the area of the cube between levels three and four on the diagonal, are the proliferation of systems that represent the current state of the art, according to their need to handle images or

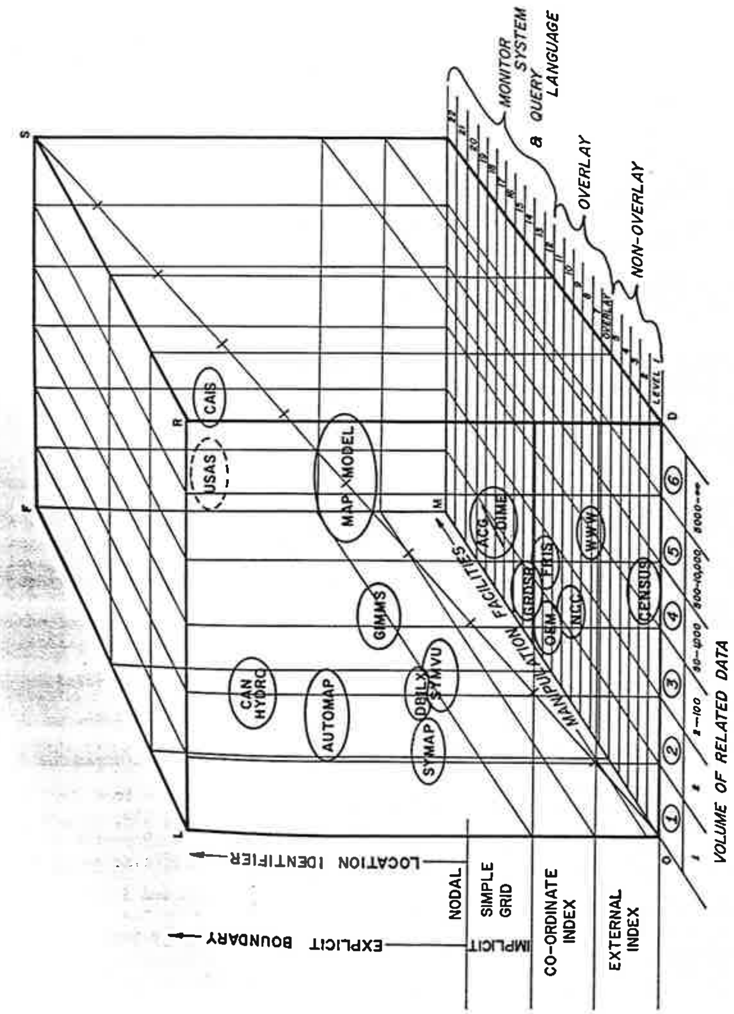

MONITOR SYSTEM
QUERY & LANGUAGE
OVERLAY
NON-OVERLAY
LEVEL
CAIS
USAS
MAP MODEL
ACG DIME
FACILITIES
MANIPULATION
GRDSR
FRIS
NCIC
WWW
CENSUS
CRM
GIMMS
CAN HYDRO
AUTOMAP
DBILX
SYMVU
SYMAP
VOLUME OF RELATED DATA
LOCATION IDENTIFIER
EXPLICIT BOUNDARY
IMPLICIT
NODAL
SIMPLE GRID
CO-ORDINATE INDEX
EXTERNAL INDEX

FIG. V

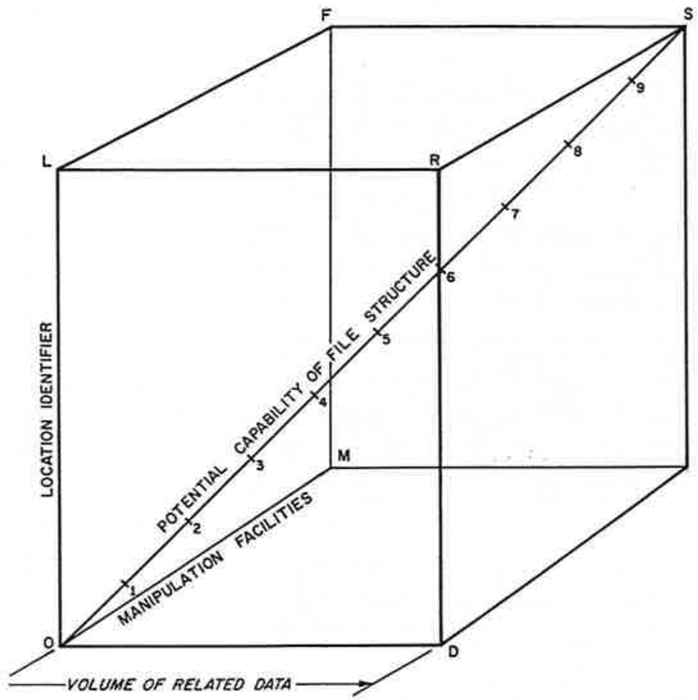

F
S
L
R
M
O
D
LOCATION IDENTIFIER
POTENTIAL CAPABILITY OF FILE STRUCTURE
MANIPULATION FACILITIES
VOLUME OF RELATED DATA

FIG. VI

related data sets. These include the SYMAP (26), SYMVU (26), ACG/DIME (1) and GIMMS (69) type of group on the image side and the GRDSR (18), OEM (42,43), NCC (54), and FRIS (8) type of group on the related data side. These probably represent the level of the file structure design coupled with machine capability that has currently found economic acceptance. The development of further systems with the various parameters that fall into this category should find no difficulty in being implemented or being efficiently used.

As one proceeds up the diagonal the file structure necessarily gets more sophisticated. The type and capacity of the data processing machine that must be used, coupled with the degree of skill with which it must be used to achieve efficiency, similarly increase. Typical of the systems at the five to seven level are the image handling systems of AUTOMAP (48), the Canadian Hydrographic System (6) and the MAP/MODEL (3) system.

Perhaps at one level higher is the type of system represented by the CGIS (55), by reason of the fact that they attempt to carry more related data than the systems predominantly concered with the manipulation of images. At this level, efficiency of operation is totally related to file structure and the optimum use of current computing facilities.

Apart from allowing some understanding of the relationship between characteristics of the various systems, the use of this approach to examining geographic information systems does not have as its primary purpose the task of establishing a hierarchy of system sophistication. Rather, the use of this approach allows the systems to be looked at in terms of available environmental data and the effort that is required to handle them. Just as the systems may be described with the parameters of location identifier, volume of related data and manipulative facility, so may the requirements to handle data to better describe and understand the environment be described in these terms. In parallel to the cube or matrix relating the systems, a similar cube or matrix can be developed to specify user data handling requirements (Fig. VII).

Given an understanding of the parameters of system capability as initially provided at the symposium and summarized in this manner, it may be possible to develop the relationships between system capability and environmental data handling needs.

Further, it is possible in the same terms to relate the geo-

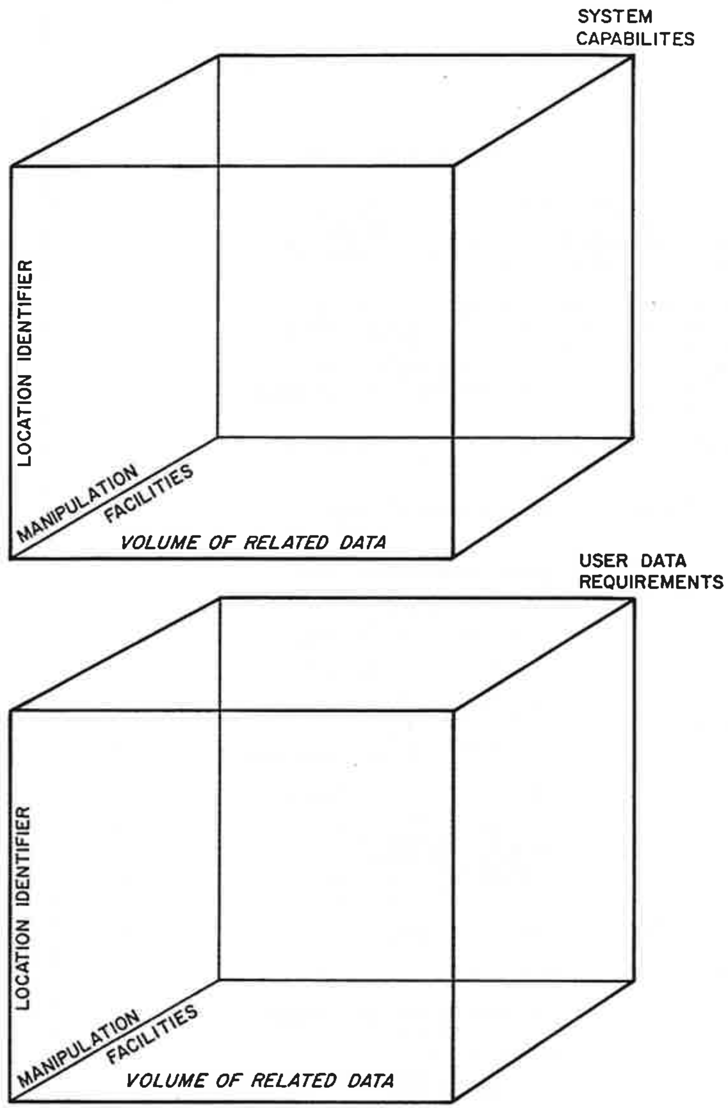


FIG. VII

graphical information system capability to the capabilities of sensors to gather data. In similar terms, it may be possible to relate the stated user requirements for data to the parameters of the decisions that have to be made from them (Fig. VIII).

During the course of the Symposium, Working Groups were established to examine the various elements of this continuum. These were added to the existing working group on geographical data sensing that was formed prior to the Symposium and had its first meeting during the course of the Symposium. The resulting structure of Working Groups is as follows:

1) Geographical Data Sensing (including photo-interpretation and remote sensing)

Chairman:	Professor D. Steiner	- University of Waterloo
Members:	R. Alexander	- U.S. Geological Survey
	U.V. Helava	- Bendix Research Laboratories
	R. Tress	- Air Photo Analysis Associates, Toronto
	W.A. Brooner	- University of California, Riverside
	C.N. Johnson	- University of California, Riverside
	F.E. Horton	- University of Iowa
	H. Gierloff Emden	- University of Munich
	A. Gregory	- Dept. Energy, Mines & Resources, Ottawa
	P.A. Langley	- U.S. Dept. of Agriculture & Forestry
	A. Rosenfeld	- University of Maryland
	T. Nakano	- University of Tokyo
	S. Schneider	- Institut fur Landeskunde, Badgodesburg

2) Geographical Information System Definition and Review

Chairman:	Professor A.R. Boyle	- University of Saskatchewan
Members:	E.L. Amidon	- U.S. Dept. Agriculture, Forest Service
	R. Van den Driessche	- ORSTOM, France
	W. Grabau	- U.S. Army Engineers, Waterways
	J. Sharp	- U.S. Army, TOPOCOM

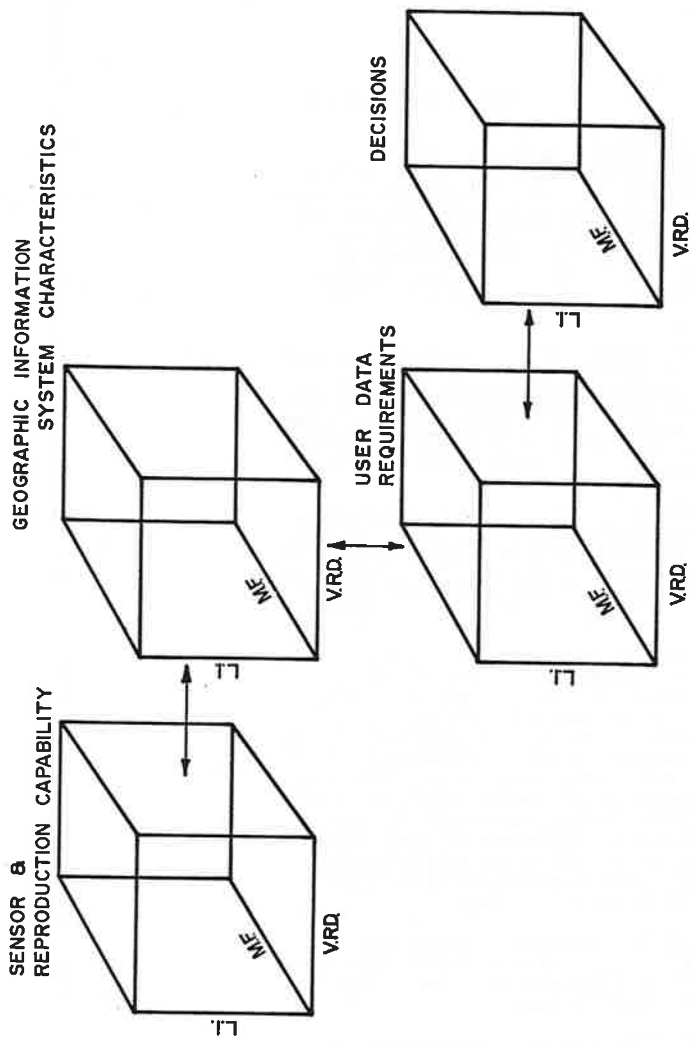


FIG. VIII

	P. Yoeli	- Israel Institute of Technology

3) Geographical Data Manipulation

Chairman:	Dr. V. LaGarde	- U.S. Army Engineers, Waterways
Members:	P. Kingston	- IBM Canada Ltd., Halifax
	O. Salomonsson	- Swedish Central Board for Real Estate Data
	D. Sinton	- Harvard University

4) Geographical Data Communication, Display and Dissemination

Chairman:	Dr. D.T. Edson	- U.S. Geological Survey
Members:	D.P. Bickmore	- Royal College of Art, London
	A.H. Schmidt	- Harvard University
	W. Schmidt	- Central Intelligence Agency, U.S.A.
	M. Guy	- Institute Francais du Petrole

5) Geographical Information Systems Use and Economics

Chairman:	Professor D.F. Marble	- University of Uppsala
Members:	A. Baker	- University of Toronto
	H. Calkins	- University of Washington
	K. Dueker	- University of Iowa
	J. Gilliland	- Dept. Energy, Mines & Resources, Ottawa
	J. Salmona	- Observatoire Economique Méditerranéen Marseille

There will be considerable interaction between the Working Groups as the elements in the continuum are examined. The output can be thought of as a series of products that are currently underway.

1. A review of the techniques available for the remote sensing of geographical data and a critique of their output.

2. A glossary of terms relating to geographical information systems - this to include norms of quality, system content, data reliability and data formats.

3. A file containing status reports of the current systems of accepting and storing geographical data, their location, capabilities and state of development. The classification and summarization of this file to result in a review of the systems status in March, 1972.

4. A monograph of geographical data manipulation algorithms documented and classified by manipulative characteristic.

5. A set of examples of graphic output from geographical information systems with the associated algorithms required to produce them, supported by a statement of items, including costs of production, time needed as a function of volume, resolution, size, neatness and number of colours.

6. A newsletter to emphasize system use and economics resulting in a review of system use and economics in March of 1972. Also an on-going census of persons involved in the development of geographical information systems classified by their current use.

7. A bibliography of the various reference works and bibliographies available on various aspects of the continuum.

The objective of the overall working group structure is to pull together the knowledge in the field and in the Spring of 1972 to publish a reference text and set of guidelines for the development of environment information systems in countries in various stages of development in the year thereafter. This text and the findings of the Working Groups will be the foundation of the second Symposium to be held in Canada between August 1 and August 9, in 1972. This second meeting will be one of an educational nature, where the experts from the first meeting will present the results of 18 months' findings to a wider audience. In particular, the second meeting will be aimed at being of direct use to those countries and agencies who have need to handle environmental data, but who need clear advice on the benefit, costs and technical problems associated with the various approaches and need to be able to refer to the existing expertise and personal experience.

OPENING ADDRESS

Given by: R.F. Tomlinson, Chairman of the International Geographical Union Commission on Geographical Data Sensing and Processing

I am delighted to have this opportunity today to welcome you to the First Symposium on Geographical Information Systems. This is the first of two symposia being held by the Commission on Geographical Data Sensing and Processing of the International Geographical Union (IGU) in conjunction with the United Nations Educational, Scientific and Cultural Organization (UNESCO). I would like to express my appreciation to UNESCO for their assistance, in particular to the staff of the Natural Resources Division in Paris and to the Staff of the Canadian Commission for UNESCO in Ottawa.

The purpose of this Symposium is exploratory: it is to determine what is being done in the field of geographical information systems and, further, to identify problem areas and to establish a reporting process, to monitor the work that will be done to resolve these problems.

We have before us rather a unique situation, in that the science of geography has developed extensively over the years, but it may now have reached a turning point. In recent years another science has pursued a parallel development. This is a science about which the general public knows very little indeed, though notwithstanding this lack of knowledge, the general public is fascinated by its methodology. This other science I refer to is data processing.

Many different disciplines have requirements to handle data, but surely geography, with its area of study being the whole surface of the earth and the universe of data that it uses being that which is necessary to describe the total environment, can use to the best advantage the capabilities offered by data processing.

The environment to which I refer is not simply the rivers and forests and natural countryside so beloved of pollution control activists. Our environment is the whole living space of mankind, with man himself, his creations and his actions as important elements within it. Data about our environment and the understanding of these relationships of the elements within it, are what we have in the past called geographical data, and the study of

OPENING ADDRESS

geography. This First Symposium on Geographical Information Systems could equally be called the First Symposium on Environmental Information Systems.

Very large amounts of effort and money are spent each year in gathering information about the environment. They have been for many years and will be for many years to come. Yet, it is probably a fair comment to say that there is less understanding on the part of the decision makers about the actual conditions of the environment and the relationships between man and the land than there is about any other subject with which they have to deal. Urban administrators are faced with a great lack of up-to-date information about their cities on which to base rational decisions about their civic problems or the direction of their future growth. Similarly, regional, and to an even larger extent, national governments are faced with in ever-increasing information vacuum. I say vacuum not because the data are non-existent - indeed there are almost too many data - but because without massive effort the data do not readily supply information that can be used. The result is an effective environmental information vacuum around our decision makers at all levels of government.

The development of geographical information systems to date has been going on piecemeal, with different users applying different approaches in different areas. Research funds are not forthcoming in adequate amounts, because the overall approach lacks direction, and is not clearly related to the overall objectives and needs at the national level that are necessary to gain support. Not only is there not enough money, but there are not enough skilled research workers to allow rational development to take place in this manner. The result of developments so far has been a proliferation of systems which are incompatible for handling the various kinds of data. There is a very real danger that our progress may become stultified unless we can develop a program for the overall direction of our work and for ensuring minimum levels of compatibility between the systems that we develop.

A strong case can be made at the national level to examine the needs of an entire country in terms of the decisions that will have to be made about the environment in the next ten years, and then to formulate a national research and development program to ensure that systems to handle the data will be available. If this is not done, the current approach of incompatible, piecemeal development will result in chaotic or, at the least, uninformed

OPENING ADDRESS

decision making. The environment is enormous; there will be too much environmental information available even with the use of current capabilities for data gathering in the next ten years, let alone the next one hundred years, for it to be handled efficiently and economically. The need, then, is to be selective about the data which will serve as a basis for the decision making process. For this reason, the data required must be chosen by the agencies responsible for their use, and not, as at present, by the agencies responsible for their collection. The development of systems to store, analyze and use the data had best be developed in the agencies and by those persons responsible for decisions about the environment, not in the fundamental data collection agencies.

The application of data processing techniques to the task of handling environmental data requires the design and use of systems incorporating both computer hardware and software elements. The particular selection of the type of hardware and software depends on the use to which they will be put. The criteria that determine which system should be used include such diverse ones as: the volume of information which has to be handled; the required response time to requests; the frequency of each request; the similarity of inquiries; the volume of inquiries; and the peak load conditions. The application of the capabilities of a computer to data handling obviously imposes a particular discipline on the thought and design of the process. It requires a clear understanding of the interdependence of the elements in that continuum of data gathering, reduction, storage, analysis and presentation. It also requires a clear understanding of the uses to which the information will be put. It is apparent that the only rational way to develop or evaluate the future acquisition of data about the earth is in terms of their use. The only way to ensure that systems are developed to provide current information that can be used for decision making purposes, and to prevent the systems from becoming clogged with irrelevant or out-of-date information, is to ensure that there is excellent understanding between data collector and data user.

How to bring about this change of attitude vis-a-vis data collection is difficult to determine. The first step is to examine the currently available technology, be it ever so different in application, and search the processes involved in the light of the need for fundamental data for decision making. This step is perhaps the responsibility of those engaged in the developing of systems for data handling. They must report on the current state of their

OPENING ADDRESS

development and formulate policy on what further work need be done. This first step is being begun here today. An interim report on the systems discussed will be produced at the end of this conference. The working groups formed will monitor forthcoming research in the problem areas and report to the Second Symposium in 1972. I hope that it may be possible to persuade the governments responsible for decisions about the environment that the answers to their questions must come from the types of technology which will be discussed here in the next few days, and that the development of systems for handling environmental data must be related directly to their needs. The task which you have accepted in coming to this conference is not, as you can see, an easy one; it is, however, a significant one. Immediate access to relevant information has become a key factor in the decision making process. We have before us an opportunity to facilitate that process, by the development of information systems geared to the needs of our decision makers.

On the first day of the Symposium, participants gave a very brief review of their current work. The following are the names of the speakers, the organizations that they represent and a capsule note of the system or work described.

SPEAKER	ACTIVITIES/SYSTEMS DESCRIBED
E.L. Amidon U.S. Department of Agriculture Berkeley	MIADS/2 - Map Information and Display System (2) An implemented grid manipulation system employing line printer output with tabulation capabilities for single or overlaid maps.
David F. Sinton Laboratory for Computer Graphics and Spatial Analysis Harvard University Cambridge	SYMAP - Synagraphic Mapping Program (26) An implemented line printer mapping system (generalized package) capable of producing choropleth, contour and proximal maps based on grid manipulation.
Arie Shachar The Hebrew University Jerusalem	The Urban Atlas of Jerusalem is completed (49). It was generated from a developed data bank of block inventory combined with a grid manipulation system. Retrieval is via tables or SYMAP-like programs.
Dieter Steiner University of Waterloo Waterloo	GEOMAP - Geographic Mapping Program (52) An implemented grid manipulation system similar to SYMAP which produces shaded maps (Choropleth or isorismic) on a line printer. The Information System for Switzerland presently under development uses a data bank of three files: statistical; grid; and line data.

SPEAKER	ACTIVITIES/SYSTEMS DESCRIBED
Robert Tweedie N.Y. State Department of Transporation Albany	(a) The Transporation Information System based (38) on grid manipulation incorporates, among others, land use and travel characteristics (geo-coded) in its data base. It is implemented and produces line printer dot maps as output. (b) The Highway Inventory Information (57) System is based on a data bank incorporating such items as the physical road characteristics, a road inventory, bridge records, traffic volumes, and has developed status.
Warren Schmidt C.I.A. Washington	AUTOMAP - Automatic Mapping System (48) - has been operational since 1967 and can produce coastlines and any form of line or point data. It is a map compilation program at the world level. TOPIC is a system allowing line-of-sight analysis based on earth profiles. It is in the implemented stage.
Pinhas Yoeli Israel Institute of Technology Haifa	The Israel Computer Mapping is not a system, (70,71,72) but rather various projects done in the areas of altimetric content, lettering content, and planimetric content on maps. In the light of this, experiments have been done on the creation of terrain models by digitizing contours as well as automatic hill shading.

SPEAKER	ACTIVITIES/SYSTEMS DESCRIBED
Bill Aumen TOPOCOM U.S. Army Washington	SNAPS - Symbol and Names Placement System (65). This is a system to place names into a magnetic tape system and plot them back out onto a film projecting plotter. In so doing, one can develop the names overlay. A three-dimensional modelling system has been implemented which uses terrain elevations as data for milling machines. This is a map compilation system.
Owe Salomonsson Swedish Central Board of Real Estate Data Sundbyberg	The Spatial Information System (8), a pilot data handling project containing 35,000 people records, is an attempt to integrate various data files by cross-indexing techniques on administrative data systems. It is experimental and based on a grid manipulation.
Reimer Herrmann University of Wurzburg Munich	The Munich Air Pollution Survey system employs remote centers feeding a data bank. From this a meteorological forecast is done. The plotting of forecast maps has not yet been implemented. It is a grid manipulation system.
Tom Waugh University of Edinburgh Edinburgh	GIMMS - Geographic Information - Manipulation and Mapping System (69). A system which combines areal information with a manipulation capability to produce maps via a plotter, SYMAP or GRID.
John Weldon Dominion Bureau of Statistics Ottawa	GRDSR - Geographic Referenced Data Storage and Retrieval System (18). A developed geographic system based on geographical indexing designed for operation with the 1971 Census.

SPEAKER	ACTIVITIES/SYSTEMS DESCRIBED
Hugh Calkins Urban Data Centre University of Washington Seattle	(a) Urban Data Centre. This system, similar to the DBS geo-coding system, has been implemented for three years. It incorporates ARDS which is a CRT display unit manipulated by computer Displays Inc. (9). (b) A Network Analysis System which is experimental, employing light pen input with on-line error correction and manipulation capabilities.
David Symons National Capital Commission Ottawa	The NCC Information System (54) is based on six modules, none of which have been implemented. They are: (i) a geo-coded data base; (ii) a geo-coding system; (iii) xy co-ordinate retrieval with point-in-polygon routine (developed); (iv) graphical output; (v) a modelling routine; (vi) capital improvements schedule.
Robert Keith University of Oregon Eugene	MAP/MODEL system (3), an implemented system based upon the segregation of data collection, incorporates overlaying techniques by a mathematical overlaying process.
Ray Boyle University of Saskatchewan Saskatoon	The Canadian Hydrographic System (6) is a graphic data handling system incorporating a universal grid co-ordinate system. It is intended to produce hydrographic charts of high accuracy from survey data digitized automatically at sea. Manual digitization is also available for existing data in that form.

SPEAKER	ACTIVITIES/SYSTEMS DESCRIBED
John Foster Dept. of Regional Economic Expansion Ottawa	The Canada Geographic Information System (55) will be utilized in the planning of land resource use. Development is almost completed on boundary capabilities. The system will incorporate soil, land use, forestry, wildlife and recreation coverages. It is a full graphic data handling system.
Dean Edson U.S. Geological Survey Washington	The U.S. Geological Survey System (41) is intended to provide base mapping for the United States and will be completed in another 10 years. It is based on a 7-1/2 minute quadrangle mapping at various contouring levels. Hardware and software are presently being developed to merge new map information onto the base maps. It is a geographic data handling system in the development stage. A National Cartographic Information Centre has recently been authorized, and this will supply the total spectrum of cartographic data to the map using public. The Earth Resources Data Centre is intended to aid in the receiving, processing, dissemination of data from the earth resources satellite.
Warren Grabau U.S. Army Corps of Engineers Waterways Experiment Station Vicksburg	The Waterways System, an integrated information system, incorporates integration of data acquisition storage, retrieval and manipulation. It stores primary data and has overlay capabilities. Some parts are operational, some are still at the conceptual stage. This is a geographical data handling system.

SPEAKER	ACTIVITIES/SYSTEMS DESCRIBED
Jean Salmona Observatoire Economique Mediterraneen Marseille	OEM is a pilot information centre (42,43). There is a block level geographic information system for each city in the south of France greater than 20,000. There is a geographic unit level data base incorporating the last two censuses. The Information system is intended to operate in real time and batch, and is in the experimental stage.
Raymond Van den Driessche O.R.S.T.O.M. Bondy	The ORSTOM (44) system is basically a data base of on-going research in soil science. The data base accepts lists in foreign languages and incorporates the following features: (i) management of the data base; (ii) updating of the descriptions within the data base; and (iii) Boolean selection of elements from the data base.
Alan Schmidt Laboratory for Computer Graphics and Spatial Analysis Harvard University	SYMAP - Synagraphic Mapping Program (26) - a grid manipulation system, can produce choropleth, contour or proximal maps. It is a fairly general program, using the line printer as an output device. GRID (26) is a system for producing line printer maps based on gridded data. It is fully implemented. SYMVU (26), an implemented but not yet released system, employs the line plotter to portray in perspective a 3-dimensional surface, which corresponds to the SYMAP or grid surface.

SPEAKER	ACTIVITIES/SYSTEMS DESCRIBED
Alan Schmidt (continued)	CALFORM (26) is an implemented system using the line plotter to produce conformant or choropleth maps. OBLIX (26), a program similar to SYMVU, incorporates several other features, such as the ability to present contour lines on the perspective view of the 3-dimensional surface.
Barry Wellar University of Kansas Lawrence	The Comprehensive Urban Information Systems is a consortium effort to develop an integrated information system. The effort consists of four stages: (i) analysis; (ii) conceptualization; (iii) design; and (iv) implementation, of which only the analysis stage has been completed.
Edward Schlosser Delaware County Planning Commission Media	The Delaware County Systems (16) are composed of: (a) The Newton Township pilot project, a grid manipulation system, incorporates one kilometer grid cell. It is based on land use information, including coding of land slope. Input is by manual methods. (b) The County system employs a standard grid, including point information and is presently under development.

SPEAKER	ACTIVITIES/SYSTEMS DESCRIBED
Michel Godron Centre d'Etudes Phytosociologique et Ecologique Montpellier	The Vegetation and Environmental systems are intended to find a precise relationship between vegetation and the environment. The following elements have been implemented: (i) a standardized data collection procedure; (ii) a way of estimating linkages between species and environments. The following elements are under development: (i) computation of the amount of information given by different aspects of the structure of vegetation; (ii) multivariate analysis.
Ronald Shelton Cornell University Ithaca	Remote Sensing and Environment Systems. The basic approach for this system is to obtain the map information by air photo or by field survey, to manually record the data, keypunch it, store it on discs. Generalized programs have been written for the retrieval of the information giving tabular or graphic output. The line printer is employed for graphics, using a package called PLANMAP, and there are approximately 25 computer programs in this implemented system. It is based on grid manipulation.
Gordon A. McKay Canadian Meteorological Services Toronto	The Canadian Meteorological Service system has 80 million records of climatological data on magnetic tape. They have started the geographic file to locate the information down to the degree and minute, and are presently doing grid mapping on a trial basis. The data base includes topographic data, and the pilot project is in the development stage.

SPEAKER	ACTIVITIES/SYSTEMS DESCRIBED
Clyde Johnson University of California Riverside	Vegetation Mapping encompasses the conversion of remotely sensed data to a geographic information system. There are three phases: (i) hardware development, presently under way; (ii) image data conversion, starting soon; and (iii) integration of the data and other information, not yet begun. The line mapping is from 70-mm or 35-mm imagery and the system will tie into SYMAP. It employs the UTM grids and is a test project.
David Bickmore Experimental Cartography Unit Royal College of Art London (note received after the symposium)	The Experimental Cartography Unit is a research institution with a variety of hardware and software capabilities. Projects have so far concentrated mainly on pilot production of existing maps for Ordnance survey, geological soil surveys, and oceanographic charts. On-going research lies in (1) Data capture (essentially by sampling surveys); (2) Computer Processing, especially interactive, editing and manipulation and data banking (data sets typically of 8 x 10^8 characters); (3) graphic output including generalization and development of techniques to evaluate effectiveness of comparative graphics.

Report of Discussion

Reported by: Warren E. Schmidt and Tom Waugh

INTRODUCTION

The rapporteurs sought to introduce some order to the thirty presentations of systems planned, under development or operational. To accomplish this some common approach was sought among the diverse efforts described. Capability was chosen for this purpose and it became the core of the arrangement below. This and the subsequent subdivisions and categorizations represent an attempt to provide a starting point for further work in this dynamic field. In no way should it be deemed comprehensive in its coverage, but represents only a categorization of those systems represented at the Symposium.

A SYSTEM FOR CLASSIFICATION

The main criterion for this scheme of classification is function. On this basis the systems were divided into Image Systems, Information Retrieval Systems, and Integrated Systems (Fig. 1).

I. Image Systems (Base Mapping)

These map image production and reproduction systems (not necessarily with cartographic accuracy), include those systems involved with topographic, hydrographic or aerial map production.

II. Information Retrieval Systems

This category can best be described as containing systems dedicated to providing information for a specific purpose.

III. Integrated Systems

This group of mixed image and information retrieval systems contains those with features of both the special Information and Base Mapping categories.

These groups were further subdivided and evaluated on the basis of their implementation, capabilities and sophistication. The degree of implementation was on a scale of six, beginning with potential and continuing through planned, experimental, feasible, developed and implemented. Capabilities were concerned with provision for

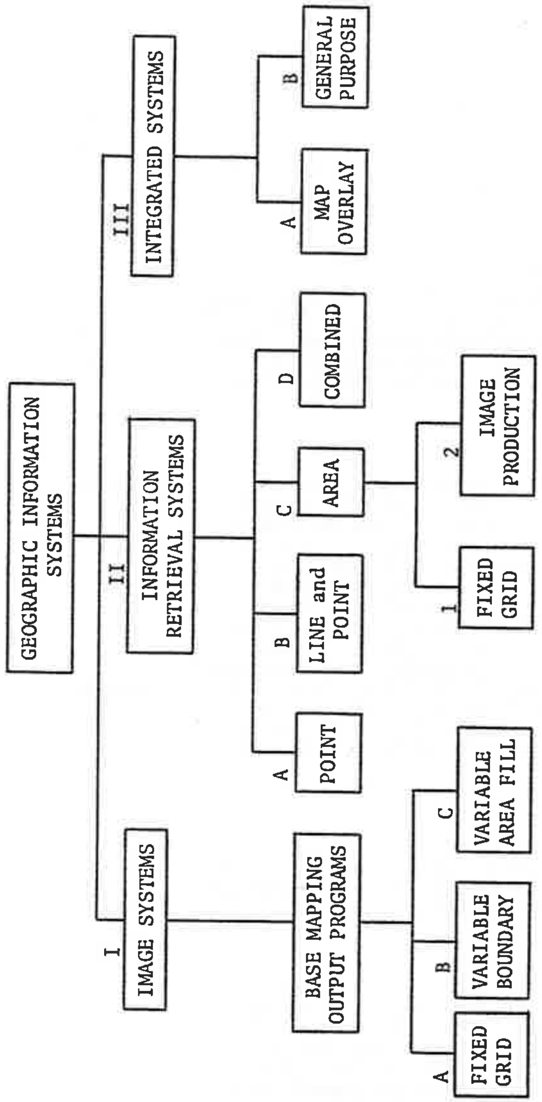


Fig. 1 - Section I

input, input storage, manipulation, manipulation storage and output. Sophistication considered degree of integration, compatibility, software efficiency, economy and other factors.

I. Image Systems

Much of the first day's discussion centered on output mapping programs. An information system as a minimum criterion stores and manipulates information while an output program has the capability of providing visual representation of this information. An integrated information system will do both of these and provide capabilities for input control.

Output Mapping Programs

I-A. Fixed Grid

The most commonly known and used of this group is undoubtedly SYMAP (26), developed by H.T. Fisher at the Laboratory for Computer Graphics and Spatial Analysis, Harvard University. This program provides great flexibility in graphic output on the line printer and is implemented and very well documented. A program of similar capabilities called GEOMAP (52) at the Department of Geography, University of Waterloo, is considered as developed. Documentation is being written. A.Schmidt D. Steiner

Also in the implemented stage is the GRID (26) program from the Department of Landscape Architecture, Harvard University, which will output maps from information in a grid matrix to a line printer. While GRID is only a mapping program, underlying its use is the grid concept which will be described later. A.Schmidt

Three plotter programs are also in the developed stage. CALFORM, developed by R. Cartwright of the State of Missouri, is a choropleth mapping program which can be used to provide more aesthetic maps than the line printer mapping programs. SYMVU (26), developed by F. Rens now at the Geography Department of the State University of New York Buffalo, produces three-dimensional views of surfaces generated by the SYMAP program. OBLIX (26), developed by A.L. Thomas, Harvard University, will, in addition, draw on the 3-D surface up to 75 contour levels. SYMVU and OBLIX have capabilities of drawing background symbolism and other information to give excellent visual presentations of quantitative data. A.Schmidt A.Schmidt A.Schmidt

I-B. Image Production Systems

Three systems were described - the Canadian Hydrographic System under development at the University of Saskatchewan in Saskatoon, the U.S. Army TOPOCOM System, and the U.S. Geological Survey System.

R. Boyle

The Canadian Hydrographic Service automation project (6) is being undertaken at the University of Saskatchewan in Saskatoon. Unlike the TOPOCOM system which, inpart, evolved, this project involves a comprehensive design to handle the entire mapping mission. The basic system consists of digitizing, manipulation and output components. The pencil follower is interactive with a PDP-8 computer and at a later stage may be used to instruct manually an automatic line follower. An experiment has been carried out to convert existing soundings on charts with an automatic character recognition unit. FORTRAN IV processing can be by IBM 360 or any large computer with output on a Gerber 32 flat bed plotter controlled by another PDP-8. A Barr and Stroud light spot projector is used on the plotter. The status of the software varies from being completed in the case of the plotter to being under development in the case of the interactive display. The first part of this highly advanced system will be delivered in 1971 to Ottawa for installation.

B. Aumen

The TOPOCOM's efforts are divided into pure information collection and automated storage and retrieval. There is much overlap in these activities. In the collection area, there is an interactive digitizing system with ten machines connected to two computers and a flat bed plotter. An automatic contour digitizer is now in the prototype stage. It employs an operator before a CRT display, plus a drum scanner. The UNAMACE system automatically scans and correlates stereo pairs and outputs the information in digital form. A semi-automatic type of placement system called SNAPS (65) and flat bed plotters with photoheads are also employed, as is a system for producing three-dimensional models from the digital terrain data. With the exception of the automatic contour digitizer, all of the systems described are operational.

D. Edson

The United States Geological Survey's efforts to provide national base mapping have been to develop both hardware and software systems that will merge new map information onto the 7-1/2 minute quadrangle bases and republish it in the shortest possible time. Hardware is under development in Washington that will scan the source documents, generate a three-dimensional terrain model and regenerate high resolution graphics. Supporting software is

also under development (41).

II. Information Retrieval Systems

These systems were considered to have storage and manipulation capabilities and most had some form of output capability, often using the output mapping programs described earlier. The systems were further subdivided into the type of geographic information that they used and the three forms are Point and Line, Area, and Combined.

II-A. and B. Point and Line

R. Tweedie

The New York State Department of Transportation uses two systems (38,57), one of which is point and one of which is line. They store information located geographically by the centroid of the area to which it refers (II-A). This system is inexpensive, but lacks the finer detail of the areal forms although it is excellent for polygon retrieval. The area systems described resort to centroids for many statistical manipulations while retaining area information for their base file. The same agency exhibits line storage, where highways are digitized and stored for subsequent manipulation (II-B). Both programs have a developed status.

J. Weldon

The GRDSR system (18)(II-A), implemented by the Dominion Bureau of Statistics in Canada, provides a street network file for geo-coding purposes. Information, in geo-coded form, is thus stored in a point mode and the system will retrieve information using a point-in-polygon approach.

II-C. Area

This deals with geographic information based on a land unit system, and this is the primary impetus towards collection. This grouping has been divided into three subdivisions - (1) Fixed Grid, (2) Variable Boundary, and (3) DIME/ACG - which reflect the type of input that the systems will accept.

II-C (1) Fixed Grid

This concept of data storage is widespread, due to the ease with which it may be implemented and the ease of statistical handling. All the data to be dealt with are applied to a grid matrix and the attributes of each cell are stored. It is not necessary to store a description of the geographic information as this is geometrically defined by the grid itself. Information may be coded manually or by machine. The ease with which grid data may be manipulated in complex statistical models makes the grid system a very efficient analysis tool. Most grid systems produce maps on the line printer as well as tabular information.

Earliest among the grid systems and the first attempt at an integrated mapping system was the MIADS/MIADS 2 (2) system. Typical of the later systems are the GRID "system" (not to be confused with the GRID "program") and the Cornell Land Use and Natural Resources (LUNR) Inventory System. Both systems have provisions for complex data manipulation for analysis purposes.

E. Amidon
W. Schmidt
R. Shelton

II-C (2) Variable Boundary

Two integrated information and mapping systems are under this heading, CMS (68) and GIMMS (69). GIMMS was written in cooperation with A.L. Thomas for Leyland Systems, Inc. CMS takes manually coded geographic data and produces map output on the line printer, while GIMMS takes digitized geographic information, edits it for correctness and processes it for output on a line printer. CMS has implemented status while GIMMS has developed/implemented status. Both systems take data related to the areas described (e.g., census information) store, selectively retrieve, manipulate and, if required, store modified information. Both systems also take the data and provide them in tabular and map form.

T. Waugh

T. Waugh

A system similar to CMS has been implemented at the Observatoire Economique Méditerranéen in Marseille, France.

J. Salmona

II-C (3) DIME/ACG

These two systems, which have the same general concepts, are the most widely implemented systems in urban data banks. While not being restricted to urban areas, it is in this context that they are most useful. When the base files are complete, they contain information about blocks, street segments and node points of urban network.

The ACG system was implemented by the City of Jerusalem, Israel (49), under the guidance of the Department of Geography at the University there. They system codes the block faces of the urban network. The system can be partly machine checked but, because blocks having no population are not included, then the complete urban network is not defined.

A. Shachar

The New Haven Census Uses Test (37) implemented a DIME file as a pre-test to the 1970 U.S. Census. The DIME system completely describes the urban network in terms of street segments. The description of the segment provides information to construct two independent networks which permit extensive machine checking on the information. The base file created by these systems is used for a variety of purposes, such as geo-coding information by address-matching techniques where, by comparing the given address with the base file, co-ordinates for the address can be approximated. Point-in-polygon retrieval of this information provides maps from formerly unmappable information.

While mapping from these files is possible, it is not easy and the City of Jerusalem has implemented a grid system linked to the ACG system which provides better facilities for analysis and mapping. Similar work is under development by Dominion Bureau of Statistics, Canada.

II-D. Combined

The Canadian Inland Waters Branch data processing system is structured into three sub-systems, two of which will be implemented this month. The system emphasizes user needs, flexibility, expandability and machine independence. Output is in the form of tables and maps that incorporate point, line and area data. Another combined system is FRIS (8) being undertaken by the Swedish Central Board for Real Estate Data. Now in the pilot stage, this ambitious project will eventually be extended to all Sweden. It not only locates land units, but also can record fixed objects or the path of moving objects within those units. Output will be in tabular or map form.

J. Gilliland

O.Salomonsson

III. Integrated Systems

This group, containing the most sophisticated systems, has many of the characteristics of the two previous classifications. Information is both stored and manipulated, as it is in the Information Systems category, and Image Mapping capability is provided. It is subdivided into map overlay and general purpose systems.

III-A. Map Overlay

The Canada Geographic Information System and the Map/Model System are examples of the map overlay category. Point, line, polygon and related land classification data are input, edited and converted for storage in a common locational form. Measurement and centroid location routines are available as is a comprehensive retrieval package that can both manipulate and generate alphanumeric or graphic output on demand. The Canada Geographic Information System (55) employs a drum scanner to input linear data, a pencil follower to record area identifiers and currently an IBM 360/85 computer; it outputs on a Gerber 632 flat bed plotter. Its input volume is large - 22,000 map sheets that include data ranging from the capability of the land for agriculture to recreation potential. Eight thousand of these sheets are now in preparation and 1,500 are entered. Hardware is operational and all the area boundary manipulations are implemented, with work proceeding on line, point and socio-economic data handling capabilities. J. Foster

Similar is the Map/Model (3) system based at the University of Oregon. This has been utilized to do urban and regional analysis, study resources and support educational planning. Like the Canada Geographic Information System, it can overlay maps internally and generates a composite graphic. It employs a digitizer, an IBM 360 series computer and a plotter. R. Keith

III-B. General Purpose

Two general purpose, integrated systems are AUTOMAP (48) at the U.S. Central Intelligence Agency and a package under development at the U.S. Army Corps of Engineers, Waterways Experiment Station at Vicksburg, Mississippi. The first-named became operational in 1967. It contains an input system that handles five projections and point or line data. The output program, CAM (58), produces any of 17 projections at any scale, and any location on earth. Grids, azimuths, range rings, circles, ellipses and many other features can be output on the plotter or CRT. Measurement routines, both linear and area, are provided as are cell, earth-profile and contour matrix programs. W. Schmidt The status of the U.S. Corps of Engineers, Waterways Experiment Station system ranges from experimental through implemented. Its core is a fixed storage format for line, point or area data that includes manipulation programs. Sophisticated hardware, including digitizers, plotters and CRT devices, is employed. W. Grabau

PARTIAL GLOSSARY OF ACRONYMS

ACG	Addressing Coding Guide
ARDS	Urban Data Center System, Seattle
AUTOMAP	Automatic Mapping System
CALFORM	Calcomp Conformant Mapping Program
CAM	Cartographic Automatic Mapping
CGIS	Canada Geographic Information System
CMS	Choropleth Mapping System
CRT	Cathode Ray Tube
DBS	Dominion Bureau of Statistics
DIME	Dual Independent Map Encoding
FRIS	Central Board for Real Estate, Sweden, system
GEOMAP	Geographic Mapping Program
GIMMS	Geographic Information - Manipulation and Mapping System
GRDSR	Geographically Referenced Data Storage and Retrieval System
GRID	Generalized Raster Information Display
LUNR	Land Use and Natural Resources
MIADS	Map Information, Assembly and Display System
NCC	National Capital Commission
OBLIX	Oblique viewing of a 3-dimensional object
OEM	Observatoire Economique Mediterraneen
ORSTOM	Office de la Recherche Scientifique et Technique Outre-Mer
SNAPS	Symbol and Name Placement System
SYMAP	Synagraphic Mapping Program
SYMVU	SYMAP Output viewed in 3-D
TOPOCOM	U.S. Army Topographic Command
UNAMACE	Universal Automatic Map Compilation Equipment
WWW	World-Wide Weather System

INDEX PAPER

Given by A.R. Boyle

It is a pleasure to introduce for discussion, automatic cartography, remote sensing, transmission methods and pattern recognition. Methods and techniques of automatic cartography are generally dependent on the specific application, and the processes themselves can best be examined from the viewpoint of their principle of operation.

Digitizing Source Data

Digitizing is required from aerial photographs, from existing drawn maps and from survey sheets. It may be automatic, semi-automatic, semi-manual or even manual. It may be required for the whole area or only for selected features.

Aerial photography digitization is usually used to produce selected outlines, contours or profiles. The contour extraction can be completely automatic, such as in experiments at Rome Air Development Centre, or semi-automatic, such as the process used by the Ontario Department of Highways. There the operator traces the contour with a standard photogrammetric reader to digitize directly, making any drawn map later from the digitized data. The output is generally a series of co-ordinates to represent the line data. Some automatic digitization units are clever enough to discriminate features such as roads and rivers from the background data; an example of pattern recognition. Lower-cost methods of data input from aerial photographs are always being sought, and the approach of the U.S. Geological Survey is of considerable interest. Recording the massive amount of data in a photograph is an essential problem of any method to be used.

Existing map or survey sheet digitization requires the digitization of lines, and alphanumeric and other symbols, but as they are in a definite form, the process of recognition is quite different from that in an aerial photograph. Two methods are available. The more common way is to use a manual tracing of the line, such as with the D-mac and Bendix equipment; the second is to replace the man with an automatic scanner.

With line tracing, the amount of operator participation is considerable, and if the amount of work is excessive the automatic scanner seems to be advantageous. However, the scanner has problems;

it requires a good quality of line without ambiguity and it takes an appreciable time to run at the desired precision - approximately 4 hours per map, independent of the number of lines. (Ed. note: 15 min. for CGIS). The scanner is almost incapable of annotating the information, and requires an additional manual operation prior or subsequent to the scanner process, when original maps drawn for normal visual purposes are being analysed. In most work, it is also necessary to use a computer program to change the scanned points of the line into sequential line form, and the cost of this can be appreciable. However, recent advances (J. Diello) in this area seem to be promising.

The optimum way may be to use a hybrid line tracing system using a complex of manual tracing units, automatic line followers, character recognition, and scanner units.

As far as automatic line followers are concerned a number are available on the market and experiments have been carried out by the U.S. Geological Survey, the Royal College of Art and the University of Saskatchewan, amongst others. To date, the general result is not quite accurate and reliable enough for cartographic work. There are two approaches to the methodology of automatic line followers. The first is to use an operator to work as a continuous controller of the 'automatic' follower, varying speed and making decisions at junctions and other ambiguities. The second is to trace-digitize the map manually and add annotations, but to reduce the workload by digitizing only at the beginning and end of lines, and junctions or ambiguities; the output is a director tape. The automatic line follower then takes this director tape and the map, and automatically fills in the missing line segments in a truly automatic process which is a mixture of following an input instruction from the director tape, and following the line itself.

The large variety of symbols on a map present a distinct problem for automatic character recognition. Experiments have provided some interesting results in certain circumstances (e.g. soundings), but the method is not easily implemented in others. It is necessary to remove any non-symbol data and this can take a draftsman several hours, a cost which might count against the use of character recognition in favour of completely manual digitizing. Names are an even bigger problem and it seems advisable to start from name lists rather than to digitize from maps. In our own experiments with character recognition, it appears that because we have well drawn digits on charts and only ten digits to recognize, we can do this simply, at fairly low cost, and on a small computer. An experimental prototype is now working. We would like to extend our work to survey sheets, but while we still only have soundings, these are hand drawn, and very closely packed. It may only be possible to resolve this digitization problem with human intervention.

Visits to many establishments and analyzing their speeds of digitization and the error rates involved, revealed a great similarity in the results produced. It is interesting to note, however, that some applications were surprisingly fast, some surprisingly slow. There is a need for manual digitizing: but there is also a need for semi-automatic and there is a real need for completely automatic digitization.

The use of specially prepared maps as a form of graphic data bank storage, and their digitization, are matters referred to later in this paper.

Direct Digital Input

There are many cases in which digitized information may be obtained directly for input. In our work with the Hydrographic Service, we will obtain digital information directly from the surveys, and many people are able to give digital information that can be fed directly into systems. On the other hand, a cartographic form of image is something which has already been generalized, and there are arguments as to whether it is right to digitize such an image for a data bank.

To recapitulate, we should now consider the advantages of digitization directly at source. There is digitization from aerial photographs; how far is this better done manually, semi-manually, or completely automatically? There is digitization from map information: what is worthwhile doing, what is not? How does the bulk of the data affect the way we do it? Do we want to mix manual and automatic? Do we want to persist in one way to the exclusion of the other?

At some time between the collection of data and its storage in the data bank the question of processing the data must be considered. Is it proper or desirable to do it, and if so, how and where? In many cases there is much useless data and this must certainly be removed for the sake of efficiency. Errors may have

occurred or may have been introduced by part of the input operation; these errors must be detected and either removed or corrected. When data is obtained from an intermediate source it will already have been processed and perhaps been generalized. Such data must now be further processed to assess the extent of its generalization and to annotate it accordingly.

Descriptions of data must normally be added by an operator. In order that the computer may carry out reasonable work on cartographic data, reliabilities and tolerances are also needed and such information is rarely available from source data.

There are very few cases where direct input of data to a data bank can occur, as the data would have to be purely additive, without overlapping existing information. A name data bank is one of the few examples where this might occur. It is however, reasonable to say that in most cases some processing is necessary.

PREPROCESSING INFORMATION vs DIRECT INPUT TO THE DATA BANK

Remote processes of selectivity and pattern recognition at source could be most valuable, and may be essential to meet the data transmission capability. Our interest now lies in assuming that further processing work must also be done at a later stage, but before the data is incorporated, and whether it should be done by a man or by a machine. It is only possible to point out a few arguments on either side, as each system will result in a different outcome. I am convinced that either extreme is inefficient, and it is a matter of finding the best compromise.

Man is particularly good at rapid visual selection of spatial data. His experience is important when it is difficult to break the problem down into simple programming rules. He is able to add an alphanumeric description to an object or line without difficulty. On the other hand, man's efficiency drops when he has to do routine tracing, routine checking or multiple simple conversion operations. In general, present day machines have the reverse properties. As a compromise, the machine should aid the man as much as possible by identifying areas on which the man should concentrate. It should check his work afterwards for logical inconsistencies. There is no doubt that the phrase "management by the exception principle" as used in commerce, is a very valid principle here. The programs should not be allowed to become excessively complicated when the matter could be more easily handled by a man. The only exception might be if the work load were too great for available skilled labour.

The economics of using a man as against a machine is much too diversified a problem to generalize. In time, however, it should be possible to assess the advantages in specific applications. Much depends on decision psychology and ergonomics. For example, how long can a man continue to make decisions at one sitting? What is his best work pattern? Is it one hour on and one off, or one day on and one off? We have a lot to learn both about the cost of a man and the cost of using a machine for complex work.

Man-machine interactive operation between a manual digitizer and a computer is one extremely important process and we have been developing such a system at the University of Saskatchewan. It is important in two ways: to get rid of useless data and to help man exercise supervision. A man must have confidence that the machine is doing a proper job and is properly recording information, and that the results look sensible and logical with respect to the procedures being carried out.

In post-digitization examination we also need man-machine visual interaction to cover the aspects where lines that should join, do not, or where pieces of information from exactly the same area and about the same item, conflict. What do we do about it? A new reference system may have been used. Is it really a better one and should we modify our whole data bank to fit? Or is it doubtful? Should we keep both old and new data and tag them with different weightings? With data already weighted for reliability, should we change its weighting in the light of new data? At present, man is the only source of sufficient expression and decision capability, and the work must be done through some form of interactive display. The computer can help by enhancing areas where man's attention must be concentrated. A composite of man-machine is perhaps the optimum present state of pattern recognition. The balance of how much work should be included in the computer program, as against a man doing it, is often quite erroneously assessed. There are existing computer programs that are very complex, when a man could have done the work easily. There are others where a man is working his heart out doing a thing which could easily have been done in a computer program. There are two methods which are very valid in this area and seem to work extremely well. The first is the type of display using variations of the Tectronix 611 or the alternative IBM 2250. The other system is the one at Rome Air Development Centre, where they are using the Concord digitizer plotter; an electro-mechanical method

of interactive manipulation. A man can play a most important role in any analysis system, but must be used wisely and without waste. It is no good giving him a display if he cannot see the area clearly, if it tires his eyes from flicker, or if it is not large enough. Moreover, he must be able to manipulate the image and change it at will (i.e. work in the interactive mode). This type of equipment is now available and some units are very good for cartography.

One of the great aspects of man-machine interactive displays is that it has at last put that method which is so frowned upon by scientists - trial and error - back into its proper place, and it is very useful in cartography. It is an aspect which has been lost for a long time in the application of automatic methods.

As a general principle in interactive editing operations, data should never be destroyed, but should merely be tagged to say that it has been deleted, either for general operation or for a specific one. Every human or machine sometimes makes mistakes, and we must be more than ever prepared for this in concentrated decision making. It is also useful to record (for documentation) a hard copy of the final decision each time one has been made.

Editing procedures of this type can take place at a number of stages in a system of cartography. These might be generalized as:

1. New data input
2. Incorporation of new data into a data bank
3. Examination of information of interest in a data bank
4. Commenting on the extraction of data prepared for a special purpose such as map drawing.

For geographic data, the problem is frequently one of the very large amount of data to be stored, which takes present digital methods to and beyond any reasonable limits. Digitally, it is possible to store either in line or scan form, although the reduction to the line state usually represents the minimum storage. A certain amount of reduction has always to be done and from an aerial photograph a considerable selectivity has to be exercised to control the bulk of the data. A map form of digital data is generally accepted as the best storage medium, as it involves definite lines and features and when stored is completely definite, although the accuracy cannot exceed the source from which it was extracted and the way in which it was drawn, and extracted. Data is normally stored as points, or as a series of points or vectors to represent a line, as in contour methods and chain encoding.

A number of papers have been written on these matters, including those by Boyell and Ruston (5), Freeman (24), Morse (40) and Diello (17). The other method of tabular representation is to record altitudes at selected lattice points on a grid covering the region of interest (4,31,32). Other publications in the field are by Carr and Lopec (7), Price (46), and Vitoshankin (67).

The next aspects we must consider are the compaction and compression of data. I call compaction "using the most efficient way of storing data", and compression "selecting out representative data only". The Canadian Hydrographic Service does not need to go to compression. Hydrography suffers from a sparcity of data, rather than too much. However, compression methods have to be devised for many other cartographic systems.

Curtis (10) describes digitized line information from the D-mac pencil follower. After referring to the removal of redundancies, he looks at some examples of curve fitting. Straight lines he finds too regular, and with circles there are discontinuities. Adding smaller circles obviously requires more information to be stored. He suggests that discontinuities up to 5° might be acceptable, although this is a decision in the realm of a perception psychologist. Some benefits that I can see of using circles are that (a) one can compact data for the information bank, and (b) one can expand information for subsequent drawing by calculable means.

One of the most important papers is that by Boehm (4) of Rand Corporation. He presents a statistical method which interpolates, between ground points, sequences of pseudo-random numbers with stated probable distributions. Though not very good in appearance, it does meet statistical requirements based on distribution and reliability functions.

Some of us are fortunate that our data banks are not too large to be left as they are, but many others have a very difficult problem which would involve tens of thousands of reels of magnetic tape. After we have made compaction as efficient as possible, how can we compress the data still further? Various compression methods have been proposed. Most methods do not at the same time consider the bulk of the data bank and the cost of

retrieval and incorporation. Regeneration of a good quality cartographic line, its relationship to, and accuracy compared with the original after compression are doubtful matters, but such processes seem to be necessary. Personally I do not discern much hope among the proposed methods of compression and would prefer to examine processes such as digital storage on photographic film, reputed to be many thousands of times as compact as magnetic tape.

The Analogue Data Bank

We now come to the question of whether it is better for the data bank to be a drawn map instead of a digital one. I do not mean any drawn map, or a map drawn for visual use, but one specially drawn by a machine for another machine to read. Could all necessary manipulations be done in scan form? This is the approach of the Rome Air Development Centre (17).

This aspect of the use of a scanner is somewhat different from the use in the Canada Geographic Information System, where manuscript maps are scanned and digitized. In Mr. Diello's application the map will be both drawn and read in machines, with the only aim the production of compact, high accuracy, data storage. I am impressed with his work. Of course there are problems, but the method does have a great deal to offer.

Pattern Recognition

We can consider two aspects of the use of pattern recognition. The first is at the data collection source, and would be used to control the examination area and perhaps the vehicle itself, or at least to arrange to reject the recording and transmission of an image which is judged to be useless. When the data transmission channel is limited, this type of work has enormous possibilities (as well as the chance of considerable failures). The work is closely associated with artificial intelligence work being carried out at Stanford and M.I.T.

The second aspect of pattern recognition is a more leisurely and critical one, being applied at the base to which all data has been sent. A reconstructed image may be used in this case, but in these days of digital data transmission, it is probably best carried out on the raw data. In all this work, we are particularly concerned with multitudinous types of imagery: radar, infra-red, photographic and others. Dr. Rosenfeld will talk later today about texture analysis of images, textures being most important characteristics to segment the pattern first of all, and then the ability of pattern recognition to classify properties.

In the analysis of original maps and certain specially drawn maps, there is every likelihood that automatic character recognition would be useful and occasionally relatively easy to carry out. At the University of Saskatchewan we have found this possible in examination of depth soundings, and it could also be relatively simple when different areas have to be numbered. The methods of character recognition are many, depending upon the number of possible characters and the quality of the drawing.

REPORT OF DISCUSSION

Reported by: A. Schmidt, D. Sinton, C. Steinitz

The discussion reflected the diverse experience of the participants. The unifying experience of the participants identified by the rapporteurs was that all were dealing with space in a two-or-more-dimensional sense and with measurable phenomena within that space. Further, they were speaking of the application of computer technology to the description and/or prediction of the behaviour of these special phenomena. The following discussion is organized in a manner which allows identification of the common issues which must be dealt with before improved geographical information systems can be achieved.

In describing the actual discussions and comments which occurred on the second day, the comments are organized to reflect the major topics which were covered, and are presented in the same sequence in which they were offered.

Some Concerns in Digitizing

Nomenclature is identified as an initial important problem, particularly with reference to the words "edit", "man/machine", "digitize", "geographical data system", "compaction" and "compression", and discussion focused on the search for a set of agreed-upon definitions. R. Boyle

The index paper has identified several methods of digitizing. The first is the direct entry of digital data as part of the data collection process, such as happens during hydrographical surveys. Another, identified as pseudo-digitizing, takes place when images are scanned for transmission by telemetry systems. The third was that of extracting information from aerial photographs or plans and maps. This generates a key problem related to the system demands caused by the use of original versus intermediate data sources. The original data sources include, for example, aerial photographs and surveyor field sheets. Intermediate sources are seen as the broad range of existing drawn maps such as contour maps, which are generalizations or reproductions of the original source materials. In a hydrographic project, contour maps are used to generalize hills: "Mariners are not expected to sail their ships up mountains, but they like to know the general shape of them". However, underwater terrain is recorded from the original depth sounding records, since a greater level of accuracy is required for these data. When dealing with terrain information where concern with accuracy is of low priority, the intermediate source is used. A variety of approaches to the digitization of R. Boyle maps were identified, all of which are fully covered in the report of the U.S.G.S. conference on digitizing (66).

Two approaches of "line following" or "line tracing" as opposed to "scanning of an entire sheet" to identify patterns of lines were then discussed. While scanning is advantageous when large amounts of information have to be digitized, it presents problems of pattern recognition that are not inherent in line tracing. Another problem that exists in much of digitization work is that of character recognition.

The following are several diagrams concerning the system being developed by the U.S. Geological Survey for digitizing their maps, D. Edson for the purpose of creating terrain models.

U.S.G.S. Map Digitization

Digitization procedures being developed by the U.S. Geological Survey were introduced; their basic task is to digitize thousands D. Edson of maps at reasonable cost. Special scanning equipment designed especially for the purpose is used to digitize each of the colour-separated plates which make up the entire map. Special problems related to the digitization of two of the colour-separated plates were described. A half kilometer square is magnified and projected on the face of a photo-emissive tube which scans an area two centimeters square. The scanning element is driven in a raster scan of 1,024 lines, each line being sampled 1,204 times, so that the system generates a sampling matrix of over one million points. The photo-sensitive element delivers a binary response simply stating that there is a line present or absent in that element location. A series of addresses at which the line exists is recorded, and a line-following routine detects the string of contiguous data points and identifies the centre of gravity of the line. Although the contours are scanned automatically, it is necessary to use a semi-automatic digitizing method to input labelling tags for contour heights. They may be performed off-line. The automatic scan and the labelling tags are merged to generate a terrain model at a grid spacing of ten meters, selected as a desirable spacing to permit large scale mapping. This does not imply that the data are necessarily valid or even warrant the establishment of a terrain model at a ten-meter spacing. It simply means that the system can operate at this level if it is felt that

valid data and uses exist (Fig. 1-6).

The second colour-separated plate which generated problems is that of the cultural features. In this case, several cultural features are presented on the same original map. It has been found necessary to use a photo-mechanical separation technique to identify separate features such as buildings, railways, boundaries or road patterns. Once these features have been separated, they can be automatically scanned as a separate overlay. The introduction of these techniques has led to a re-evaluation in the U.S.G.S. as to their symbolism key for a D. Edson variety of cultural features and map preparation procedures; the long range aim is the development of machine-readable symbols for use in mapping.

Pencil Following or Line Digitizing

The CGIS and TOPOCOM systems also use scanning methods of J. Foster digitization, and the case for the pencil follower method of D. Edson digitization was presented. The system is simple, and allows D. Bickmore for a single operation that does not require the redrawing of the original map. Of particular importance is the possibility of being able to work with original field sheets. In many cases, the data as collected in the field result in manuscript map sheets which would be in no way suitable for the precision required by automatic scanning devices, particularly when such sheets contain the blood, sweat and tears of the field data collector. For a small operation there seems to be little reason to use anything other than a manual digitizer. It was W. Schmidt noted that the original experience of the U.S.G.S. in attempting D. Edson to develop a tailor-made line-seeking device has been set aside, for the time being, in favour of scanning, and that experience indicates that manual digitizing is more efficient than line-seeking digitizing. D. Bickmore

Interactive Input Checking or Editing

The needs of the CGIS in terms of the volume of documents J. Foster to be digitized does indeed justify the use of a scanner, along with appropriate software. Gaps in the availability of software for automatic problem solving are of particular concern. Typically, three major types of problems may arise with scanning devices: (1) closely packed lines, such as contour lines, at a precipice, which cause confusion in the system algorithms; (2) merging of data from adjacent map sheets; and (3) systematic

U.S. GEOLOGICAL SURVEY — SCAN DIGITIZING TECHNIQUE
DEAN EDSON
10-70

REFORMATTING OF GRAPHIC DATA

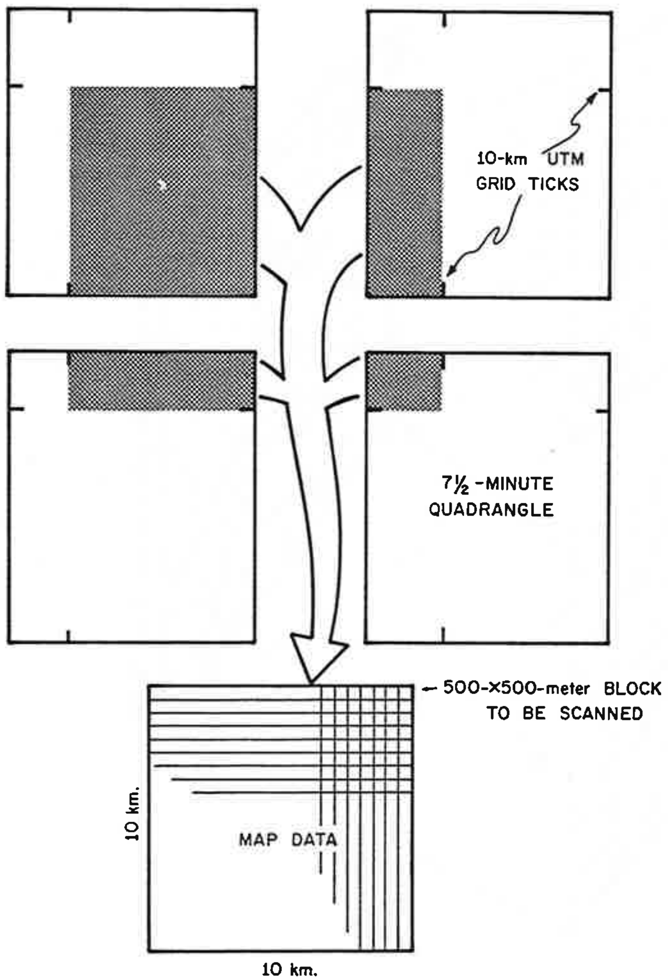


FIG. 1, SECTION II

SCANNING THE GRAPHIC DATA

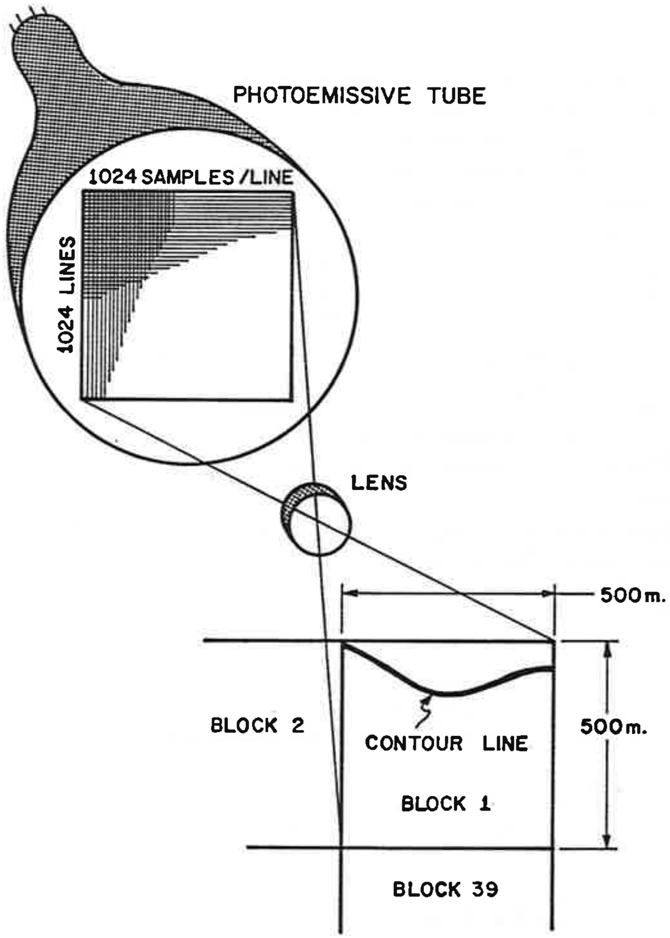


FIG. 2, SECTION II

LINE PROCESSING

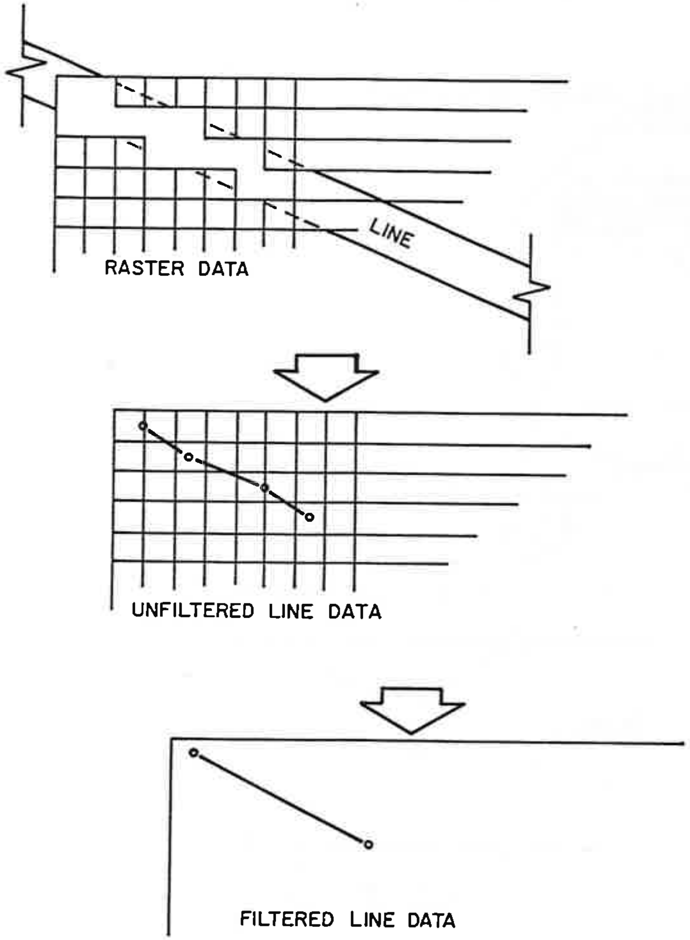


FIG. 3, SECTION II

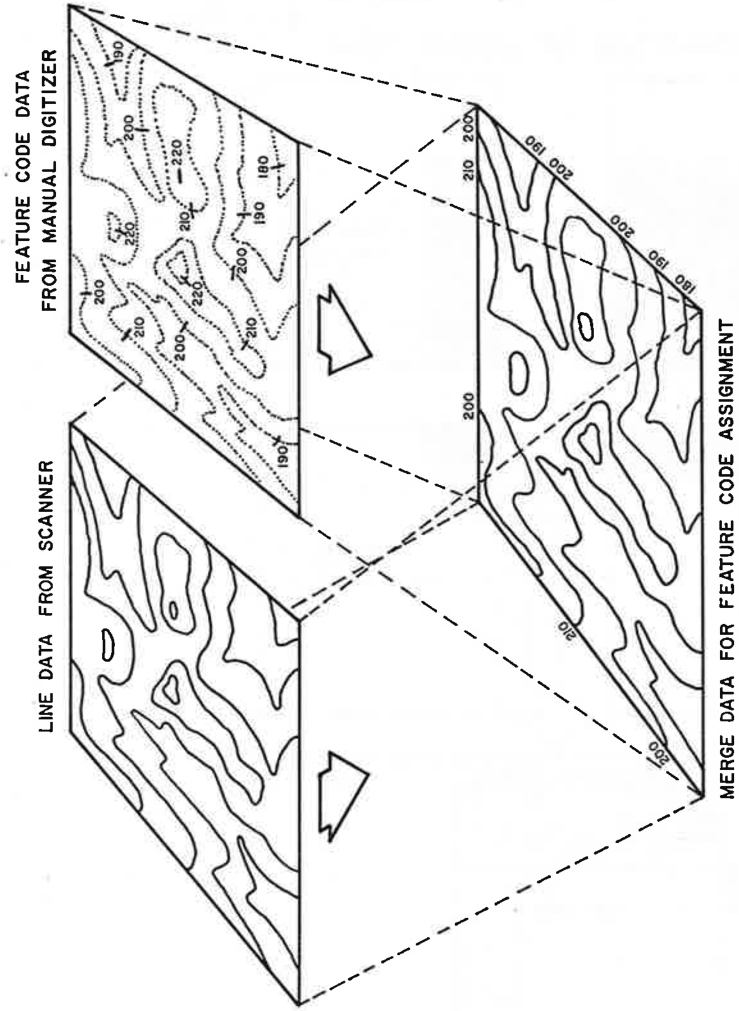
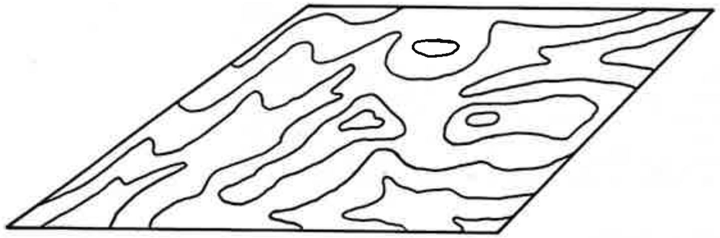
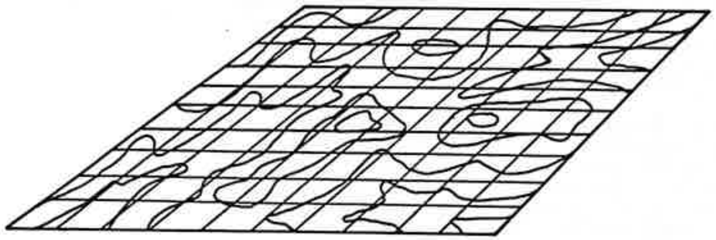


FIG. 4 – SECTION II

TERRAIN MODEL GENERATION



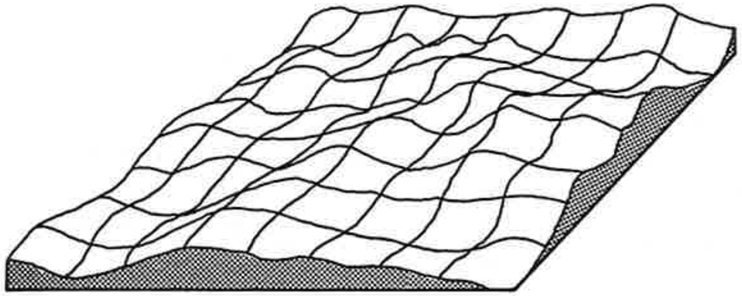
CONTOUR DATA IN DIGITAL FORM



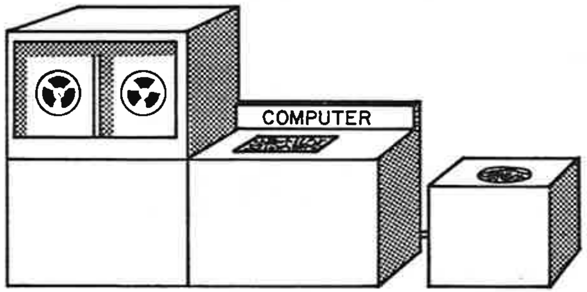
10-METER GRID OVER CONTOUR DATA

COMPUTE MOST PROBABLE ELEVATION FOR EACH GRID INTERSECTION.

FIG. 5-SECTION II



TERRAIN MODEL IN DIGITAL FORM



STORE IN DIGITAL FORM X,Y,Z VALUE FOR EACH INTERSECTION

FIG. 6-SECTION II.

errors which exist within the original source data. The need was expressed for greater man/machine interaction, rather than placing a greater importance upon the need for automatic problem-solving software. V. LaGarde

A system has been in prototype operation for over two years which involves a man/machine interaction in conjunction with an automatic contour scanning process. The contour lines as scanned are displayed on a video device, which is monitored by an operator who has a joy-stick for use to correct lines being produced and displayed by the computer. In addition, the operator is able to alter the lines as they are defined. Over two thousand maps have been digitized by this process at TOPOCOM. W. Aumen

The use of the man/machine interactive environment is also beneficial to the operator in that it allows him to maintain a degree of confidence in the work as it progresses. R. Boyle On the question of automatic problem-solving software, software to anticipate and solve all possible problems does not yet exist and the currently available software should be considered solely as a tool for a person and not a slave for a machine. D. Edson

General Comments on Editing

It is important to perform two separate types of checking, for positional accuracy and for feature coding and other editing. D. Bickmore While the second can be checked in a variety of ways, for example, CRT, etc., the first can be checked only by overlaying an exact graphic playback over the original map. There are very real dangers of perpetuating errors through the over-precise digitization of already generalized source data, and the need is apparent to use reliability codings.

It appears that the ultimate definition of "error" is in most cases referred back to a senior draftsman for determination, and in almost every case the editing process is essentially one of trial and error.

Smaller Scale Systems

The need for an editing system based on redundancy checking of a polygon's segments is apparent, and for the small user a S. Arms number of considerations are important. These include such things as needing to use whatever equipment may be available from a T. Waugh service bureau, rather than being able to purchase a specific piece of equipment which may be more attractive, such as those

capable of on-line editing and reciprocal display of the information as it is digitized. For the purpose of matching adjacent map sheets, the needs for accuracy of the small user could perhaps be sufficiently well served by digitizing edge points common to each of two adjacent map sheets, and transferring locations from the first to the second. (Note: Small systems often use only a few sheets.) T. Waugh

Some Issues in Terrain Modelling

Regarding the basic philosophy of digitizing, and especially that of digitizing contour data, doubts were expressed that the ultimate purpose of the digitizing process is mainly for storing functions, because, if so, it would be much cheaper and easier to make a microfilm of these functions. P. Yoeli If the user wishes to produce them in any scale, he can then blow them up again. Rather, the purpose of digitizing should be to achieve a digital terrain model. The question is not primarily how to digitize contours, but what to digitize in order to achieve the digital terrain model which will answer all those questions of the relief presentations which we need.

Contours are the result of the manual work of cartographers, and therefore the cartographer's personal interpretation of the relief surface. P. Yoeli The user has no idea of the accuracy of the contour system itself, of the flow of the contours, or whether they are generalized, or whether they are sufficiently detailed, not to mention the simplification of the relief which is the result of the choice of a specific vertical interval by the cartographer. As a result, to digitize the contours of such a map is basically wrong. The primary question should be: "If terrain models of the earth's surface are desired, at what density are right points needed and which of the sources yield the most accurate and suitable information to achieve this specific model, and are they available?" P. Yoeli The approach recommended is to digitize air photos, numerically, at a sampling density which has been decided or calculated as suitable for terrain models approximating the true earth's surface, which otherwise cannot be defined mathematically.

The complex issues involved in dealing with climatological systems support this approach. G. McKay Among the complexities are inequalities in the spatial distribution of data sources, locally inconsistent data sources, the need to deal with data in both an archival and a predictive sense, and a need to incorporate both short and long term time dimensions in the input, storage and analysis of data.

The Man in the Field

It was noted that the man in the field is only the first part of a system that should end with the user, and it was confirmed that a data system not tailored to a user is useless. The requirements of a user cascade down through the system to the data collector in the field, highlighting the importance of correlating field data collection procedures with the remainder of the total system. In particular, devices such as soil compaction field probes produce records directly in digital form on a magnetic tape and are immediately compatible with input digital-form methods, with minimal prior manipulation. Reference was also made to a field data collection device used to record hydrologic time series data via analogue to digital conversion techniques.

R. Tomlinson
W. Grabau
J. Gilliland

Data Generalization

On the subject of data generalization, the potential use of sampling techniques in data acquisition for obtaining accuracy of definition, and the need for a measure of generalization within any given data base, were discussed. Procedures for data generalization should be reversible although this requirement is not always met in practice. It is important to acquire and maintain data at its finest scale and then generalize to the level of aggregation appropriate to any specific application. Discussion of the necessity of a trade-off between data accuracy and the cost of its acquisition follows, and mention was made of a technique for factor complex map nesting which involves the compilation of multiple areal data files into a single data file. The final "map" can also be envisioned as a multi-valued variable as a function of co-ordinate position. The related question of the need to investigate the relationship between spatial scale and classification scale was also raised.

P. Yoeli
W. Tobler
V. LaGarde
C. Steinitz

Data Storage Efficiency

The distinction between compaction and compression of data was outlined. A compaction seeks the most efficient way of organizing the data to be stored, while compression involves the actual elimination of information. Compression should be undertaken only when conditions permit the regeneration of an adequate approximation to the original information, or "a good looking line".

R. Boyle

Compression as a special-purpose generalization is one approach to the solution of massive volumes of data. The approach instituted by the Rome Air Development Centre, dependent on the analogue storage of information in a map form, was described. The information is drawn on a specially prepared film base for complete digital retrieval through a scanning device.

R. Boyle

A mathematical procedure which involved the fitting of cubic splines to random line shapes (e.g., coastlines) was discussed and dismissed as too expensive in computer processing time to be useful. The "tolerancing" procedures in use for the compaction of line data through the use of adjustable "tram lines" was also described. A response of the Canadian Geographic Information System to the large data volume problem identifies the point and vector method as a compaction system for efficient description of line data.

D. Bickmore
D. Bickmore
P. Kingston

Highlighting this point, and acknowledging the intuitively compelling argument in favour of compaction, it was pointed out that "one must be careful to realize that compactness is not the ultimate goal. Usability is the ultimate goal and for certain types of applications the saving that is achieved by using an ingenious compaction scheme may be lost due to increases in processing time." Distinction was noted between compacting lines which are intended for recovering as lines, and compacting information for recovery in an unknown way for an unknown use.

A. Rosenfeld
W. Aumen

The desirability of "flagging" to retain recognizable caricatured forms and the suggestion that "saliency" was related to perceptual problems were noted, and it was observed that very little research had been done in this area. It was suggested, however, that this concept of saliency seemed antagonistic to the tendency in data analysis to apply mathematical smoothing techniques.

D. Bickmore
J. Foster

Methods to solve some problems of archival data were described, and it was demonstrated that the identification of a direct relationship between two sets of data items can permit the complete regeneration of the item that is compressed on the basis of an item that is retained.

G. McKay

Sampling Data

It was suggested that for certain types of data the problem of compaction should start right at the point of collection,

D. Bickmore

through sampling techniques. The experimental work by the Soil
Survey Group in England was cited, in which soil data were being
collected on a one hundred meter square grid; this represents
the generalization of data at its source. It was also noted
that the problem in determining sample density was the definition W. Grabau
of homogeneous population, and some of the difficulties in
indentifying appropriate sample densities for soil mapping were
described. Alternatives to square matrices were considered, and J. Gilliland
the advantages and disadvantages of each, such as triangles and A. Rosenfeld
hexagons, were described. The square grid appeared to be the
best available compromise.

Pattern Recognition

During the last hour, discussion involved issues related to pattern recognition, identifying the differences associated with recognizing gray tones as contrasted to identifying lines. Comments were based on the following schema: A. Rosenfeld

Gray Scale Image	Map and Line Drawing
Noise limiting digitization data rate (20,000 per/sec. for 6% gray level).	Much faster rates possible.
Film Storage and rescan highly desirable for compactness.	Possible to compact digitized data; non-repeatability of the digitizing is the problem.
Integer array vs. bit.	Chain encoding?
Parallel processing (not necessarily optical).	Segmental Processing (tracking, etc.)
Rescaling by sampling or blurring.	Rescaling must preserve correctivity and smoothness.
Cartooning (segmented, edge prediction, etc.) a central problem.	Input or already in a cartoon form.

INTRODUCTION TO INDEX PAPER - P. Kingston

The following is a summary of the process of data manipulation and retrieval which will generally indicate the order of the topics covered in this section of the symposium proceedings:

A. Information Storage

1. File Structures
2. Classification (Characteristics)
3. Complex Data Classification
4. Density Problems
5. Dimensional Storage (0, 1, 2, 3, 4)
6. Linkages between Data Sets
7. Relevance
8. Reliability - manipulation effect

B. Information Manipulation

1. Centroid Allocation
2. Classification Change
3. Contouring Automatically (Trend Surface Analysis)
4. Dissolving Selective Data Lines (Boundaries)
5. Distortion Elimination (Linear, Non-linear)
6. Generalization
7. Generating Lines, Bands, Polygons, Circle
8. Intervisibility (3-D)
9. Measurement - count, area, length
10. Merging - Non-overlapping Data
11. Overlay - Overlapping Data
12. Projection Change
13. Scale Change
14. Search - Given Profile
    - Sub-set Creation
15. Statistical Routines and Applications
16. Simulation Models

C. System Control Facilities

1. Monitor Systems
2. Manipulation and Query Languages

D. Information Retrieval Facilities

1. Display Techniques
2. Listing Data of Given Profile
3. Man/Machine Interaction for correction, updating and query.

In addition to the processes of data manipulation and retrieval given above, it is also useful to think of three types of map storage systems which can be generally described as follows:-

1. Geographical Indexing Systems, which may also be called Simple Grid Manipulation Systems, or Point Systems.

In these systems, all information is stored with reference to specific point co-ordinates to represent an area grouping. These points may be the centroid, block face or similar arbitrary location. This category also includes all "discrete fixed grid" or "variable grid" systems. The two options within this classification are the fixed grid areas (e.g., square kilometer, square mile or square unit every 1/10 of a degree), or variable size grid areas (block face or enumeration area centroid).

2. Map Compilation Systems, or Map Reproduction Systems. These are generally "continuous line" systems.

These systems store image data as lines for map reproduction, display and retrieval purposes. Few, if any, manipulation facilities are provided except for contouring, scale and projection changes. Maps may be stored, updated and merged, but the logical connections of the lines to form area boundaries with classifications are not maintained. Hence, area calculations, logical overlays with area segmentation, classification manipulation and position locating are not possible.

3. Graphic Data Handling Systems, or Image Manipulation Systems.

These are continuous image systems with a sophisticated design philosophy. However, they tend to be the least developed in many regards. These systems maintain point, line and area data with their logical connections, inter-relationships, positions and data classifications. As a result, the full range of manipulation facilities including reclassification, automatic generalization, selection, measurement, logical overlay and point-line-area interaction and search are all possible. Due to the extensive volume and complexity of information held, these systems usually maintain

multi-file data banks with complex file structures and linkages. Many maps covering different areas, and different information about the areas, must be stored and subject to comparison and analysis.

The above definitions will be used during the remainder of the discussion on data use and manipulation.

INDEX PAPER

Given by: P. Kingston

PART A. INFORMATION STORAGE

A1. File Structures

Definition

The file structures of a system should be defined by the intended data use. Many structuring options must be considered, including such factors as:

1. single or multi-file organizations
2. sequential, direct-access or structured data sets
3. sequencing by co-ordinate, classification, size, arbitrary reference number, or a relative location positioning
4. continuous image, generalization by variable grid reference point or centroid co-ordinate, or arbitrary grid generalization.

Each of these four factors has a significant effect on the potential data use.

1. Multi-file organization affects the variety and volume of information that may be maintained, its flexibility and its accessibility for rapid or easy manipulation. These advantages of a multi-file organization are partially offset by its complexity and the difficulties of cross-referencing.

2. The storage mode of the data sets is also closely tied to speed and flexibility of manipulation and data use. Sequential formats are simple, usually less expensive and require less sophisticated design. Direct access files are readily accessible for inquiry, allowing fast processing of multiple inquiries to a large data base, but are more costly to maintain. Structured data files commonly are in the form of sequential sub-files within a direct access grouping. Structured files may be illustrated, for example, by sequential storage of land classifications grouped by province or state, with each provincial group directly accessible by reference to an index of storage locations for province or state name. Another example would be detailed census data grouped by census tract within storage location calculated on the basis of tract number. Census storage by data file strings, as in the DBS Geo-coding system, provides a unique file organization for a large data volume. Such structured files are normally the most sophisticated storage technique with flexibility and reasonable access speed, but frequently they either have inefficient use of storage or are difficult to update.

3. Sequencing is commonly carried out in a strictly sequential manner on the basis of some data characteristic, as indicated above. This is often dictated logically on the basis of intended data use (e.g., land potential for agriculture by area size within quality classification); however, it may result in severe limitations in the use of the system's facilities. More general sequences, such as by co-ordinate (UTM or longitude-latitude), are helpful but often require lengthy searches, as they are linear or strip-based storage techniques. One of the more sophisticated sequences is the area-based "Morton Matrix" organization (Fig. 1) which provides a minimum sequential search (on a statistical basis) between any two adjacent areas in any direction. Similar extensions of this concept to three or n-dimensional storage are evident. A few other similar concepts have been developed but not published or implemented.

4. The most frequent filing system is generalized storage by an arbitrary grid. This groups all data for areas of one kilometer square, or one minute square of longitude by latitude, and assigns to it either the most significant or average characteristic (e.g., soil type or crop being grown). Some systems allow multiple classifications with percentages attached. Often the area generalization is tied to a specific grid co-ordinate. The next level of development uses fixed shapes (normally square), but of varying sizes. Some systems allow any size or shape and maintain the information by centroid (or a reference co-ordinate) and area size. This makes it very difficult to relate areas to one another or to produce image output. However, it tends to give more accurate totals for land inventory. The most powerful technique is that of a continuous image. This allows both accurate statistics and image displays. However, the requirement of maintaining all area boundaries, to a fine scale, with their relationships to one another, produces a large data volume and a complex organization.

State of the Art

There is a "file structure triangle" (Fig. 2), a convenient

Y																
	85	87	93	95	117	119	125	127	213	215	221	223	245	247	253	255
	84	86	92	94	116	118	124	126	212	214	220	222	244	246	252	254
	81	83	89	91	113	115	121	123	209	211	217	219	241	243	249	251
	80	82	88	90	112	114	120	122	208	210	216	218	240	242	248	250
	69	71	77	79	101	103	109	111	197	199	205	207	229	231	237	239
	68	70	76	78	100	102	108	110	196	198	204	206	228	230	236	238
	65	67	73	75	97	99	105	107	193	195	201	203	225	227	233	235
	64	66	72	74	96	98	104	106	192	194	200	202	224	226	232	234
111	21	23	29	31	53	55	61	63	149	151	157	159	181	183	189	191
110	20	22	28	30	52	54	60	62	148	150	156	158	180	182	188	190
↑ 101	17	19	25	27	49	51	57	59	145	147	153	155	177	179	185	187
Y 100	16	18	24	26	48	50	56	58	144	146	152	154	176	178	184	186
3 011	5 0101	7 0111	13 1101	15 1111	37	39	45	47	133	135	141	143	165	167	173	175
2 010	4 0100	6 0110	12 1100	14 1110	36	38	44	46	132	134	140	142	164	166	172	174
1 001	1 0001	3 0011	9 1001	11 1011	33	35	41	43	129	131	137	139	161	163	169	171
0 000	0 0000	2 0010	8 1000	10 1010	32	34	40	42	128	130	136	138	160	162	168	170
	000	001	010	011	100	101	110	111								
	0	1	2	3	X →											

Fig. 1 (Section III) "Morton Matrix"

THE FIRST 256 UNIT FRAMES

FILE STRUCTURE TRIANGLE

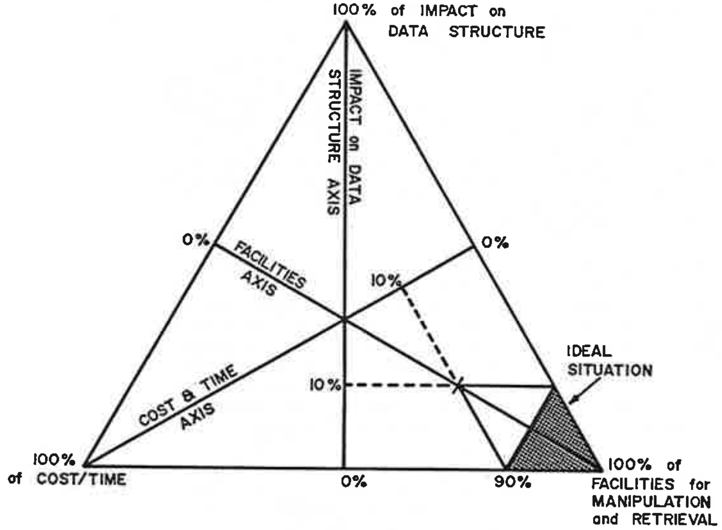


FIG. 2, SECTION III

method developed by R.F. Tomlinson and the author for ease of reference in discussions on the subject, in which we have a three-dimensional graphic system expressed in a planar triangle. There are three axes - a cost/time axis, a facilities axis, and an impact on the data structure, with 0% located on the sides and the 100% end of the axis located at the corners. The costly situation, for example, is anything that is on the lower left hand corner. The severe impact on data manipulation and retrieval are at the bottom right corner. We are ideally trying to pick out a situation that has as low as possible a cost, very little impact on the data structure in order to maintain flexibility, and yet maintain all the facilities that are required.

The object of a system is to optimize data content with regard to information to be gained from it. There must be trade-off between the facilities for manipulation and retrieval, the impact on the data structure, and the cost and time of development. With any information management system the ideal situation would be 100% of the facilities, with 0% impact on the data structures, at no cost for development. Most systems sacrifice many of these facilities, resulting in unsatisfactory compromises and very restrictive design. As a result the systems must be largely or totally rewritten to add new features. However, a slightly greater initial investment in design and organization will result in a much better system. Far more work can be and must be done in the field of systems design of file structures before available and recently announced equipment will be effectively utilized, and before it is possible fully and efficiently to implement truly sophisticated systems with many manipulation facilities. In the discussion on these topics participants will be asked:

Why were your file structures chosen?
What compromises and costs were involved?
What limitations have been imposed on the system?
Were they worth making?
What results from it?
Do you believe more compromises should have been made for an earlier, although limited and possibly closed-ended, implementation?

REPORT OF DISCUSSION

Reported by: A. Schmidt, D. Sinton, C. Steinitz

A. INFORMATION STORAGE

A1. File Structures

The term "facilities" as used in the index paper was defined by the author of the index paper as "the capabilities of being able to provide manipulations of a data structure". P. Kingston

The first presentation described the "Morton Matrix" which has been referred to and illustrated in this report. The Morton Matrix is a file structure for organizing large numbers of records describing geographic data. The particular problem for which it is designed is that of storage of such two-dimensional medium. Regardless of whether the data are stored on a tape or disc, it is possible to gain access to only a single record at any given point in time. The issue here is that unless the next item of data required is physically "near", time is taken to move much tape or to seek it on a disc. In both cases, the times are proportional to the separation which exists in the file between successive data items being sought. The problem concerning the sequence in which such records should be filed arises from the fact that a user frequently will wish to manipulate the file in order to extract information for a series of land areas which are adjacent to one another in the real world. This should require as little file manipulation as possible, and should not disturb the file pointers which Mr. Kingston referred to earlier. The purpose, then, of the Morton Matrix file structure is to minimize in a physical sense the distance within the file between any two data elements which refer to adjacent locations in the real world. As can be seen from the diagram which illustrates the Morton Matrix structure, the individual elements are arranged in order to satisfy this requirement. It should also be noted, however, that the extent to which this requirement is met is not uniform but is satisfied better around the origin, and in many other local points, and less so in the center of the matrix. An additional benefit of such a file structure is that the actual record address of a desired geographic location can be directly computed by manipulating the binary representation of the geographic X, Y co-ordinates as shown on the Morton Matrix diagram. The address is established by interleaving the binary representation of the X and Y co-ordinates in such a way as to combine G. Morton the individual digits of each axis in sequence from a high order to a low order.

The Dominion Bureau of Statistics, Canada, string concept of file organization was developed in order to provide efficient access to the Dominion's census data, which originally were recorded on 100 reels of magnetic tape. The original data file consisted of 22 million records, each of which had 120 fields of data. There were two types of capabilities desired for the eventual system. These were: (1) election of records by subject matter; and (2) retrieval of records by geographical area. It was recognized that a random access capability would be required and that it would be extremely important to minimize the number of accesses involved when retrieving data from the file. J. Weldon

If we think of the original 22 million records, each having 120 data fields, as being conceptually a large matrix of 22 million by 120 elements in size, then the proposed solution to the problem can be thought of as essentially a redefinition of the total file so as to have not 22 million records of 120 data items each, but rather 120 records, one for each of the data subjects, each record now being 22 million data items in length. A number of efficiencies result from such a file structure. When responding to a query concerning one or more data items, it is not necessary to search records containing 120 different data items. However, it is likely that many of the 22 million questionnaires will need to be tabulated. With this file structure, one need only read the collection of fields required for one's tabulation. However, additional savings have also been realized by taking advantage of the fact that many of the data items can be stored in less than a full word, i.e., in one or more bits. For example, classifications of sex can be stored as a single binary bit and educational characteristics can be stored as two bits. The over-all effect of this file reorganization and change in coding structure has been a 60% reduction in the total amount of storage space required, the essential content of the original 100 reels of data being compacted onto 22 discs. Also important is the fact that the data can be retrieved more easily since far fewer individual records have to be approached, there now being only 120 conceptual records, rather than 22 million physical records.

In order to achieve a geographic referencing capability, a separate directory is maintained. Each record in the directory or dictionary contains an X, Y co-ordinate and a pointer to the sequential location of a group of records in the large data matrix.

The X, Y co-ordinates describe the block face centroid locations at which the data were originally collected. It is important to note that no geographic references such as X, Y co-ordinates, are contained within the data base itself. In order to retrieve data for a specific geographic area, one first extracts from the dictionary pointers to all those groups of records which satisfy the definition of the area for which information is to be retrieved, i.e., all X,Y groups that fall within the interest window or polygon. The imputed aggregates are then reported for the areas in question.

The question was raised as to whether or not the directory is regarded by DBS as being analogous to a data element dictionary, which is usually comprised of files, records and data elements. It was resolved that the directory is a subset of a dictionary, and that the documentation of the DBS data string concept of file organization is in its early stages. B. Wellar J. Weldon

Historical data such as that which might be obtained from each of the subsequent censuses would be dealt with under this file structure and could be stored essentially as a third dimension in the existing two-dimensional file structure. J. Weldon

The DBS data structure appears to be parallel to the principle of bit plane storage developed at the University of Maryland, except that the DBS structure is concerned with data records, while the University of Maryland storage is concerned with data words. A. Rosenfeld

The GIMMS system (Geographical Information Manipulation and Mapping System) input consists of line segments which are prefaced by the labels of the areas which the line segment separates. This file is useful for editing, both for boundary checking to ensure the integrity of the map sheet, and to change and modify the information contained in the file. This file contains the information needed to connect the areas on either side of a given segment. The input file structure, however, is not useful for area manipulations. As a result, it is necessary to modify the file structure prior to performing area manipulations on the data. This is done by creating a file where the information for each zone is linked together so that connectivity around a particular zone is achieved. The new file format also enables special cases such as islands, "doughnuts" and completely enclosed areas to be identified. In addition, the system provides for multiple labelling of areas so that one line segment may be the boundary between two parishes, and may also be the boundary between two counties. This allows county boundaries, or parish boundaries, or combinations thereof to be retrieved for mapping purposes. It was not felt that there was a need for inputting redundant information for checking the topological integrity of a file. However, the need was recognized to input redundant information so that more accurate information could be computed.

It was pointed out that the MIADS input form consists of drawing around the complete boundary of each area, rather than just inputting the line segments. This is easier for relatively unskilled digitizer operators. E. Amidon

On the question of storage of labels as separate from the storage of the X,Y co-ordinates themselves, the GIMMS system labels are arbitrary reference labels tied to the areas concerned. The X,Y co-ordinates in the information relating to the geographic areas are maintained separately and are merged in the plotter mapping program which provides the final graphic output. T. Waugh

A taxonomy for classifying different types of data was suggested which had been developed during the study of very large volumes of data already collected and recorded on maps or in volumes of statistics. The taxonomy is based on three attributes of a data item: its location, its value and the time at which it was collected. The three-dimensional nature of this definition of data can be related to the three-dimensional nature of many of the simple systems of computer storage, that is, data file, data record and the location within the data record. The parallel dimensions of computer storage technique and the attribute of a data item should also be considered when developing structures for the storage of digital data. D. Sinton

Many of the decisions arrived at in the design of a file structure are, however, determined by hardware necessity. Described in terms of a general problem, one of the greatest difficulties is the attempt to overlay multiple map sets to create composite maps. In taking the case of 30 maps, each of which is divided into three zones, there exists the mathematical possibility of 3^{30} combinations. Fortunately, nature intervenes and prevents many of these combinations from occurring. However, even if there are only 3^{20} combinations, efficient file structures are still an obvious necessity. This problem as stated was referred to in later portions of the day's discussion, particularly in the discussion on map overlaying. The question of identifying such enormous numbers of combined variables was not clearly resolved. V. LaGarde

Request frequencies are not uniformly distributed. The possibility of structuring a file to reflect the probability of the need to retrieve any particular data item was discussed; the Morton Matrix, it was noted, is such a file organization. As is true with almost all such files, however, it is a static structure, and a heuristic file structure would be self-reorganizing on the basis of experience in its actual use. R. Boyle W. Tobler

From the viewpoint of cartographers whose primary interest is that of graphic output, two classes of maps can be identified. These are: topographic or physical maps in which the goal is to represent a piece of the earth's surface; and thematic maps, in which thematic information is imposed on the base information of the map. In terms of data storage the first group can be characterized as permanent information. It includes (1) a geodetic base, (2) topographic information, (3) name content, (4) hydrologic information, and (5) planimetric information. The second major grouping of information was characterized as statistical data. The role of a map editor was described as one in which a selection is made from the various groups of information mentioned above and a map is the final product. The expectation is not to produce the whole map from output hardware, but to be able to create the basic data files which would allow one to select and produce maps from a set of mixed technical procedures. P. Yoeli P. Yoeli

SECTION III

DATA USE AND MANIPULATION

INDEX PAPER

PART A. INFORMATION STORAGE

A2. Classification (Characteristics)

The data classification is the description of the area represented by the data record.

Classification characteristics are supplied portions of the actual data content of data set records. The only exception is in the remote-sensing scanners that identify the area classification through the use of an input scanning device. Techniques of inputting the data through character, pattern or colour recognition were covered elsewhere in the Symposium.

Our concern is in the actual data representation that is maintained. A single specific data classification (land use, soil type, forestry type, or climate class, even with an extensive coding structure, causes many limitations on the type and accuracy of data maintained. As it is impossible to classify every point on the earth, a compromise is required. This is recognized in most systems, allowing multiple classifications (or sub-classifications) with percentages attributed to each to be assigned to a stored area. It seems reasonable in most situations to limit this to a maximum number of characteristic factors.

A flexible classification scheme and the ability to join together classifications from different input source maps is absolutely essential in order to overlay the combined source maps and to answer the questions concerning the interaction, or to provide a map comparison facility. In these situations two separate, possibly unrelated maps are combined and each land area is multi-classified (Fig. 3). An example would be a combination of soil class with rainfall in a study of agricultural potential.

These concepts have already been developed in detail and are partially implemented in some systems.

A3. Complex Data Classification

The capability to maintain a complex set of classifications for stored data is a frequent requirement. Census data may have as many as 1800 inter-related variables that it is desirable to store and manipulate on various bases. For planning purposes it

PRESENT LAND USE

Classification items A and B represent the characteristic of the land involved.

NATURAL RESOURCES

Classification items C, D and E represent a different set of characteristics of the same land.

COMBINED COVERAGE

In the combined coverage, classifications are combined and the area segmentation that can occur is shown.

This is desirable in order to answer the combination of questions as, in this example, "Where is the ore deposit under a bed of water, and is it difficult to reach it?

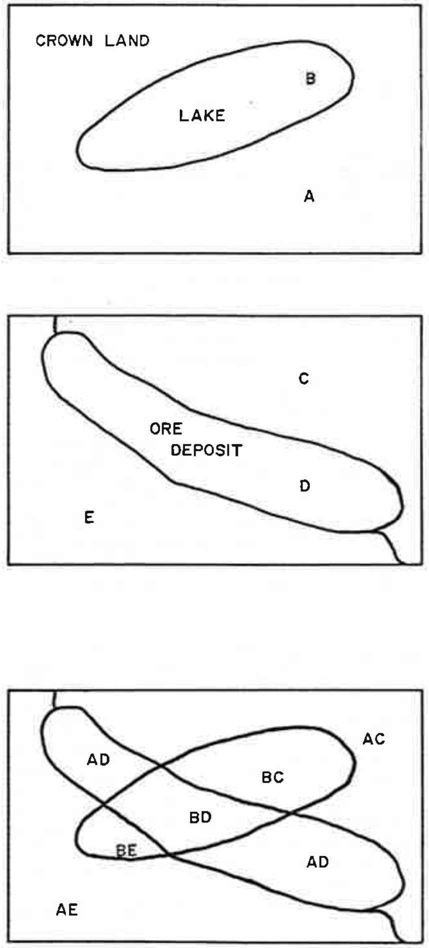


FIG. 3, SECTION III

is important to tie this information to land potential and use, and thus geographic relationships must be associated with it. This can be done within one system, with a file structure based on a common grid or continuous image. However, it is only possible with a variable data structure for classification information. This includes having such facilities as: a variable number of classification characteristics; a variable length of classification information; optional information types (numeric, binary, alpha-numeric characters); variable structures or arrays with sub-classes; and the ability to reference the variables by field name, which is extremely important in retrieval operations.

A4. Density Problems

Density problems are associated with the quantity of information to be handled, especially in high-population areas, or areas for which there is a need to maintain detailed information. In fixed grid systems, this does not cause great difficulty as there is no facility for holding more information about one unit than about another. In a continuous image system, it compounds the problem already in existence of maintaining the volume of data (especially for large land areas such as all of Canada) in one file.

Some facilities exist for handling this situation. The simplest is that of generalization, or assigning a multiple classification with percentages to grouped areas; this is a form of avoiding the problem in a continuous image system, but it is the only solution for fixed grid systems. Techniques exist for reducing the volume of information by as much as 50 to 1, through compaction of line or boundary information to a vectorized notation. These routines are used through necessity for continuous image systems. No better system has yet been developed for situations where the area units are very small and these compaction routines are not as effective. These facilities are being used more to fill the gap than to solve the problem. Some of them are quite efficient, but for very small and very segmented areas even those storage techniques provide limited data compaction.

The state of the art for the required facilities is, at best, experimental.

REPORT OF DISCUSSION

PART A. INFORMATION STORAGE

A2. Classification (Characteristics)

A3. Complex Data Classification

A4. Density Problems

J. Sharp

A major problem in digital mapping concerns the transfer of line-map data from a stereo photogrammetric instrument onto a compilation sheet for pre-editing. Transforming the data to computer storage formats precisely and economically for retrievable storage, display, change and plotting on separation negatives needs to be done with a numerically specified quality. Such specification of a mapping system can be best defined as one which performs operations on a series of documents used or created in the process. The content of map-system documents at present can be estimated by the magnitude of its information in bits. This magnitude implies the resolution, positional precision and therefore quality of the documents. The major documents referenced are tabulated with data content defined below:

Document	Maximum (neat)	Bits (magnitude)	Form of Data (digitized)
Roll of Film (aerial)	9" x 9000"	10^{12}	Facsimile
Photograph	9" x 9"	10^{9}	Facsimile
Control Sheet (geodetic)	48" x 48"	10^{5}	Vectors & Data
Line Sheet	48" x 48"	10^{8}	Facsimile
Name Sheet	48" x 48"	10^{5}	Vectors & Data
Display Section (12" x 12")	4" x 4"	10^{4}	Vectors & Data
Separation Negative	48" x 48"	10^{8}	Facsimile
10-Colour Map	48" x 48"	10^{9}	Facsimile

INDEX PAPER

PART A. INFORMATION STORAGE

A5. Dimensional Storage

Zero dimensional storage is point data in which there are two co-ordinates to represent the point, but the point itself is zero dimensional as regards definition. There are many facilities available to use this type of data.

One-dimensional storage is line data such as highways or railroads that may be plotted, but cannot be manipulated in terms of their relationships to one another.

Two-dimensional storage refers to area boundaries with their bounding and neighbouring relationships, to permit area calculation and manipulation.

Two-dimensional extended storage, which most people call three-dimensional, is area boundaries plus altitude information maintained as "warped plane". This increases the data volume very slightly over two-dimensional storage, thereby permitting contour mapping. Frequently, storage is maintained as zero dimensional (points) which are on request submitted to trend surface analysis. Systems with this facility have been developed for plotting purposes only, without other manipulation facilities, although particular systems permit digital terrain models. We have heard about some of these systems in the past two days, and they tend to be in the developed stage, with reference to their facilities.

Three-dimensional storage maintains air current and weather conditions, land with altitude, sea currents, ocean bottoms and geological information.

Four-dimensional storage includes the time factor, or historical information, about the areas that are concerned.

However, the last two produce far too high a storage volume for any existing systems, and the computers and programming logic required to handle such a mass of complex image data have yet to be created.

REPORT OF DISCUSSION

PART A. INFORMATION STORAGE

A5. Dimensional Storage

The question was asked whether, contrary to the statements in the index paper, one could indeed manipulate line data in one-dimensional storage. The author acknowledged that one could change scale and projection in this system. He also clarified the nomenclature question of what constituted two-dimensional storage, and included network analysis within this grouping. D. Marble

INDEX PAPER

PART A. INFORMATION STORAGE

A6. Linkages between Data Sets

Data set linkages are required for multi-file organization schemes. The example here (Fig. 4) is a simple linkage, for a separation of image and descriptor data. Arbitrary sequence numbers are assigned to portions of the image, to be used in the descriptor portion of the data set as a cross-referencing number. It is usually required to have some way of re-approaching the image, and a more sophisticated technique, such as the one used in the Canada Land Inventory based on the Morton Matrix, provides a linkage, given the location of the area or the centroid, back to the image portion of the total image file in which that data item would be contained.

There are few multi-file information systems, and as a result data set linkages tend to be peculiar to their individual systems. Very little work has been done on a general basis to develop an effective or optimal theoretical base for linkages. Those that exist are more the result of the particular system than an enhancement or base for its further expansion.

SIMPLE DATA SET LINKAGE

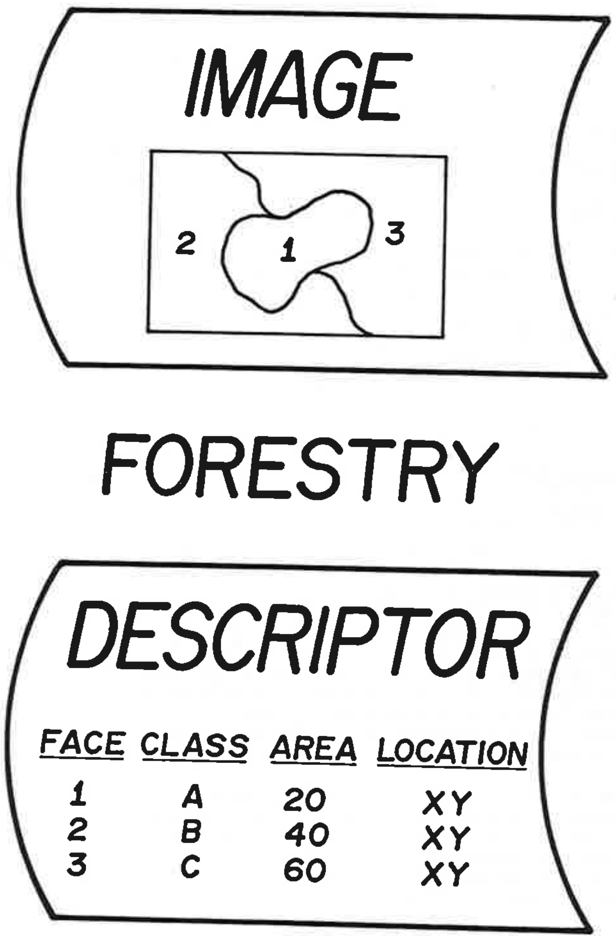


FIG. 4, SECTION III

SECTION III DATA USE AND MANIPULATION

REPORT OF DISCUSSION

PART A. INFORMATION STORAGE

A6. Linkages between Data Sets

A set of linked data files was described that is being used to reduce part of the tremendous data storage problems implicit in a study now being undertaken. This study concerns the design of vehicles to operate in an environment similar to that of Northern Alaska. Once the test study area is selected, a linked data file system then identifies the correct quadrangle to be studied and within that quadrangle, identifies the UTM grid cell. The numerical identifier assigned to that grid cell links the system to a library file which contains information on surface configuration, vegetation characteristics, and other significant data characteristics. The library file must be continuously updated as new overlaid maps are prepared, because each new overlaid map generates a new area number which will be referencing a new set of composite characteristics in the library file. V. La Garde

During slight semantic confusion, the very important point was raised that linkage to library files produces efficient storage only when the amount of information to be stored in the library is greater than the amount of space required to store the address characteristics. R. Shaw

The RESCAN system developed by two students at the University of Michigan in 1968 was mentioned. This system has handled some of the problems of large scale data storage by storing simple "on and off" map characteristics as binary bit patterns. These on-off bit characteristics are linked to a library file system that identifies the characteristics of the "on" state. E. Amidon

INDEX PAPER

PART A. INFORMATION STORAGE

A7. Relevance

One of the greatest problems facing automated data base information systems is that of data relevance. This is very closely tied to the question of intended data use. The problem is similar to the storage of library abstracts before it is known what all the specific interest profiles of the users will be. It is necessary to make the system as general and flexible as possible so that the information content may be modified, as user needs mature and become clarified, without loss of the system's capabilities. It is also necessary to determine, as clearly as possible, before entering data into a system, or even designing the manipulation facilities, what the data use will be.

When a system is operational there is a need to determine if the information contained is relevant to the changing needs of the user, and to the subjects being studied. Do any systems maintain processing or inquiry profiles? Ten-year-old census data may not be relevant by themselves, but ten-year trends may be. How do we determine if land descriptions from surveys 50 or 100 years old are still relevant? Do users indicate known inaccuracies, and is this used as a feedback technique indicating frequency of occurrence, to determine data relevance? What is being done to tie in relevance of data to their actual uses?

A thirty percent accuracy rate on a library abstract retrieval facility is considered successful. Much higher demands are being placed on geographical information systems, but less is being done to determine the actual conditions. There is an immediate need to purge systems of useless data, determine automatic purge techniques, maintain data relevance indicators and lay out a discipline of data relevance techniques. The state of the art is in the infancy of experimental development.

REPORT OF DISCUSSION

PART A. INFORMATION STORAGE

A7. Relevance

The chairman described relevance as being based on the temporal decay rate of data in the system. Many reasons were given why one should not throw data away. The British Agricultural Department has thrown away data and has a current inability to develop certain trend analyses. There are also legal constraints to throwing away data, which center on the fact that many of the data in the public domain are owned by the government. The question of who owns and who throws away is in this case, perhaps, dominant, but the issue of when to throw something away is, in other cases, of at least equal importance. D. Marble R. Boyle J. Coppock G. McKay

Temporal relevance might best be appreciated by considering three relationships: (1) between demand for and availability of data; (2) between data on file and the status of phenomena that they represent; and (3) among elements of data sets, that is, the time span during which data are collected for the respective sets. B. Wellar

The chairman redirected his definition of relevance away from the discarding of outdated information to merely reducing the level of accessibility of a given set of information, perhaps by changing the file structure within which it is kept. In experiments involving 150 years of daily hydrographic data, in which data compaction has been attempted on the basis of trend curves, the criticism that this would lose important exchange values was acknowledged, but it was noted that long-term recurrent values were the intent of the experimenters. Discussion of the issue of relevancy was redirected toward identifying criteria for the selection of data to be input and data to be retrieved. This important aspect of accuracy was not developed in the discussion. D. Marble D. Sinton R. Shaw D. Sinton K. Dueker

INDEX PAPER

PART A. INFORMATION STORAGE

A8. Reliability

Data reliability is a subject that is often neglected as much as data relevance, yet it is a subject that is of similar concern and importance to the user. Information reliability is affected by input, storage, and manipulation techniques, in addition to the original survey sources. It is difficult, if not impossible, to do any more than estimate the reliability of the original input that is provided. However, a measure of the accuracy maintained or lost as a result of digitization of the continuous image from scale and projection changes, from manipulation and retrieval facilities or from generalization (smoothing, grouping and regrouping), is mathematically calculable; the effect of various manipulation techniques could be compared automatically. This would provide an indication of the relative accuracy of the information (such as a plus or minus indicator to go with areas, e.g. 10 square miles plus or minus 0.2), which the inquirer could use as a reliability indicator. If this is possible, why do so few systems provide the facility? The cost of producing the computer program code and the storage requirements for maintaining the information have always been traded for considerations of time or for other facilities. However, when the systems are implemented the first questions asked concern the reliability that may be placed in the information produced. Is it through neglect that these facilities have only reached an experimental state, or is there a misunderstanding of the relative value of the requirements?

REPORT OF DISCUSSION

PART A. INFORMATION STORAGE

A8. Reliability

The discussion which occurred on the issues of reliability tended to be organised in four groups. They were: (1) general issues; (2) input reliability; (3) mechanical reliability; and (4) output reliability.

The study of the inherent reliability of data is a neglected field. A procedure had been presented some time ago for identifying with four levels of accuracy, the reliability of spatial data which appear in map form. Unfortunately, the procedures proposed were never officially adopted. J. Coppock D. Bickmore

On the issue of reliability of input, there is a distinction between precision and accuracy, precision referring to space in the surveying sense, and accuracy emphasizing the observed data. The measurement of accuracy, it was proposed, is the measurement of distortion, and eight orders of accuracy are identifiable. J. Sharp

Repeatability is an additional concept of data reliability, in particular with reference to changes in data which may occur due to hardware malfunctions or distortions due to time. Output reliability or accuracy can perhaps be improved by displaying data in a fuzzy form, when this is appropriate, as a measure of their actual variance. The precision steps in the gray scale would be identified as being related to orders of fuzziness. However, "fuzzying" data would not improve them, and one always has to make the best possible sharp discrimination within data. R. Shaw A. Rosenfeld

INDEX PAPER

PART B. INFORMATION MANIPULATION

B1. Centroid Allocation

A reference co-ordinate is maintained in most systems. It is usually the data identifier in fixed grid or point data systems (e.g., the central point of a one kilometer square unit, or a block face number). In a continuous image system it is used to give a representative co-ordinate, or as a labelling point for plotting purposes. The difficulty in using a centroid is that it may be external to the area which defines it. (Figs. 5 and 6).

The reference co-ordinate may be assigned on the original map input and be simply maintained by the system. The other option is to calculate and assign automatically an internal point or centroid. The calculation of the true centroid is fairly straightforward and readily available in manipulation systems. The calculation of a central point that must fall within the defined area may be found quite easily by counting boundary crossings. However, to ensure that the point falls within the largest note of a stretched-out, complex shape is much more time consuming a calculation and usually is discarded in favour of an approximation. No efficient algorithm has been developed beyond the experimental stage for internal "best-point" assignment. Mathematical centroids and other calculations have implemented status.

REPORT OF DISCUSSION

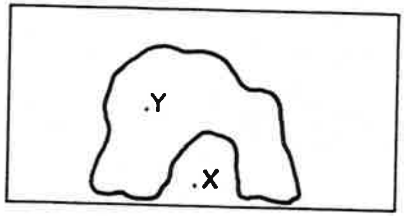
PART B. INFORMATION MANIPULATION

B1. Centroid Allocation

References to publications concerned with centroid allocation in general or point-in-polygon procedures in particular included several reports available as a part of the Lund series on geography published by the University of Sweden.

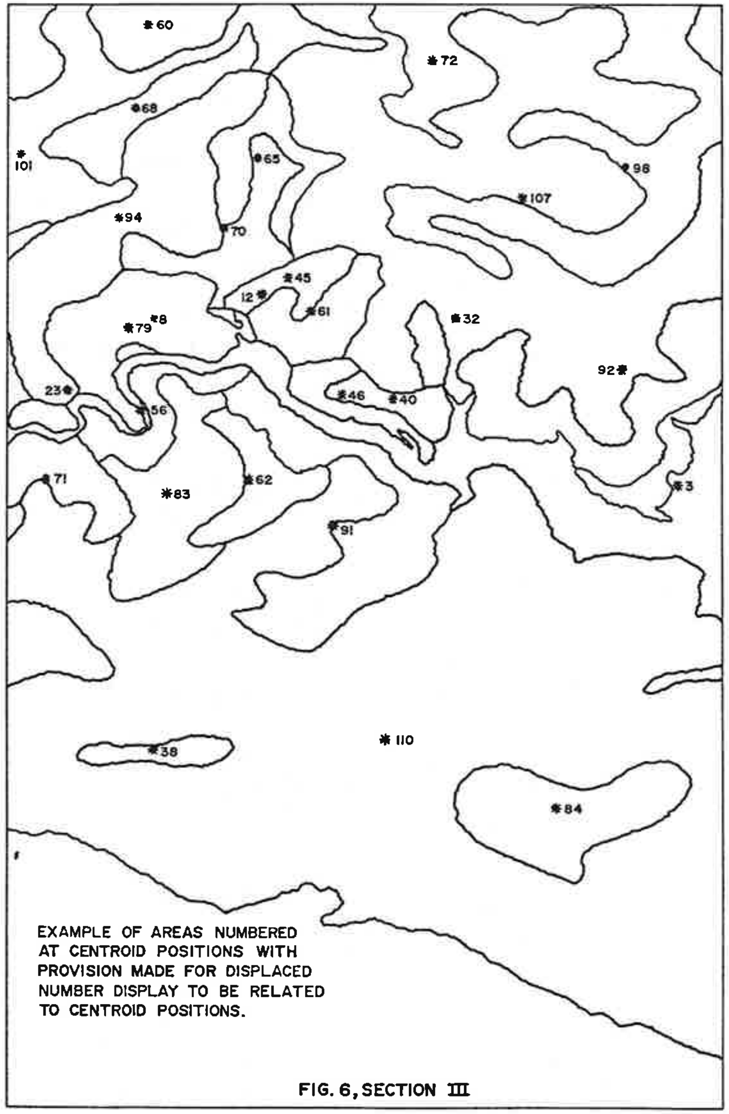
In response to the problem of a computed centroid which would fall outside of its zone, such as in the case where we have a crescent-shaped zone, one technique which can be used is that of A. Schmidt

REFERENCE CO-ORDINATE–CENTROID RELATIONSHIP



X– centroid
Y–reference co-ordinate

FIG. 5, SECTION III



EXAMPLE OF AREAS NUMBERED AT CENTROID POSITIONS WITH PROVISION MADE FOR DISPLACED NUMBER DISPLAY TO BE RELATED TO CENTROID POSITIONS.

FIG. 6, SECTION III

locating a centroid so that the point is within the boundaries of the polygon and at a maximum distance from its perimeter. The time for calculation of such a point is reasonable for a simply connected region. However, for a continuous line system, it may become extensive. A calculation that takes more than one-tenth of a second per area becomes excessive when dealing with tens of thousands of areas with multiple boundaries (e.g., doughnuts). A technique for placing items into a simply connected region is:

Watkins

Method: Construct a 1-1 mapping of the region on the unit square.

Applications:
(i) Pseudo-centroid ↔ (1/2. 1/2);
(ii) Place items at given density, uniformly or at random ↔ similar placement in the unit square;
(iii) Name placement ↔ name in unit square.

The Mapping: Assume a "sweeping" of the region by "shortest" chords, say n in total. Then (U,V) ↔ point on the ith chord in ratio V:1 - V where i = nearest integer to U_n.

Sweeping Algorithm: From any start point, parametrize two curves. (Fig. 7a).
1st chord via: step once from Start on each curve.
Knowing chord n, there are three candidates for the (n + 1) chord: advance only on curve 1; or advance only on curve 2; or advance on both curves. Choose that chord of shortest length (or rather length-squared, to avoid taking a square root). There is a mild difficulty in resolving a tie.

Detecting trouble of backtracing: Detect sweep turning back on itself (Fig. 7b) via:
Compute signed area A of the quadrilateral (or triangle)(Fig. 7c) determined by successive chords by the equation:

$$2A = x_1y_1 + x_2y_2' + x_2'y'_1 + x_1'y_1 - x_2y_1 - x_2'y_2 - x_1'y_2' - x_1y_1'$$

and detect change of sign.
Mark the trouble point and continue. If we finish without turning back again, we are in case I (Fig. 7d).

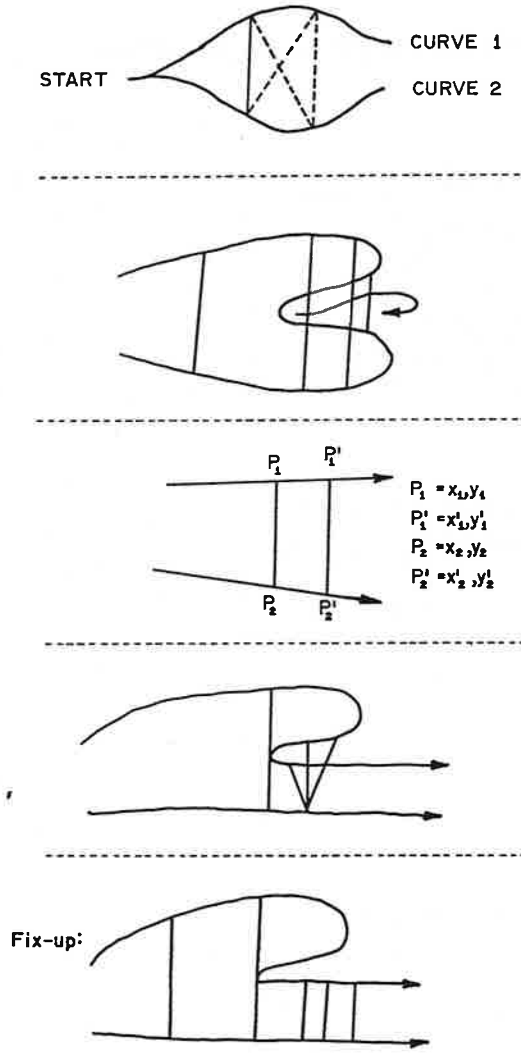


FIG. 7a

FIG. 7b

FIG. 7c

FIG. 7d CASE I

FIG. 7e CASE II

FIG. 7, SECTION III

```
      SUBROUTINE PSNTRD (IX, IY)
C,.   RETURNS PSEUDO-CENTROID
      INTEGER x 2 N1, B1 (5000), N2, B2 (5000)
      COMMON/WNGLN/N1,B1,N2,B2
      INTEGER x 2 M, I TAB (5000), J TAB (5000)
      COMMON/PTRTBL/M, I TAB, J TAB
      M = 5000
      CALL CHORDS
      MH = M/2
      I = I TAB (MH)
      J = J TAB (MH)
      IX = (B1(I) + B2(J))/2
      IY = (B1(IH) + B2(JH))/2
      RETURN
      END

      SUBROUTINE CHORDS
C,.   WORKS THROUT IN FIXED-POINT. NO DIVIDE
      INTEGER x 2N1, B1 (5000), N2, B2 (5000)
      COMMON/WNGLN/N1, B1, N2, B2
      INTEGER x 2 M, I TAB (5000), J TAB (5000)
      COMMON/PTRTBL/M, ITAB, J TAB
      INTEGER LSQ(3), IT(3)/2, 0, 2/,JT(3)/0,2,2/
      I = 1
      J = 1
      K = 1
1.    LSQ(1) = (B1(I + 2) - B2(J))**2 + (B1(I + 3) - B2(JH))**2
      LSQ(2) = (B1(I) - B2(J + 2))**2 + (B1(IH) - B2(J + 3))**2
      LSQ(3) = (B1(I + 2) - B2(J + 2))**2 + (B1(I + 3) - B2(J + 3))**2
      IL = 1
      IF (LSQ(1).LE.LSQ/2)) GO TO 2
      IL = 2
2.    IF (LSQ(IL).LE.LSQ(3)) GO TO 3
      IL = 3
3.    I = I + IT(IL)
      J = J + JT(IL)
      ITAB(K) = I
      JTAB(K) = J
      K = K + 1
      IF (K.LE.M. AND.I.LT.N1. AND.J.LT.N2) GO TO 1
      M = K - 1
      RETURN
      END
```


Fix up: Restart at one end point of the turnover chord. Otherwise, we are in case II (Fig. 7e): Mark the second turn and scan node for longer chord, discarding chords as we go. Amputate the thumb (discard, or do later) and continue until new trouble.

INDEX PAPER

PART B. INFORMATION MANIPULATION

B2. Classification Change

Facilities for changing the classifications of areas and regrouping areas into new classifications may be separated into three categories. The first is classification substitution. This consists of changing the description of the contents of existing areas; the file may also be resequenced to reflect these changes. This facility has been implemented in various systems through file copy and update routines, In conjunction with utility sorting programs.

The second technique involves the grouping of contiguous areas into a single area, and assigning of new classifications. This is discussed as part of the "Dissolve" facility.

The third file updating technique requires the modification of boundary lines as well as classification changes. This may be done for error correction purposes or as a method of keeping the information current. Off-line facilities are available for specialized and limited purposes that permit this form of manipulation, but these facilities tend to be very complex and time consuming to use. The best methods of accomplishing these changes are by using on-line graphical display devices. A less cumbersome technique is the use of cathode ray tube display systems. Several different on-line image manipulation and file updating facilities have been developed which permit several levels of detail, image expansion, generalization and modification. There has been no work of note beyond the feasible state of the art.

B3. Contouring Automatically

Automatic contouring, as a subset of trend surface analysis, is one of the oldest geographical manipulation facilities. Programs have been available for the IBM 1620 to analyze a set of coordinate values defining a surface, and to plot the resultant contour map. More extensive programs are now available on new computers (e.g., IBM 1130) for contour map plotting and the associated trend surface analysis calculations. Contouring facilities are also available for discrete image printouts in conjunction with the SYMAP programs.

Elementary routines are used for numerical approximations, polygonal curve fitting, smoothing, interpolation and other surface analysis techniques. The routines are quite efficient (but quite time consuming in computer usage) and are well documented.

An example of a limitation is given in Fig. 8, where a contour defined by the set of points shown in A may be drawn as B, C or D. The algorithms for considering the adjacent data are elementary, and the surface "trends" towards the options, except where it specifically intervenes to make one situation impossible.

Most analysis routine options are only manually specifiable, and do not automatically analyze the data content to make the decision as to whether they should be invoked, or if so, with which parameters. As a result, for proper usage they require that an individual be quite well trained. These available stand-alone facilities are in a developed state of the art; however, their use in other automated systems in which they could be implemented as an option is only at a planned stage.

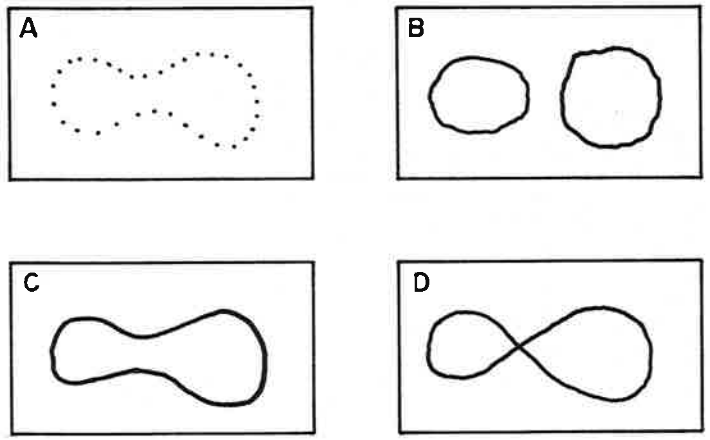
Automatic contouring on input through the use of stereo viewers or other devices was not considered as a manipulation facility and is not within the scope of this review.

REPORT OF DISCUSSION

PART B. INFORMATION MANIPULATION

B3. Contouring Automatically

It was pointed out that the three figures originally included in this section (Fig. 8a, 8b, 8c) should also include a fourth diagram (Fig. 8d) which shows the contour lines as being self-crossing. Such a situation might occur at the pass located between two peaks in a given terrain model.



PROBLEM CASE ASSOCIATED WITH AUTOMATIC CONTOUR ANALYSIS; B,C and D, ALTERNATIVE CONTOURS INTERPOLATED FROM SET OF POINTS, A.

FIG. 8, SECTION III

INDEX PAPER

PART B. INFORMATION MANIPULATION

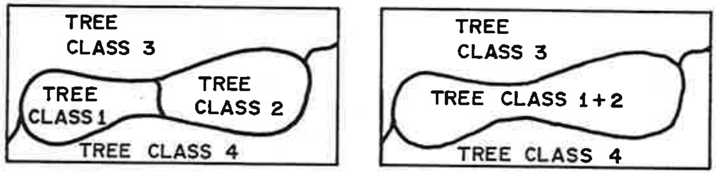
B4. Dissolve

The "dissolve" facility permits the removal of selected boundary lines in order to group contiguous data groups or areas which are to be treated as single units (Fig. 9). Normally this is done for a specific purpose, such as the grouping of the three top agricultural land classes as "prime" quality land in order to determine the quantity, size and location of such land areas. Optionally, this facility may be used as a technique for creating a new map file, such as by dissolving all the land boundaries to form a skeleton land and water coverage. The facility should permit the specification of the boundaries to be dissolved, either by indicating the two contiguous areas between which the common boundary is to be removed, or by indicating that at all contiguous occurrences of specific land classifications the boundaries are to be removed (as in the above examples). At the time of removal of the boundary, one of the two contiguous land areas must be removed from the data file or reclassified so that the two match. It is desirable to alter the classification of the remaining area in order to derive the maximum benefit from the facility.

This is a very powerful and complex routine. It requires the existence of a multi-file map storage system of continuous image data or variable area discrete image data. The system must permit the manipulation of one map to create another and enter it into the data bank.

B5. Distortion Elimination

Elimination of linear and non-linear distortion is required by and unique to highly sensitive input systems. These facilities permit the removal of minor errors that would otherwise occur in the input maps as a result of temperature changes, input document orientation (on digitizer table, scanner drum, or bed), input document stretching (equivalent to temperature variation) or other similar effects. Linear error elimination may be done by a co-linear transformation that will maintain the existence of straight lines. This is done by mathematically shrinking one (or more) end of the map data so that it will match other adjoining maps of the same known size.



FORESTRY

"DISSOLVE" FACILITY TO REMOVE BOUNDARY LINES BETWEEN CONTIGUOUS AREAS AND REGROUP.

FIG. 9, SECTION III

These routines are easily implemented but often avoided as the computer time involved for their application to every point on a continuous image map is substantial. If a coarse storage system is used, the errors of ten to twenty mils (1 mil = 1/1000 of an inch) may not be significant unless multiple maps are to be joined together on a common co-ordinate system.

Non-linear distortion is more difficult to identify and remove. The causes are similar but extend to include lens aberration of viewing devices and the earth's curvature effect on remote sensing or photographic input. The requirement for error removal is more pronounced as the matching of information for adjacent areas is essential. The recommended procedure is selection of data from the central portion of the input and its subjection to line fitting routines, used in conjunction with the linear error elimination routines.

Linear elimination is at the feasible state of the art, whereas non-linear distortion removal is at the potential stage.

REPORT OF DISCUSSION

PART B. INFORMATION MANIPULATION

B5. Distortion Elimination

The mechanical reliability of computer hardware was questioned and a data bank being maintained by the CIA was cited, which included the Danish islands of Punin and Zeeland and which, for some unknown reason, lost these islands. It was not until six months later that anyone even noticed that the two islands were missing from the data bank, which was in daily use. R. Shaw W. Schmidt

It might be more appropriate for contour lines to be shown as fuzzy, rather than sharp, lines on a map, in order to represent more correctly the accuracy with which this information should be maintained, or rather was recorded. A student at Harvard, Mr. Adrian Thomas, during the past year, has been experimenting with the production of contour lines displayed on a perspective view of a three-dimensional surface, where these lines are in fact represented as fuzzy lines. This is accomplished by representing the contour lines by use of Moire patterns. Equally important is the fact that the lines as shown vary in their width as a function of the slope of the surface; in relatively flat areas the lines W. Tobler T. Waugh

are quite broad, whereas on steep slopes the lines are quite narrow. Such a representation is in keeping with the actual differential variance of the contours as they exist in the real world.

INDEX PAPER

PART B. INFORMATION MANIPULATION

B6. Generalization

Generalization is the process of grouping many small areas on a map and assigning to them the classification that is most predominant for the group, or assigning the group a multi-classification with associated percentages to indicate the average description of the area. This is commonly done in regions of mixed small land areas where the information would be too dense if maintained on a more detailed scale. The technique is required for land classification of areas stored in discrete grid groupings.

Automatic generalization as a manipulation facility is required only in very dense areas as a result of a scale change or an overlay. Normally the input data will be prepared in sufficiently large groupings for the system. The difficulty associated with an automatic generalization facility which would be invoked when necessary as a result of a scale change or overlay is that it is tied to the data classification. If all land classifications were stored in a fixed format it would put a limitation on the freedom of classification (see "Complex Classifications") organization and detail. If that is sacrificed, then totals may be taken, throughout the region to be generalized, of land area by classification. A new classification may then be formulated representing the resultant combination that is formed.

These facilities include: smoothing image data; area grouping; data removal (e.g., islands); and feature retention (e.g., passages or peninsulas vs. islands).

There is no information systems currently under development that include this as an automated manipulation feature. It is at the planned stage of development.

B7. Generate - Select

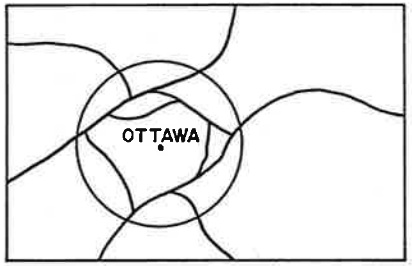
The generation of pseudo-maps for manipulation, selection and retrieval facilities is a desirable feature. Circle generation permits inquiry of the form: Find all areas greater than 10 acres in size within 50 miles of Ottawa that have land classification "A1 B3" (Fig. 10). Polygon generation permits the user to specify (probably by longitude-latitude pairs) the extreme points of a polygonal figure defining his area of interest for a similar inquiry (Fig. 10). Band generation allows a request to be formulated that would select all the land within a given distance of a given line. An example might be to determine what the assessed value of all land (from a map of property assessments) within a half-mile of an existing road would be to ascertain expropriation costs for building a new throughway. Line generation would be associated with a facility to list all areas that were "cut" by a line joining a specified set of points. This is slightly different from the polygonal specification that results in a request for all areas "contained" in the indicated region, but a natural extension of the concept.

Several systems in operation are preparing a limited use of generated circles, polygons, grids, ellipses or other figures for selection purposes. Significant theoretical work has been done to describe the techniques and methods for implementing these search routines in continuous image systems. There is much more effort required, however, before these algorithms will be economically feasible. It is best described by considering the circle and polygon generation-selection routines to be in a feasible state of the art, and describing band and line cut selection as experimental.

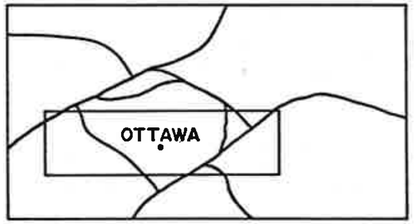
B8. Intervisibility and Route Determination

Three-dimensional intervisibility between two points and the more complex pathfinding techniques are in many ways similar to automatic contouring and trend surface analysis. Both subject groups consist of a more specific field of general interest and a broader, more all-inclusive subject of more specialized interest. The mathematics associated with the topics is a similar branch of topology. The most advanced work in the two fields has been done by the military. Both subjects are used by the military in war games simulation routines. Both require altitude information as well as co-ordinates (three-dimensional data storage). Few systems that do have such facilities are tied into broad data-base

COVERAGE AGRICULTURE
GENERATE CIRCLE



COVERAGE AGRICULTURE
GENERATE POLYGON



ADMINISTRATION
GENERATE BAND

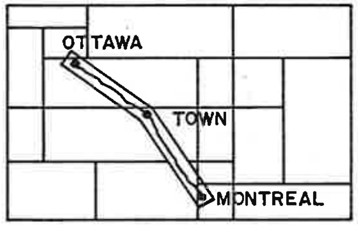


FIG. 10, SECTION III

geographical information systems. The routines tend to be heavy in computer usage and hence fairly efficient, as they have been optimized significantly.

Intervisibility between two points means that there is no obstruction between those two points that would hinder an individual located at one of the points from seeing the other. Basic hill shading and radar shadow determination routines have been developed. Associated routines include locations of the nearest point from which a second is visible when intervisibility does not exist. Another facility tied to the same basic set is that of ensuring that there is no intervening obstruction to a projectile's trajectory from one point to another. Pathfinding, which is the generalization of all the facilities, involves the movement of an object from one point on the earth's surface to another according to a set of specified rules. The rules, for example, may not permit leaving the earth's surface or a particular contour height except under specified conditions. A "shortest" route requirement must consider hills and valleys.

The definition of what such a facility is to include is clear. The development of the algorithms to handle automatically the analysis and solution are in their infancy. Several basic facilities are available but not extensively developed. For comparison purposes, these facilities are only in a feasible state of the art.

REPORT OF DISCUSSION

PART B. INFORMATION MANIPULATION

B8. Intervisibility and Route Determination

Intervisibility is used in a digital terrain model for the purpose of placing radar stations in order to detect incoming aircraft. The same application is also concerned with the identification of flight paths which would have minimum exposure to known radar installations.

It is possible to determine the landscape view areas visible from any selected station point. This work is being done in connection with a highway route location project and utilizes a program originally developed by Mr. Amidon. Cathode ray tube images of structures in an urban area can also be shown in colour. C. Steinitz

The purpose of the latter work, being done by Kamnitzer at U.C.L.A., is to allow the user to "drive through" an urban area and essentially to experience the urban scene from various vantage points. It is also possible to do such things as: (1) given a digital terrain model, identify graphically the peaks, pits, passes, pales, courses and ridge lines on that surface and (2) display the magnitude and direction of slope implicit in a digital terrain mode. A. Schmidt

Intervisibility is merely one sub-category of a larger group of special analyses features that might be applied to a digital terrain model. It is a step beyond the sieving process which is a somewhat more primitive analysis operation. Numerous spatial analysis programs have been developed at Harvard over the past few years, and a recent inventory of such programs turned up 24 discrete programs. Many of these had been developed independently of each other and as a result were not totally compatible for use in a sequential fashion. As a result, one of the projects for the current year is to develop a package of interchangeable and modular programs which can both deal with various kinds of spatial analyses and at the same time be used sequentially when this is desirable. C. Steinitz C. Steinitz

INDEX PAPER

PART B. INFORMATION MANIPULATION

B9. Measurement

Various characteristics of stored data may be automatically calculated and maintained for future reference. Usually these measurements are only made upon the request of an inquiring program, but occasionally the most frequently required ones are kept with the classification data.

In the case of fixed grid systems the area is a fixed amount. For continuous image systems the area calculations are straightforward and the area is maintained in many systems with the classification. Length calculations for boundary or perimeter distance of areas and for road lengths on line maps is not quite as simple as may be expected. Apart from minor errors due to hills, latitude or map projection, there is a major error potentially introduced if smoothing is used to reduce data volume in a continuous image system. A movement of one unit north plus one unit east is 1.4

times as great as a movement of one unit diagonally northeast. This is one of the factors that must be considered. Simple approximation calculations exist but few take all factors into consideration.

Counting facilities are normally called in only as response to particular inquiries. They then are used to determine the occurrence of a classification of interest. A sample inquiry using count, area and length calculations would be: "How many areas greater than two square miles in size have boundaries that may be enclosed with less than ten miles of fencing?"

B10. Merge

Merge facilities permit the combination of maps that do not overlap but do contain the same information characteristics about the land they represent. The facility combines the separate data files to form a single data file representing the larger area. Normally this is only required for contiguous maps that represent a continuous area for study, but it may be used for any two map sets. An example of its use would be in the process of building a data bank representing a large area. The only limitations on this facility will be those of the design of the system itself. Standard manufacturer-supplied software programs are readily available to merge two files into one larger one.

The system must provide the routines that will update the master indexes or directories, maintain intra-file linkages and cross-references, ensure that reference numbers are not duplicated and ascertain that classification data do not become confused for areas that flow from one map to another on contiguous maps.

These facilities have been quite clearly defined but not completely implemented except to a feasible level in any system. The file merging is done readily enough, but often there are limitations on the matching of contiguous maps.

B11. Overlay

The overlay facility permits the combination of maps that contain different information about the land they represent (e.g., agricultural potential and political boundaries) and may overlap partially or totally, providing information concerning the same land areas (Figs. 11-15).

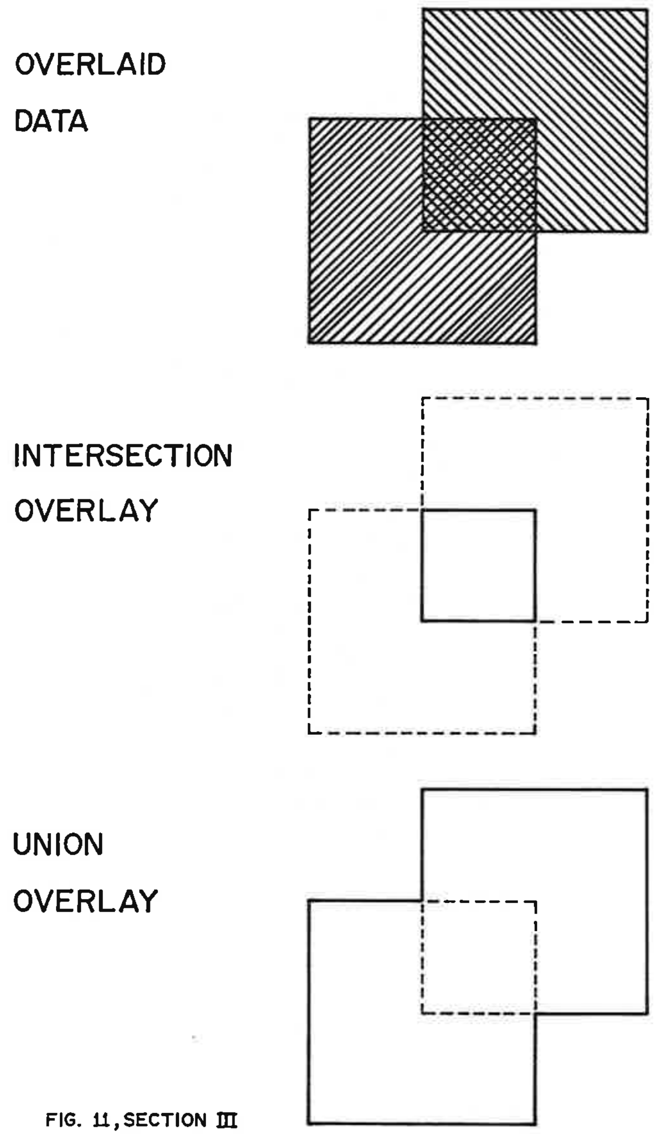
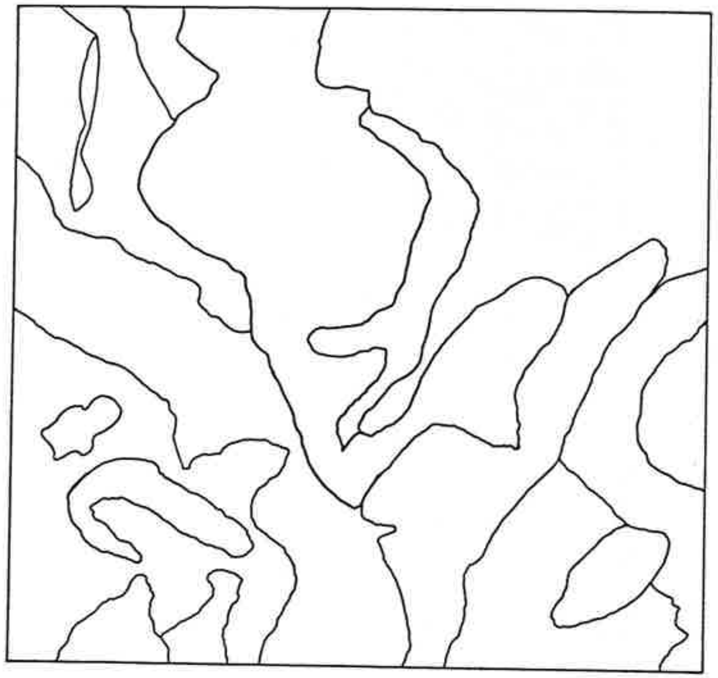
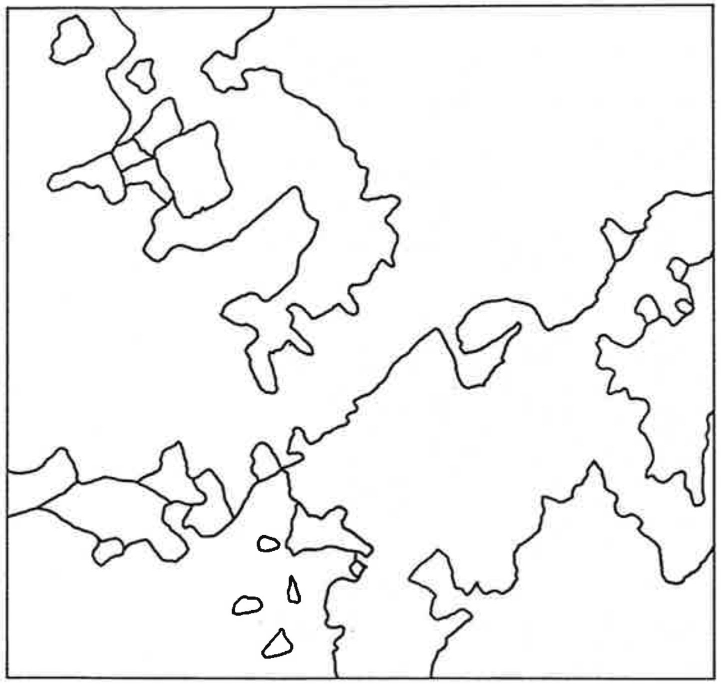


FIG. 11, SECTION III



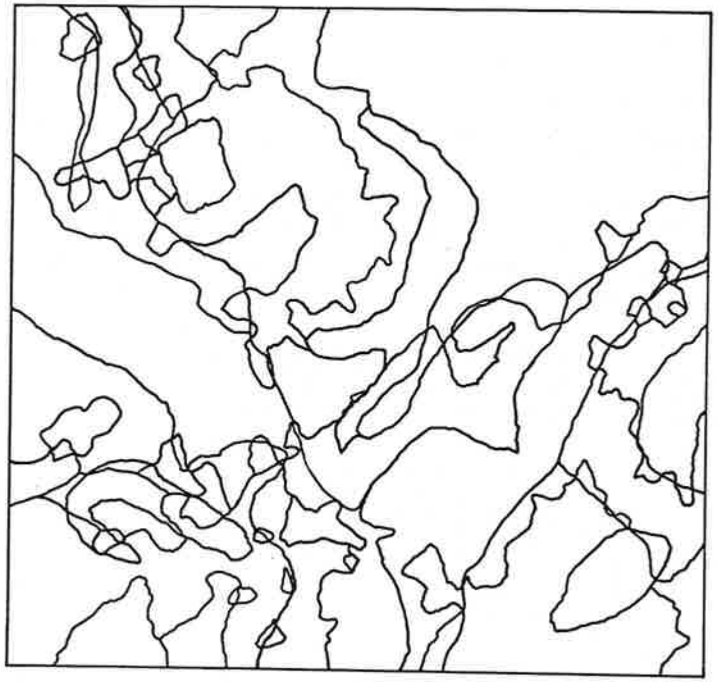
EXAMPLE OF BOUNDARIES ON TYPICAL SOIL CLASS MAP

FIG. 12, SECTION III



EXAMPLE OF BOUNDARIES ON A TYPICAL PRESENT LAND USE MAP.

FIG. 13, SECTION III



RESULT OF COMBINED SOIL CLASS AND PRESENT LAND USE MAP BOUNDARIES.

FIG. 14, SECTION III

ADMINISTRATION
+ CENSUS +
GENERATED CIRCLE

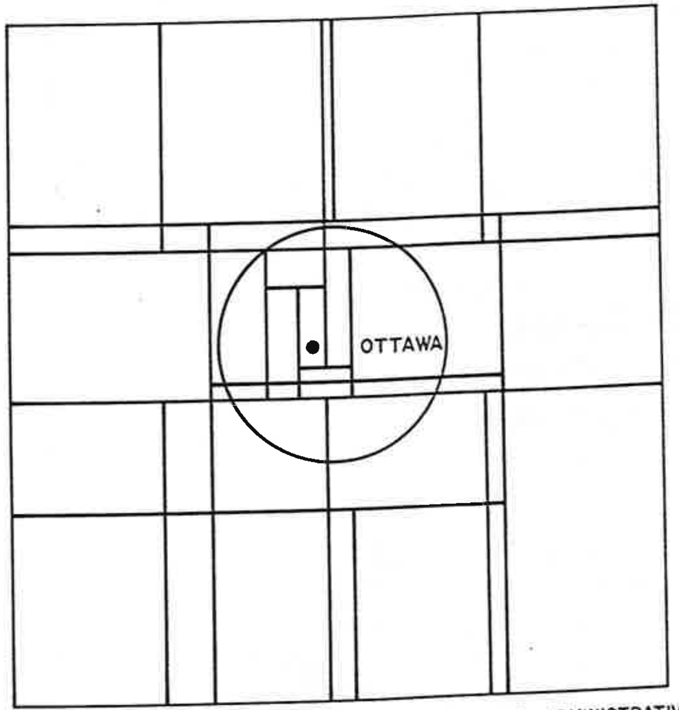


DIAGRAM ILLUSTRATING OVERLAY OF HYPOTHETICAL ADMINISTRATIVE AND CENSUS DISTRICTS COMBINED WITH A GENERATED CIRCLE.

FIG. 15, SECTION III

DATA RETRIEVAL

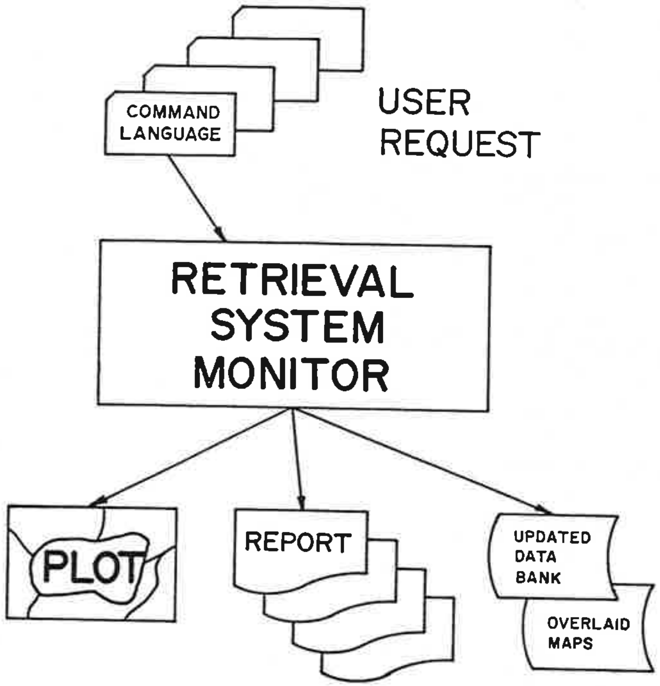


DIAGRAM SHOWING THE RELATION BETWEEN USER REQUEST, COMMAND LANGUAGE, RETRIEVAL SYSTEM MONITOR AND EVENTUAL OUTPUT OR UPDATES.

FIG. 16, SECTION III

On discrete fixed grid systems this facility simply requires a combination of the data about the discrete units to form a longer classification. It is developed only to a feasible stage for these systems.

For discrete variable grid systems where areas are represented as a single point, this facility is limited. As the area boundaries are not fixed, and also not stored, their intersection cannot be determined. Hence, the effect of an overlay of another such area in creating new areas with intersecting boundaries cannot be calculated.

In a continuous image system this becomes an extremely powerful and very desirable facility. It is also, without doubt, one of the most complex facilities to provide. By the creation of a single data file, with all the information contained in several separate map files plus the effects of boundary intersections in causing a multiplicity of small multi-classified land areas, many complex questions concerning the areas represented may be asked. Examples would be: (1) Find the location and size of land areas within Rideau county suitable for growing high yield crops (Agricultural class 1 or 2) that are currently used for recreation, are undeveloped, or are under forest (Present land use of R, U or F). (2) Combine census maps and agricultural maps to determine areas with high quality land but low average income. (3) By adding maps with watershed, climate, land potential for forestry, wildlife, recreation and geological data the variety of questions which may be asked is limitless.

This facility has currently reached a developed state of the art in the Map/Model system and the Canadian Geographical Information System.

Difficulties encountered are associated with extremely dense maps, as the volume of information is compounded as a result of an overlay. The other major limitation is the quantity and variety of maps stored within the data bank. In the time available, priorities must be set and areas selected from all of Canada for inclusion.

REPORT OF DISCUSSION

PART B. INFORMATION MANIPULATION

B11. Overlay

When combining the data from two separate maps to create a third map one must be able to transfer the values from the first two maps and transform them onto the third. In doing so, it is necessary to recognize that the values may be of two types, either absolute data or density data. For absolute data, the values as they appear on the third map in any geographic zone must be shown as a proportion of the original value, where the proportion is a function of the relative size of the new zone to that of the two original zones. Density data, on the other hand, is by definition expressed as a ratio value and as such may be transferred directly. T. Waugh

The procedures used by the Map/Model system for identifying and recording the areas which are created by the combination of two or more original maps, were explained. S. Arms

INDEX PAPER

PART B. INFORMATION MANIPULATION

B12. Projection Change

Map image information is usually stored in one projection for manipulation purposes. Projection changes may be done on input maps to convert them to the common storage base, or to the output maps that have been requested, just prior to plotting them. Some map reproduction systems will store image data in any projection and convert it to any other for plotting, but will not provide other manipulation facilities as they do not maintain data associated with classification or area.

Projection change of some form on input is easily done by a sub-routine in most continuous image systems, so that conversion from UTM, Mercator, or polar projections to the storage medium (UTM or longitude-latitude) may be accomplished. Just prior to feeding data to a plotting routine, the reverse conversion will provide plotted output in any desired projection for which a

mathematical definition exists.

These routines are available in a sophisticated state of the art in some systems such as the AUTOMAP system with 15 optimal output projections, and could be incorporated into many others as required.

REPORT OF DISCUSSION

PART B. INFORMATION MANIPULATIONS

B12. Projection Change

Several computer programs developed by the CIA for the purpose of computing changes in map projections are available. When converting input data, their system is able to transform the co-ordinates into any one of five different map projection systems. For output purposes, they can work with up to 16 different map projection transformations. W. Schmidt

INDEX PAPER

PART B. INFORMATION MANIPULATION

B13. Scale Change

Scale change facilities are used both for manipulation and output purposes. They are simpler than projection changes and easily implemented for plotting or output facilities. Similarly, image expansion for manipulation is simple and worth little investigation. Difficulties occur with image reduction (or reduction followed by expansion).

At this point a short note of justification for manipulation scale changes will be inserted. Often maps at different projections come into a system that it is desirable to compare, overlay or match together (merge). It is also logical (if possible within the system) to store maps with a lower density of data at a coarser scale to reduce storage volumes. In these instances, it is advantageous to use the computer's power to do the work.

The difficulties that occur with image reduction are associated with the merging of lines that were formerly separate, and the density of resulting data. A cartographer will decide what lines should not touch, a small area or island should not disappear or that certain data are either less important or extraneous. This results in avoiding the problem on a manual system; lines do not merge, important data are retained, yet the map is not cluttered due to data density. However, a set of general rules that may be converted to computer code and that will simulate the cartographer's decisions has yet to be found. The existing techniques either do not let anything disappear, or let everything disappear (creating islands of peninsulas, etc.). A more sophisticated algorithm will vary its decision based upon an importance indicator attached to the data, in conjunction with a factor based on the extent of the scale change, and generalization routines to group many small areas into a single large one. The cost of developing such facilities has been deemed prohibitive in the past, relative to their value.

Considering the expansion of the image following the reduction, it becomes apparent that the existing facilities are as practical as the more sophisticated theoretical routines would be. It is not reasonable to retain information concerning what has been removed (volume of data is usually the reason it was removed). Therefore, when the image is expanded the data are not reinserted, and lines which were kept apart are separated farther than they would have been originally.

Reduction or expansion of an image by scale change is a logically required facility and at a reasonably implemented stage, as it is now, without generalization routines. Any requirement for a more sophisticated implementation that may exist, especially requiring image reduction followed by expansion, must be critically examined for the logic and true need behind it. There may be a simpler alternative, such as using the original map.

B14. Search

Three types of search concepts may be defined. The simplest is the search for a given profile of characteristics. The second type is to search the data bank for all points, lines or areas contained, cut by or touching a given set of points, lines or areas. The third form is the closest search concept used to find the nearest point, line or area of a given profile to a given set of points, lines or areas.

The result of any of the search routines defined may be either

a list in report format or a sub-set of the total data in standard data bank file format. This facility may be used to answer questions such as where land suitable for camp-grounds touching a lake within five miles of a main highway is located.

The ability to search for data of a given profile and to create a sub-set file or to list the data is available, being implemented on many systems. The other selection facilities are only planned for development.

B15. Statistical Routines

Statistical routines may be valuable when analyzing the contents of a data bank. Such routines vary from simple linear regression and variance calculations on information content (classification data, especially census data) to more complex multivariate regression and correlation analysis. The mathematical routines are available, many as pre-coded sub-routines in the IBM Scientific Subroutine Package. They may be tied into any system that has manipulation facilities sophisticated enough to which new routines may be added, and that maintains the data content required in the classification to permit it to be submitted to statistical analysis.

B16. Simulation Models

Simulation models that use and manipulate information stored in a geographical system could provide valuable tools for decision makers. Land use models could forecast potential crop yields if land was subjected to improved fertilization. Loss of income and food crops versus transportation costs and land use capability could be studied in planning locations for new cities or industrial centers. This would alter our current misuse of valuable agricultural land. Urban planning and cost of road development may be determined.

These are the basic underlying purposes of many of the geographical information systems under development. These are noble and meaningful aims which have yet to be implemented. Some of the systems are now operational on a basic level. The purposes for which they were being developed must not be forgotten. Further development must continue until the results for which the systems were intended are being produced.

Most systems are open-ended and permit the inclusion of routines for simulation of altered conditions on the data they contain.

REPORT OF DISCUSSION

PART B. INFORMATION MANIPULATION

B16. Simulation Models

A consensus was sought on the meaning of the term "simulation", as used here. In the opinion of the author, the use of the term should include those operations normally referred to as "models" as well as those of a probabilistic or Monte Carlo variety. It was proposed that the term be used in its broadest sense, but that it not include physical or analogue models. — D. Marble, P. Kingston, D. Sinton

Several slides of data drawn from a data bank maintained for the south-east sector of Boston were shown. These slides illustrated the manner in which a variety of data, such as water pollution and land development, can be combined by logical operations for the purpose of evaluating the suitability of land for future developments. — T. Waugh

One of the most important search operations which can be performed on a geographic data base is that of optimizing route location. A computer program designed to assist in the identification of optimum routes for purposes of locating a highway within a transportation corridor has been developed in the Department of Civil Engineering at Purdue University. — D. Sinton, D. Marble

Research is being done at the University of Kansas, in a transportation geography course, that relies heavily on techniques for simulating, analyzing and evaluating transportation networks. Similar work is being carried out at the U.S. Army Electronic Proving Ground, Fort Huachucha, Arizona (64), that includes optimizing techniques (linear programming, dynamic programming, Monte Carlo, etc.), though does not include parametric programming. — B. Wellar

A number of people expressed an interest in optimum route location procedures, and in response to a query as to who might like to obtain more information about this topic and possibly correspond in the future, the following people asked that their names be recorded:

E. Amidon	D. Marble	D. Steiner	B. Wellar
K. Dueker	W. Schmidt	C. Steinitz	
V. LaGarde	D. Sinton	W. Tobler	

Mr. Sinton stated that the following would also be interested:

Mr. W. Warntz,
Harvard University,
Laboratory for Computer Graphics and Spatial Analysis,
Cambridge, Mass.

Mr. G. Elsner,
U.S. Forest Research Center,
Berkeley,
California.

In addition to the several programs available at Harvard, there also exist a number of fairly large geographic data banks describing suburban and rural areas at a scale ranging from one kilometer to 1/100th of a kilometer in grid cell size. These data banks will be made available to others who might be interested in using them in their experiments for such purposes as optimum route locations. — C. Steinitz

INDEX PAPER

PART C. SYSTEM CONTROL FACILITIES

C1. Monitor Systems

Monitor systems are used as system control facilities for information systems with many files and many possible routines (Fig. 16). They permit the user to call for the routines and data files required without manually setting up special sets of command language cards to extract the programs desired from libraries, and to describe the files to be accessed. All this will be done for him automatically, as a result of an analysis of the specific query which he formulates. These automated systems are very complex, yet they remove the housekeeping burden from the query programmer and place it on the machine where it belongs.

Elementary monitor facilities have been developed in the Canada Geographic Information System. Further development is under way: however, it is limited until all the facilities which are to be tied together by the monitor are completed, defined and at least partially implemented.

C2. Manipulation and Query Languages

Another feature that facilitates the user interface to a system is an inquiry language. This facility permits a user to formulate his question in a simplified format that is interpreted by the computer. The computer will in turn,under the control of a monitor system as just described, call upon the required data and programs to provide the response. This reduces, and in some cases eliminates, the need for a programmer to act as an interface between the user and the information system.

In order to clarify the distinction between a monitor system and a query language (although they are usually implemented together), several examples will be given.

Let us suppose that an inquiry may be to report on all land areas greater than 10 acres in Rideau County within Ontario that have agricultural classification "A", giving their reference number, centroid location and acreage. Through the use of a query language that inquiry might be coded as:

REPORT (REFERENCE, LOCATION, ACREAGE)

IF (PROVINCE=ONTARIO, COUNTY=RIDEAU, AGRICULTURE=A1, ACRE>10)

If, however, a special subroutine had to be programmed to provide the same logic, and it took several hours to do so, then a query language does not exist. More complex requests requiring the use of several manipulation facilities may involve several lines of coding and either special training or the use of a reference manual.

If a monitor system exists, then the inquiry above would be combined with four or five control cards which would call interpretive routines. These routines would analyze the inquiry and decide upon the data files and manipulation programs required to satisfy the request. Automatically, all the required control cards and instructions to the computer operator to access the required data files would be generated, processed and the results produced.

In the event that a monitor system is not being used, the inquiry must be combined with sizable card decks containing programs and operator instructions. Also, additional cards would have to be prepared, depending upon the information required by the inquiry, to indicate to the computer which files must be called out of the library to provide the necessary data. In many cases this would be a multi-step operation involving much card handling and preparation, and would be subject to human errors.

A basic query language has been implemented in both the Urban Geo-coding and Canada Geographic Information systems. The DBS Urban Geo-coding language is specialized in purpose, for inquiry rather than for manipulation, but provides very flexible inquiry capabilities. The Canada Geographic Information System has a powerful, but not completely implemented, manipulation language. As it was designed for manipulation flexibility with many files it is somewhat more complex, requiring specialized training for use. It does not provide simple report inquiry facilities. It is very much open-ended in design and these routines could readily be added. As these facilities are subject to extensive further development, they must be classified as in a feasible state of the art.

REPORT OF DISCUSSION

PART C. SYSTEM CONTROL FACILITIES

C2. Manipulation and Query Languages

U.S. Corps of Engineers Waterways made the decision to rely upon the FORTRAN language as their basic query language. On the other hand, the Canada Geographic Information System, the DBS Urban Geo-Coding System, the Map/Model system and others were using the PL/1 language or languages based on PL/1 for this same purpose (34). — V. LaGarde, P. Kingston

Neither FORTRAN nor PL/1 are suitable for use as user-oriented query languages. A query language should require minimum effort and programming knowledge on the part of the user. Higher level languages such as FORTRAN are suitable for the preparation of a query language, but should not be considered as query languages in and of themselves (19,20,21). — K. Dueker

Languages called DATMAN written by Waugh; DATA-TEXT, a social science research language prepared by the Department of Social Sciences at Harvard; SPSS (Statistical Package for the Social Sciences), originally developed at Stanford but now being distributed by the Public Opinion Research Center at the University of Chicago; and GRIDS (Geographic Retrieval and Information Display System), developed by the U.S. Census Bureau and currently being used by their offices at the Southern California Regional Information System Studies (SCRIS) have been used. — T. Waugh

INDEX PAPER

PART D. INFORMATION RETRIEVAL FACILITIES

Information retrieval will be treated briefly, as it overlaps into the realm of data output which is covered in the next section of the conference.

D1. Display Techniques

Display techniques involve basic printed reports, plotting, CRT displays, solid black-white or colour film printing and others, all of which will be covered on Day Four. Here we are more concerned with the manipulation routines that precede the display. Included in this category are some of the facilities such as scale and projection change, mentioned earlier. Other facilities include the SYMVU routines that calculate perspective viewing for contour data to be plotted. These routines are well developed and provide dramatic results. Very little has been done to interface the many potential output facilities that could be tied into major geographical information systems. The displays provided by CRT's, plotting devices, SYMAP line printer maps, SYMVU and automatic contouring could all be tied in with the display facilities of any continuous image geographical information system. Only minimal work has been done in the linking of these facilities (themselves often more highly developed) to total systems. In this area the state of the art is at a feasible level, with much thought having been expended but little implementation accomplished.

D2. List Data of Given Profile

Listing data of a given profile results from the selection procedures of the search facility described earlier. Automated reporting facilities that produce this information are particular to individual file organizations. Associated with query languages (as opposed to manipulation languages), there may be an automated report-generating facility. Often these languages are not easy to learn and use.

A lot more work may be done to simplify the use of most systems in providing data lists. More easily used query languages and report generators would increase the data accessibility to the users for whom the systems are intended. Although much effort has been expended in this area, a significant amount of additional work must be done to bring the user-oriented interface beyond the feasible stage of development.

D3. Man/Machine Interaction for Correction, Updating and Query

Man/Machine interaction is best accomplished by some form of on-line computer with a feedback loop to the user. Frequently this is achieved by the use of a computer terminal, a CRT display or an on-line digitizing device. In an off-line operation it may consist of submitting a request to a computer program which produces an analysis with or without error messages and a report back to the user calling for further action.

The interaction is required for one of three purposes: (1) input editing of data and error correction; (2) file updating or modification; or (3) inquiry and response.

To some extent, similar routines may be used for input data and file updating. Much of this may be handled from a typewriter terminal. However, where a continuous image is involved it is much simpler and quicker to use a display facility such as a CRT to correct or modify map image data. Although CRT displays are preferable, for cost reasons typewriter terminals are usually implemented.

As hard copy output is desirable, a typewriter terminal more frequently meets the requirements for inquiry results. Plotted output provides a permanent record of map data and may easily be requested from a terminal for off-line production. Devices are now available to produce 8-1/2 x 11 inch copies of IBM 2250 CRT displays; however, only a limited amount of image may be shown on the CRT screen which is smaller than the average-sized plotted output that would be requested.

For editing and correction purposes, the Canada Geographic Information System has a sophisticated user "Manual Error Correction" language for off-line batch submissions. This permits extensive image manipulation and modification for error correction. An extension of the concepts could be adapted to terminal usage and data bank updating.

An on-line CRT for image editing and correction has been developed for the TOPOCOM system. On-line interaction and editing directly linked to a digitizer has been studied by the Oxford

Experimental Cartography Unit, producing valuable and interesting comparative results. Similarly, interesting work has been done for the Canadian Hydrographic System.

Updating techniques involving man/machine interaction are often similar to the editing facilities, but usually not as well developed. Much experimental work has been done in both inputting and updating. The results seem consistently to indicate that a totally automated entry system is virtually impossible, and therefore man/machine interaction is essential. However, the results are so extensive and diverse that, although several systems have facilities implemented, the optimum result is not clearly definable and the field is still under development.

In the area of man/machine interaction for query, several systems have been developed without on-line facilities but with allowance for them to be added. In these systems, batch or remote inquiry is used, providing similar facilities with slower turn-around. On-line inquiry has been developed, however, for systems that are intended to operate in a real time environment. The AUTOMAP and Waterways systems have interactive inquiry facilities. CRT inquiry and updating of image data has been experimented with, using aeronautical charts, for the U.S. Coastal and Geodetic Survey. The value and convenience of on-line inquiry are undisputed. Manipulation facilities associated with the inquiry systems are extremely limited. Significantly more facilities will have to be developed before truly sophisticated on-line inquiry and manipulation of geographical information will be possible.

REPORT OF DISCUSSION

PART D. INFORMATION RETRIEVAL FACILITIES

D1. Display Techniques

D2. List Data of Given Profile

D3. Man/Machine Interaction for Correction, Updating and Query

Mr. Kingston preferred not to begin discussion on this section, since it was of primary interest to the following day's agenda.

INDEX PAPER

Given by: W.R. Tobler

Data display and presentation range from the treatment of specific hardware to visual perception, with several other considerations along the way. It is obvious that this is a big subject. I also believe that almost everyone in this room has at least five years' experience, some maybe considerably more, in geographical processing and display, so that the collective experience is so much greater and more extensive than my own in probably all specific items, I will to some extent retreat to a philosophical position and give a professor-like lecture. In the suggested classification of stages of development I believe this would come under "potential" rather than "implemented".

In particular, being at a university I have no operational responsibilities except teaching. This naturally colours my point of view. For example, there is not much point in teaching undergraduates how things are done today because the technology, particularly hardware, will have changed by the time they graduate. It is sometimes estimated that the half-life of knowledge is about five years, that is, one-half of what one learns will be obsolete five years from now. To teach effectively it is therefore necessary to attempt to identify fundamental concepts; that is, concepts with a half-life approaching 20 to 50 years. As you would expect, this is not at all easy, but let me give you an example. It is now approximately 20 years since the first maps were produced by computer. I know that most of you are aware that the Swedish meteorologists were producing contour maps on a line printer as early as in 1952. My own experience goes back to about 1957 when I first used a Benson Lehner plotter to draw an outline map of the United States. The point is that since that time there has been only one development which has surprised me, and that was six years ago. Frankly, I would like to be surprised more often, but the reason that I am not is because of the fact that I accept the general sense of Turing's theorem; i.e., that any procedure which can be specified in a finite number of steps can be automated. It is thus a priori clear that almost anything is potentially capable of being automated, and certainly this is true of most cartography. My experience is that a large number of cartographers were very surprised when Mr. Yoeli demonstrated that we could do shaded relief by machine. I was very glad to see that he had been able

to work out the specific details of how to do it, although it did not really surprise me because, as I have just stated, it was a priori clear that this had to be possible. Of course, Turing's theorem does not say anything about speed, cost, memory requirements or the technological state of the art, and so on. These difficult topics lie in the area of development and implementation, and have really been the concern of people working in the operational agencies.

The university budget, of course, reinforces the contemplative approach. The annual equipment budget at my personal disposal is about $500.00. I would be interested in Mr. Aumen's answer to what he could do in his organization with an equivalent budget of $500.00 per year. Such a budget may be considered either a liability or an asset, the latter particularly since the half-life of equipment is also short. One of the plotters which I use in teaching, for example, costs about five dollars. It is portable, slightly fragile, and it happens to be small enough that I brought it along today. On the other hand, we do have a very excellent university computer system so that I can use my home telephone to produce contour maps on a plotter.

Before giving an outline of some of the various types of display equipment given on your list, I wish to comment on some of the parameters. Consider first the types of data with which we are concerned. The cartographic bias which colours the view of most geographers suggests a classification into points, lines and areas. This turns out to be useful for many data processing procedures. But data have many facets. They can be classified as to subject matter, for example. Others have suggested use of a topographic/thematic dichotomy. Some find it convenient to use a distinction between spatially discontinuous or continuous phenomena. In the latter case, one may distinguish between scalar fields and vector fields, both of which are simple special cases of tensor fields. Another useful point of view is to consider the type of measurement involved in obtaining the observation. This determines the class of permissible operations. Stevens and Coombs suggest that it is this set of permissible operations, what one can do with the results, that determines what kind of measurement you have. The four types of measurement usually treated are the ratio, interval, ordinal and nominal scales. For example, nominal data, which in the layman's terminology is usually considered a classification, satisfies the axioms of an equivalence sets. In fact, if you come from the physical sciences, you are familiar with the procedure known as dimensional analysis where you make sure that everything comes out right. Well, you can do the same sort of thing with these different types of measurement. I have found this a useful kind of way to look at data. There are, of course, other classifications of data, for example, deterministic or stochastic.

It may also be useful to recognize that we are discussing a logarithmic range of magnitudes. This applies to budgets which range from 10^0 to 10^7 dollars per year. The volumes of data seem to range over similar exponents, resolutions go from topographical data taken every centimeter to population by 5° quadrilaterals of latitude and longitude for the whole world. If one were attempting to make a table categorizing by system, stage of development and so forth, I think one should also have categories in terms of the exponents of the magnitudes of the data involved and the areas involved. This exponential system can, of course, be extended to other categories. For example, there are similar time ranges; contour mapping of a portion of the ocean surface from stereoscopic photographs differs from contouring of the earth's gravity field, which changes slowly. A system designed to display maps of a forest fire in real time, thus allowing optimal fire fighting procedures to be attempted, is very different from a census which involves a complete replacement of data only every decade. Somebody has suggested that political real time is about one week. This gives us another idea of the ranges involved. Clearly, the objectives of the various systems are very different. Take at random any two systems represented here. I suspect they will be trying to do different things; yet we are also involved with similar ranges, such as the sizes of maps or the sophistication of users. A related statement might be made about accuracy. I think we all agree there is no such thing as an error-free measurement. A contour line is a line only in theory; in practice it is a probability density function which can be described in approximate form by the variance (assuming Gaussian error statistics). But this PDF is shown as a fine, solid line on a map. It would be much more reasonable to show it as a fuzzy line. When I show slides of maps to my students I usually present them in defocussed form, since this is really more realistic. Perhaps more precisely, the resolution of the display device should be less than the standard error of the data, i.e., the size of the bathymetric symbol on the map is larger than the scaled standard error, at least I hope it is.

A bathymetric measurement - considering now only its planimetric position - has a very small standard error relative to the

size of the symbol as shown on the map. But if you enlarge the map up to the scale of one inch equals a millimeter on the ground, then the symbol on the map is going to be smaller than the standard error of the data. I think it is because most contour maps have the line on the maps wider than the standard error on the ground that it does not cause problems. When you get into census-type data, we make the maps a convenient size (usually page size) irrespective of data error. If I follow this consistently and make a contour map of the U.S. population from census data, maybe I ought to show that map only at a size of one centimeter by half a centimeter, but we do not like to make maps that size. However, I am sure this is a point which we can discuss at length at another time.

Much of our discussion may be taken up by consideration of display equipment, as given on today's outline. I thus will not go into any great detail.

LINE PRINTER

The line printer is the most common and graphics, of a sort, have been produced on such devices since the days of the wired accounting machine. The well known characteristics of these devices, which I prefer to think of as functionally equivalent to a typewriter, are their low cost and low resolution. It turns out that there are a lot of valid uses for these. Mr. Yoeli, for example, did shaded relief on one; after having done his overprinting, he reduced it to 1/10 of the original size so that 100 printer units were 1 inch. It was quite effective. The role of such devices changes in a time sharing environment, because the device is essentially under the control of any user with a telephone and not the computing center. Thus, the possibilities of changing symbols, or inserting coloured ribbons, or printing on a special paper, i.e., a base map, are greatly enhanced.

PLOTTERS

Plotters, the next class, are essentially printers with high resolution; these are extensions of the facsimile type, but greatly improved. Next are line drawing flat bed and drum plotters which can use ink in various colours, or scribing, or light, with which to draw lines or position symbols. Several of these are described in Appendix A. Unfortunately, this description only includes items available in the United States. I might also refer to the Journal of the Society for Information Display (28), which contains advertising describing many types of equipment, and also the July 1970 issue of Modern Data (39).

CATHODE RAY TUBE

The cathode ray tube is, of course, familiar to all because of its use in television. The characteristics of present devices include small image area, high speed, relatively high cost, high resolution, but low accuracy. Let me remind you that CRT's were originally designed as an error amplification device. Some time in the future I think we can expect solid state devices which will do essentially similar things without the drift problems, etc. Today's versions of the CRT, however, as you all know, include the vector and symbol generating tubes, gray level tubes and some colour tubes. The speed of these devices is what makes them cheap on a volume production basis. It also makes them attractive as a movie producing device, and we have done a few at the University of Michigan. I think there is a conference on computer movies coming up sometime shortly.

Let me mention two movies that we have made. One was a movie that Professor Pollock in the Geology Department did based on a magnetic tape from the Coast and Geodetic Survey of all the earthquakes, by latitude, longitude, intensity and day for a seven-year period. Then we simply merged this with a map outline of the world and displayed a little dot on the CRT for every single earthquake for a seven-year period, by latitude and longitude, and made the brightness of the dot proportional to the intensity of the earthquake; the decay characteristics of the phosphorus seem to be somewhat analogous to the decay characteristics of the aftershock of earthquakes. It made a very nice movie. It is available for purchase from the Geology Department of the University of Michigan. Let me point out what the purpose is of making a movie like this. After all, we have static maps showing the sum of all of these earthquakes and see very clearly where the earthquakes occur on the earth's surface. The purpose was really to see if we could detect any pattern in this space-time series. We had long discussions with some statisticians on how to detect patterns in multi-dimensional time series in a non-Euclidean space, and they really cannot tell us in an effective way how to analyze this pattern. So what we are saying is: "Let's show a picture of this and see if the human mind can detect any space-time patterns". I do not think we did, but that was the purpose.

I made a somewhat similar movie of the growth of the City of

Detroit, showing the population as one of these perspective block diagrams. We ran it at several speeds - once we used a time step of 0.05 years, which is a twentieth of a year or about once every two weeks, and then we ran it at a time speed of 0.5 years, which is half a year or six months; the purpose again was similar. Could we see anything that would be suggestive in a hypothesis-generating way which would lead to interesting questions, which we could investigate further? One of the problems I had was that I did not know a priori what the velocity of a city is. It turns out, in retrospect, that for showing at 16 frames per second the right speed is about one frame per month - so that it does not just sit there and you cannot see anything because it is too slow, and it does not go by so fast that you cannot see anything. But apparently a city grows, or the city of Detroit in the past fifty years, grows at a rate such that approximately one frame per month, at 16 frames per second, is the right visual speed. I had no experience in making movies or cartoons and I could not find anybody in the urban field who could tell me what the velocity of a city was. So we had to determine this empirically. A description of my Detroit movie, by the way, has been published in case you are interested in looking it up.

There is a movie language which Ken Knowlton wrote at Bell Labs about five or six years ago. He has such programming statement as Zoom, Merge and other terms which are familiar to the people in the movie industry. This technique seems to be growing very rapidly in application. A lot of people have now made movies on the cathode ray tube.

Another application of the cathode ray tube has been described in Ivan Sutherland's article in the June issue of Scientific American. One of the things he does in there is to have two very tiny cathode ray tubes, about one inch in diameter, and you wear these as glasses. You turn your head and the computer tracks where your head is and it shows you what you should see over here if you were looking at an imaginery scene. This is similar to work done at U.C.L.A. That is a very interesting kind of work. At the moment it is all experimental but it does suggest the ranges of possibility.

FACILITIES

Consideration should be made of the kind of things one actually does with these devices. The early applications were essentially a sort of mechanized draftsman, and this is still a very common usage. In fact, it is quite economical. I have used a CalComp plotter in my system and I have worked a great deal with map projections using a tape of the world coastal outline. To draw a 17 x 22 inch map of the world on some obscure projection costs about $4.00 for the plotting; it costs about $6.00 for the computing and it costs perhaps $10.00 for the setup time: you have to write out control instructions. This is quite cheap. I do not know any place you can get a draftsman to do a 17 x 22 inch map of the world on some strange projection for less than $100.00. So, in this case, it is economical. But others of you here, I am sure, can address the economics more effectively than I can because I have a very simple system: individual work. My observation has been, however, that most of the work with plotters has been attempts to replicate what we do by hand. There have been very few cases in plotting, just plain drafting, where we try to do new things. For example, I think about half the group here have programs that draw perspective diagrams of terrain, population surfaces, etc. These are very attractive and popular; we all know the variants: anaglyph, stereopairs, etc. The basic projective equations, however, have been known since about 1600 so that I do not find that being able to do this by computer is a particularly striking intellectual advance, although it is obviously of great practical importance. I personally like them, especially since they are so inexpensive when done by computer.

If you look at what has been happening, particularly in the automation of cartography, it appears that the difficult areas are such relatively mundane things as name placement. This has been one of the really hard things to automate. It is not clear even now whether it is economical to automate it, or even a saving in time. The difference between what is difficult to do by automation and what is easy to do by automation, the difference between what is easy for a person to do and what is easy for a computer to do, is quite striking in this case and it really suggests the magnitude of the difference between the two.

The next step in graphic display devices is the interactive manipulation; graphical manipulation where you do not have to know any programming language but you have to have a feel for design or art, geometry, sketching, etc. This goes beyond the simple drafting. The examples mentioned at this conference include editing of map data for input; then the more interesting circuit analysis which Mr. Calkins has described as applied to a street and bus system, where you are actually manipulating the data, not quite in a real time sense, but to solve an information kind of

problem. The typical configuration here includes graphic display, usually a CRT, some kind of graphic input device, which is perhaps a sketch-pad or light pen or Rand Tablet, usually a keyboard and, because of the data rates involved, a small computer hooking you up to a big computer. This seems to be a fairly typical configuration. It is very difficult for me to say how implemented these systems are today. I know they have one at the University of Michigan, a little PDP-8 interfaced with a big computer, and they do all sorts of things over there, but it is mostly research. It is quite clear that these systems are exciting in the long run, because you could have a display of the City of Ottawa and propose a new freeway by sketching it in on the tube and then, if one can model how things happen, letting the system evolve over the next fifty years (in 50 computer minutes) and see what occurs, and then erasing that road and putting it in someplace else and see what the difference is. This sort of thing is quite attractive. Technologically this capability has been available since about 1962, at least for small systems. In terms of economics, I would be very interested in hearing from Mr. Calkins, for example, whether the bus system in Seattle finds it economical in terms of what work they are doing.

FUTURE DEVELOPMENTS

I want to mention, before leaving equipment, two devices which I would personally like to have. One is a wrist watch longitude and latitude indicator; I am thinking particularly of the man David Bickmore referred to the other day as having muddy boots. He wants to know where he is when he records his data. It seems to me that something like that would be very useful. I say wrist watch because I do not want to have to take along a convoy of trucks to have this. It has to be quite precise because you want to know where you are within a few feet. Perhaps it will be available some day.

The second kind of device that I would like and which I cannot find on the market anywhere at a cost accessible to me is something that can take an aerial photograph, preferably in colour, which I can feed into a slot like an ozalid machine and out of the other end comes a magnetic tape. This I would also want to work backwards: feed in the magnetic tape and out comes a photograph. I would like this to cost maybe 15¢ per copy. My resolution requirements are not absurd - perhaps 1/20 inch. These are two kinds of devices I would like.

Let me go on to the ultimate use of systems like this: display systems, geographical systems. It seems to me that this is a critical thing. I want to give you an analogy which is not particularly new, but I think it is helpful to keep in mind. If you have recently been in an airplane cockpit, it is really very impressive. The thing is full of instruments and controls, and I do not see how it is physically possible to look at every instrument once in an hour. Some of these are graphical display devices. Some of these are tactical, particularly some of the controls are tactical. Some are accoustical - a buzzer goes off if there is a fire. We could go into how these instruments are ranged in a hierarchy of priorities, etc., but it is quite clear that some of these display the results of measurements on observable variables, whereas the knobs, levers and gears, etc. allow changes in the control variables. The pilot interacts between these two, taking what is, one hopes, the appropriate action. In some cases, the display devices are bypassed completely and an analogue computer of some kind is taking the appropriate response, not the pilot.

Now, I want you to shift this to think of a mayor's office, and I'll use a little science fiction. Imagine the mayor sitting in an easy chair; all around him he has display instruments and he is watching what is going on in the city. You can carry the analogy on yourself.

It seems to me this analogy suggests the whole point of geographical information systems. If we compare the mayor's office and the pilot's cockpit, there do seem to be some difficulties. For example, the aviation industry seems to have evolved to the point where they know what things they want to measure. This is continuously changing, of course. It is not clear to me that we always know what we want to measure when we sample the environment. We obviously have the same distinction between observable variables and control variables. Think of space ship earth - I would suggest that it does not have a pilot, but rather all the passengers are the ones who are attempting to produce the appropriate responses. This has an implication for the display system, namely, instead of having one cockpit with lots of dials, we have 3.5×10^9 people and we need that many dials for each variable. This is perhaps why we use graphics so much; because a lot of this audience can read only simple kinds of dials or is not accustomed to looking at various kinds of dials. It is also my suspicion that, not only do only a few people know what variables to measure, but we really do not know what the appropriate responses

are. This is, of course, the objective in many of the sciences: to find out what the appropriate responses are.

It is, furthermore, often the case that the mayor sitting in his chair is not seeing what is happening right now but what was happening a decade ago. If we imagine him steering his ship, he is making a correction for something that happened so long ago that his correction is likely to be in the wrong direction. This question comes up all the time: is air pollution, for example, causing a warming of the earth? Air pollution is apparently going up so fast that by the time we find out whether it is warming the earth, it will have done so much damage as to be irreparable. We have many problems of this type. I have made the statement before that we seem to be measuring things that were considered to be important some years ago. For example, as our theoretical understanding improves, we find that Variable A, which is now being measured, is not what we should have measured but we should have been measuring Variable B. By the time we can change one system, we will have discovered that it is not really Variable B, but Variable C we should have been measuring. This phase difference seems to be kind of endemic. I do not know how you correct for this. I would suggest that, as in Mr. Tomlinson's emphasis in his introduction to the conference, the user of the data should perhaps have a little more direct control over the collection system and over the display system and, particularly, in the flexibility with which these systems change.

REPORT OF DISCUSSION

Reported by: V. LaGarde

I. Introduction

As expressed in the index paper of today, the collective experience of the participants at this symposium is both extensive and deep in the areas of data collection, storage, retrieval and use. The lag between geographic system possibility and system implementation is as serious here as in other scientific sub-disciplines; but this meeting of such diverse interests testifies that the magic time has been reached for the beginning of co-ordinated effort. In attempting co-ordination, superficially distinct but essentially connected components within the concept of a geographic information system must be considered. A statement of some of these components has been made. These included a mention of the types of data with which we are concerned: points, lines and areas; subject matter; a topographic/thematic dichotomy; spatially discontinuous or continuous phenomena; scalar, vector and tensor fields; deterministic or stochastic, and so on. Also mentioned were the logarithmic range of magnitudes; in budgets, in volumes of data, in ground resolutions and in areal coverage; there are similar ranges of the time constants. Clearly, the objectives of the various systems are very different. A related statement might be made about accuracy.

These are only some of the considerations immediately necessary to this group in the development of display systems, and are relatively minor compared with such questions as the ultimate use of the envisioned system.

The following section discloses the participants' particular activities regarding display devices. General references to listings and/or comparisons and/or descriptions of various data output systems are contained in the References.

II. Specific Discussion

A. Line Printers

There was general agreement that line printers provide a ubiquitous and inexpensive means of performing graphic display, and that graphic methodologies so developed are easily transferable among installations. The reason given was the continuing long-standing contact between all computer users and this most primitive of graphic devices. There was also general agreement that the major constraints are print and line spacing and symbol limitations. Many expressed the opinion that this form of output is not suitable as a finished product, and is better used as a data consistency check or edit procedure in the development of the finished product.

There are some exceptions to this opinion. One such exception was presented in the form of a multi-coloured overlay map. This was achieved by producing separate line printer plots, using a photographic method to produce equivalent colour transparencies, and photographing the resultant set of transparencies overlaid. T. Waugh J. Coppock

Several others described similar output achieved using different colour carbons or ribbons on successive printer plots. One imaginative solution (70,71,72) to the line printer's

limitations is a photodensimetric study of areas, homogeneously filled with individual line printer symbols, yielding the equivalent of twenty gray levels. Since the intensity is a logarithmic function, additional sub-gray levels not in the primary set can be achieved by overprinting. The reduced size photocopy of output tends to blur the symbols and erase the inter-symbol spacing, yielding a visually acceptable map. In the example shown, the "map" is a perspective view of a terrain surface configuration model shaded according to sun shadow. P. Yoeli

Two dangers associated with using line printers for graphics are seen. The versatility and ease of using the line printer becomes a disadvantage in the temptation to submerge actual needs under output quantity. Over-selectivity can lead to redundancy and wasted effort, since there is a lag between printing and selection of printed matter for some use in many situations. The common sense of the developer seems the only answer to this dilemma. D. Bickmore D. Sinton

Other graphic non-map uses of the line printer include manuscript printing, in which the input method is by punched cards. The result of an economic study was very favourable over traditional typewriter-text editing procedures, but this automated method of editing has not spread. Aesthetics, particularly the absence of upper and lower case, seems to be the deciding factor. This cannot be corrected by using a "golf-ball" head since the head is not engineered to withstand high speed printer use. E. Amidon

The existence of two pseudo line printers, one in development and the other commercially available, was brought to the participants' attention. A cheap graphic digitizing device with magnetic tape output and an associated ink-squirt printer are in the one-year development stage by students in the High School of Technology, Lund, Sweden. The digitizer is a low resolution fiber-optic/photodiode affair. The printer ink droplet stream forced at pressure through a capillary tube, is on-off controlled by electrostatics. The droplets must pass through a mask before reaching the paper; the placing of a charge on the droplets between the capillary and mask sufficient to overcome surface tension explodes the droplets. The printer system developed by Data Corporation is essentially identical to that above. A digitizer associated with this plotter is a negative transparency drum type with resolution in the one to three mil range. Printer gray levels (128) are achieved by laying down various patterns within eight mil blocks (sub-block size is two mils). The O. Salomonsson V. LaGarde estimated printer cost is $90,000, including support facilities and software. W. Grabau

B. Analogue Devices

A question was raised concerning the existence of special analogue devices as a substitute for digital devices discussed over the past days. There was a general lack of knowledge concerning any such devices, and opinions were expressed that those using analogue devices were primarily engaged in very specialized tasks.

One inexpensive field plotter (30) allows graphic production of areal "maps" by plotting along equi-potential lines on a conductive paper. This device is commonly used for teaching purposes and appears useful in such problems as simplified population distribution, heat transfer and stress analysis. The price is in the $500.00 to $1000.00 range. Brooner

C. CRT Display

There was a consensus of opinion that CRT (black and white) display provides more flexibility and speed than line printers or drum or bed plotters. This is particularly evident in symbol placement. CRT's allow higher resolution in high event density plotting, and negate the requirement for hardware changes for resolution change. The requirement of flat surfaces as needed in drum and bed plotters to reduce distortion is not a requirement for CRT's. D. Bickmore

CRT devices are not trouble-free, as problems occur, principally in the optical system. To stress the point of resolution, easily adjustable line widths and character placement, the printed result of a 35-mm output with 64 gray levels over a 4096 by 4096 grid was shown. Simonett (now in Australia) and Kamnitz (at U.C.L.A.) are very knowledgeable in the area of colour CRT displays. Various symposium participants mentioned development stage programs they are involved with; one is using a colour display for discrimination methods. P. Kingston R. Boyle

An interesting limited-use overlay procedure using coloured areas on transparencies was proposed, which would allow very rapid calculation. D. Steiner

Three multi-colour operational display systems were Brooner

described. The first is the SDC (51) system which accepts digital data (for example, single vidicon or infra-red scanner), groups the gray scale into 16 levels, presents each level successively on a normal CRT display, and uses colour filters and multi-exposed colour film. An adverse comment is that the system often lacks calibration. A second system, IIS (47,62) allows multiple input (up to four flying spot scanners) for both black-and-white and colour CRT display of 16 gray levels. Another system, the IDECS (11-14,36,45,59,61,62), was originally developed for geoscience research and is considered state-of-the-art. In particular, multi-spectral analysis studies have a prominent place in its use.

The registration problem in multi-gun CRT tubes places a limit on resolution and accuracy. J. Sharp

D. Holograms

There was a distinct lack of information concerning the use of hologrammetric display devices. Development is in initial stages and no currently usable system is known. The feasibility of display of digitally produced terrain models is theoretically possible, but technically undesirable because of computer time involved. A calculation of the time involved on a large computer to produce a one centimeter square hologram resulted in an estimation of 16 days. Helava W. Grabau

E. Character-Symbol Generation

Again, there was agreement that the multiplicity of character and symbol sets, sizes and placement on "maps" has developed historically along aesthetic lines, and such multiplicity is not really necessary to the information content. This seems to be the major stumbling block in this area of development. For instance, the most sophisticated character symbol generation procedure allows an operator, in real-time, viewing a CRT display of a map, to light-pen input a block into which teletype-punched symbols are placed. Only a rectangular block may be specified; one cannot "follow the river". One possible answer to this general problem is to produce the most aesthetically pleasing character and symbol notation within reason, and educate the users to the cost-benefit of simplicity. J. Sharp

Two other character placement systems are very similar and consist of searching within an areal boundary to check available P. Yoeli V. LaGarde space vs. required space for symbols. One system possesses multi-level decision ability, in that if space is lacking, the software routine searches for the closest available open area for input and draws a line to denote the area marked. V. LaGarde

Various suggestions were advanced as alternatives to this problem of symbol placement, chief among which was that this could be reserved for manual insertion. D. Bickmore

The frequent problem which occurs on colour coded maps is that of "islands" with too small an area to be coded properly. Inserts are sometimes effective in this situation. D. Bickmore R. Boyle

F. Model Carving

Model carving is an effective presentation methodology for many situations, but the cost and the general state of individual system development do not warrant its implementation. The one device possessed by a participant is dedicated to production of surfaces which are to be used for plastic relief map production. The principle of operation is a computer-controlled turret lathe. W. Aumen

One ultra-low-resolution alternative advanced was the use of inflatable rubber models with skin thickness a function of slope. A second alternative of this type is the three-dimensional postcard. Although these alternative produce models of an impressionistic nature, several participants noted their wide-spread acceptance on several occasions during the past 30 years. P. Yoeli, W. Grabau, W. Aumen W. Grabau

G. Special Systems

During the course of the day, outlines of two systems of special interest were presented. The first is a sophisticated map production facility, and its description (50) contains an interesting selected bibliography. The second system demonstrates an ingenious use of simple concepts in a plotting system. J. Sharp D. Bickmore

G1. Map Production

The input system consists of a fiber-optic, photodiode drum scanner with two-mil resolution. The scanner drum is 18 x 24 inches, and the scanned plate is a negative transparency. Editing features are available with CRT display and lightpen real-time interaction. Editing includes choosing the gray tone for various features on the output product. Output is performed on a photographic negative J. Sharp

mounted on a rotating drum, and exposed by a fiber-optic pipe engineered to produce sharp edges on square two-mil spots on the photo. The definition of "sharp" in this situation is that the full rise width to maximum exposure intensity is approximately 10 microns. Shading of areas is performed by dot spacing; line width is controlled in the same fashion. Examples of output were viewed by the participants. The spot edge sharpness is of sufficient aesthetic importance to warrant the extended efforts made in this direction. A principle feature seems to be that the coarseness of Gaussian shaped two-mil spots is associated with the eye performing a centroid-search; sharp edges alleviate this problem.

The system is not performing production runs as yet.

G2. Plotter

Input is by magnetic tape of line-follower data, and plotter control is by a PDP-9 computer with 16K memory. The basic elements of the plotter are a computer controlled flat bed plotter and an optical head system. The flat bed plotter is an AEG produce with 0.01-mm return accuracy and one to seven cm/second speed. The optical head system consists of a computer-controllable disc mask (48 slots) with a light piped light source of variable intensity (tungsten bulb). The optical head is changeable with a scriber or pen head. The recording medium for the optical head is a photo negative of 1.4 square meters maximum size. Gray tones are provided by length of exposure; dot and dash patterns of adjustable length and weight are available. Varying line size is performed by shifting (automatically) the disc mask to the proper size aperture. The optical head was constructed at an estimated cost of $5,000.00. Various samples of output were displayed to the participants. D. Bickmore

This plotting system general approach appears to offer more simplicity and flexibility at moderate cost than any other system discussed thus far. V. LaGarde

H. Economics

During the course of the discussion, it was evident that system economics are of primary interest to all involved. It was equally evident that this problem of placing a price tag on system development, or even on products, is almost impossible at this stage of information systems development. The comparison of

D. Sinton,
D. Edson

systems frequently resolves into the problem of a comparison of system goals. In addition, information systems are in general meant to aid in decision processes, and one frequently cannot place values on the success or failure of the decision in question.

R. Boyle

One participant reported engaging in a cost benefit-type study of automated and manual (and various combinations thereof) information systems. The result showed that one could come to no conclusion since, in general, one could not assign even comparative values to parameters of the decision.

D. Bickmore

J. Sharp

Two participants demonstrated the ultimate vulnerability by venturing to place cost estimates on products they are associated with. The manually produced Atlas of Britain cost approximately 50 shillings per square inch, while the sample presented to participants at the symposium was estimated at 36 shillings per square inch. The ultimate cost goal for the cartographic production step of a 20 x 20 inch negative was estimated at $400.00.

B. Wellar

Although it may be too ambitious an enterprise to attempt to document fully totals, system capabilities and system costs, the users/consumers of the systems being discussed frequently have to base decisions on totally inadequate information. The participants form an appropriate group to investigate the systems under discussion and to establish an inventory of benefit/cost criteria associated with the respective systems so that the users have a better idea of what systems builders think they should look at when relating goals and systems.

III. Conclusions

The topics of day four, in fact of their immediate interrelationship with geographic information systems (GIS) development, forced a consideration of basic issues. Those most basic, and therefore most elusive considerations of (1) who needs the data, (2) why the data are needed and (3) data format presentation requirements, appeared repeatedly throughout the day. The last consideration, data display and presentation, was the primary concern and was covered in somewhat more detail than the others. In particular, various hardware types and general presentation facilities were discussed with attempts to rationalize their use on the basis of a combination of need, economics, product resolution and display presentation.

The index paper of the day aptly demonstrated the two major groups identified with GIS, those interested in methodology (the theorists) and those interested in methodology and application (the applieds). These groupings are by no means unique even within the same working group, but the vast majority of systems are obviously developed under one of these labels. A sub-grouping, to which the above comment can again be applied, is along the direction of interest. Some are interested solely in cultural features and/or information immediately related to cultural features. Examples are participants solely concerned with urban development or medical problems. Others are concerned with cultural and non-cultural information, with emphasis on the non-cultural. Examples of problems of interest to those in this set are intervisibility or engineering questions such as data for road construction.

Those within these four groups showed, at times, a remarkable similarity of approach to similar problems. While this was so, the actual implementation of systems they possessed frequently progressed along dissimilar lines, due to other major considerations.

Major conclusions were that there is insufficient information and development to define ultimate destinations of any one group, and that the goodness of a system approach depends on its applicability to the problem at hand. In addition, non-graphic output is not desirable for every information system application.

REFERENCES AND NOTES

1. ALMENDINGER, V. 1969. A geographic base file for urban data systems, Corporate Communication, System Development Corporation, 2500 Colorado Avenue, Santa Monica, Calif. 90406, U.S.A.

2. AMIDON, E.L. 1966. MIADS 2, an alphanumeric map information assembly and display system. U.S. Forestry Service Research Paper PSW-38.

3. ARMS, S. 1970. Map/Model system - System description and user's guide. Bureau of Governmental Research and Service, University of Oregon, Eugene, Ore., U.S.A.

4. BOEHM. Feb. 1967. Tabular representations of multivariate functions - with applications to topographic modelling. Proc. A.C.M. Nat. 1967, pp. 403-415 and Rand RM 5081PR Feb.1967, 58 pp.

5. BOYELL and RUSTON. Hybrid Techniques for real time radar simulation. Proc. 1963, Fall, J.C.C. Las Vegas, U.S.A.

6. BOYLE, A.R. Automation in Hydrographic Charting. The Canadian Surveyor, Vol. 24, No. 5, Dec. 1970, pp. 519-537.

7. CARR and LOPEK. Terrain qualification phase 6. Surface geometry measurement Texas Inst. Final Report Contract AF 19(628)-451, Project 762, Dec. 1962.

8. CENTRAL BOARD FOR REAL ESTATE DATA. 1970. FRIS - a spatial information system - A pilot study. Sundbyberg, Sweden.

9. COMPUTER DISPLAYS, Inc., Cambridge, Mass., U.S.A.

10. CURTIS, A.R. Characteristics of digitized line information from the d-mac pencil follower. Perception Psychologist.

11. DALKE, G.E. Automated processing of multi-spectral images. CRES Technical Report 61-16, University of Kansas. NASH Contract NSR 170004-0003.

12. DALKE, G.W. Implementation of pattern recognition techniques as applied to geo-science investigations. Computer Contribution 7, State of Kansas Geological Society.

13. DALKE, G.W. and ROUSE, J. Orbital remote-sensors - Data processing for geoscience investigation. Transactions of the American Astronautical Society, Vol. 2.

14. DALKE, G.W. et al. The study plan. Phase I; Phase II; Final. Technical Reports of multi-image Correlation Systems Study for MGI, for U.S. Army ETL, Geographic Information Branch. Contract DAAK 02-67-00435.

15. DATA CORPORATION, Dayton, Ohio, U.S.A.

16. DELAWARE COUNTY, Pa., U.S.A. 1970. Delaware County Planning Commission Information System, Delaware Co. Panning Comm., Thomas A. Curran Bldg., 2nd & Orange Streets, Media, Pa. 19063, U.S.A.

17. DIELLO, J. Investigation of digitial cartographic techniques. RADC (Rome Air Development Center), Rome, N.Y., U.S.A.

18. DOMINION BUREAU OF STATISTICS. Geocoding - Information Bulletin No. 1, Feb. 1969, Ottawa, Canada.

19. DUEKER, K. Spatial data systems I. Technical Report, Department of Geography, Northwestern University, AD652005, Clearing House for Federal Documents, Springfield, Va. 22151, U.S.A.

20. DUEKER, K. Spatial data systems II. Technical Report, Department of Geography, Northwestern University, AD652006, Clearing House for Federal Documents, Springfield, Va. 22151, U.S.A.

21. DUEKER, K. Spatial data systems III. Technical Report, Department of Geography, Northwestern University, AD652007, Clearing House for Federal Documents, Springfield, Va. 22151, U.S.A.

22. ESTES, J.E. Multi-image correlation systems for geographic information. Ph.D. Thesis, University of California, Los Angeles, U.S.A. 1969.

23. FRASER. Computer Journal 10: 165-167, May 1967.

24. FREEMAN. Techniques for digital computer analysis of chain encoded arbitrary plane curves. Proc. Nat. Electronics Conf. Chicago, Oct. 1961, pp. 424-432.

REFERENCES AND NOTES

25. GILLILAND, J.A. and TREICHEL, A. 1968. GOWN - A computer storage system for groundwater data. Canadian Journal of Earth Sciences, Vol. 5, No. 6.

26. HARVARD UNIVERSITY, Laboratory for Computer Graphics and Spatial Analysis. Selected projects 1966-1970. Cambridge, Mass. U.S.A.

27. HOUSING AND URBAN DEVELOPMENT. Urban and regional information systems. Government Printing Office, Washington, D.C., U.S.A. 1968. 441 pp.

28. INFORMATION DISPLAY: Journal of the Society for Information Display. Monthly. 647 North Sepulveda Blvd., Los Angeles, Calif. 90049, U.S.A.

29. INTERNATIONAL IMAGING SYSTEMS (IIS). Mountain View, Calif., U.S.A.

30. INTERPRETATION SYSTEMS, Inc. P.O. Box 1007, Lawrence, Kansas, U.S.A.

31. ISELI. ECAC Topographic data processing capability. Electromagnetic Compatibility Analysis Center Report. ECAC-TL-64-21. Sept. 1964.

32. JANSOON. TERROR - The influence of terrain on optical and radar visibility. Research Institute of Nat. Defence, Stockholm, Sweden.

33. JONES, K. 1970. An outline of the proposed Norwegian Geodata System. Paper prepared for 5th International Seminar on Information Systems. Czechoslovak Inst. for Regional Planning.

34. KINGSTON, P. 1968. A computer oriented geographic information retrieval system. Paper presented at the Canadian Operational Research Society Conf., Toronto, Canada. May 1968.

35. LARSEN. Data filtering applied to information storage and retrieval applications. Comm. ACM. 9, 11:785-789, 793. Nov. 1966.

36. LENT, J. An application of automated multi-channel data interpretation for wild-land resources. International Remote Sensing Inst., Sacramento, Calif., U.S.A. 1970.

REFERENCES AND NOTES

37. LEYLAND, G. 1968. Implication of the New Haven census uses test for transportation and land use planning. Highway Research Record D. 229.

38. MAERCKLEIN, L.A. 1968. The development of a State-wide mapping program and projection grid system. Program Analysis Bureau, Planning Div., N.Y. State Dept. Transportation Publication TP030001.

39. MODERN DATA MANAGEMENT. July 1970.

40. MORSE. Computer storage and analysis of contour map data. Air Force Office of Scientific Research. Report 400-106 (AFOSR 65-0592) Feb. 1965.

41. NATIONAL AERONAUTICS & SPACE ADMINISTRATION, Washington, D.C., 3rd Annual Earth Resources Program Review. To be pub. mid-Feb. 1971.

42. OEM Jan. 70. Fichier Ilot - Orientations. Institut National de la Statistique et des Etudes Economiques, Direction Régionale de Marseille et Montpellier. Observatoire Economique Mediterranéen, 10 rue Léon Paulet, Marseille, France.

43. OEM. Feb./Mar. 1970. Tables de Correspondence - Addresses - Ilots - Coordonnes Lambert. Tome I, Tome II. Institut National de la Statistique et des Etudes Economiques, Direction Régionale de Marseille et Montpellier, Observatoire Economique Mediterranéen, 10 rue Leon Paulet, Marseille, France.

44. O.R.S.T.O.M. Glossaire de Pédologie description des horizons en vue du traitement informatique. Office de la Recherche Scientifique et Technique Outre-Mer, Paris, France. 1969.

45. PETERSEN et. al. On examples of image discrimination, enhancement, combination and sampling (IDECS) in Remote Sensing in Ecology, 1969.

46. PRICE. A discussion of terrain analysis techniques. Raytheon, March 16, 1964.

47. ROSS, D. and JENSEN, R. Assorted papers referring to IIS system, circa. 1969.

48. SCHMIDT, W. 1969. The AUTOMAP System. Surveying and Mapping, 29(1).

49. SHACHAR, A. 1970. Mapping of Jerusalem by computer. Computers and Automation, May 1970.

50. SHARP, J. Development of photogrammetric data bases. (U.S. Army, Topocom, Washington, D.C.) Presented to Working Symposium of Commission II, International Congress of Photogrammetry, Munich, Germany, 1970.

51. SPATIAL DATA Corp. (S.D.C.) Santa Barbara, Calif., U.S.A.

52. STEINER, D. 1970. FORTRAN Program "GEOMAP". Department of Geography, University of Waterloo, Waterloo, Ont., Canada.

53. SUTHERLAND, I. Computer graphics. Scientific American, June 1970.

54. SYMONS, D.C. 1970. A parcel geocoding system for rural and urban information. National Capital Commission, Ottawa, Canada.

55. TOMLINSON, R. 1967. An introduction to the geographic information system of the Canada Land Inventory. Canada Dept. of Forestry & Rural Development, Ottawa.

56. TOMLINSON, R. 1970. Computer based geographical data handling methods. Presented to the World Land Use Survey Meeting, April, 1970, London, Eng.

57. TWEEDIE, R. Development of informational sub-systems with emphasis on computer graphics, Publ. No. TP00500-01, N.Y. State Dept. of Public Works, Albany, New York, U.S.A.

58. UNITED STATES (C.I.A.) Cartographic automatic mapping program (CAM). Document PB 191779, Program PB 191780, U.S. Dept. of Commerce, National Technical Information Service, Springfield, Va., U.S.A.

59. UNIVERSITY OF CALIFORNIA, Berkeley. On further comparative examples of IDECS and Philco-Ford systems in various reports by the Forestry Remote Sensing Laboratory. (Attention: Dr. G.A. Thorley).

60. UNIVERSITY OF ILLINOIS. Conference on character reading, 1968.

61. UNIVERSITY OF KANSAS. IDECS Image Discrimination, Enhancement, Combination and Sampling system. Center for Research, Engineering Science Div. (CRES), Lawrence, Kansas, 66044, U.S.A.

62. UNIVERSITY OF KANSAS. Further examples of IDECS in various reports by CRES. (Attention: Dr. R.K. Moore).

63. URBAN REGIONAL INFORMATION SYSTEMS ASSOCIATION. Annual Proceedings.

64. U.S. ARMY ELECTRONIC PROVING GROUND. Fort Huachucha, Aris., U.S.A. Investigation of model techniques. DA Project No. 3H98-00-800. Clearing House Document AD-275549, 349 pp. $21.00.

65. U.S. ARMY TOPOGRAPHIC COMMAND, Topographic Station, Washington, D.C. 20315, U.S.A.

66. U.S. GEOLOGICAL SURVEY, Topographic Division. Computer Contributions Paper No. 5. Proceedings of Symposium Map and Chart Digitization. U.S. Department of the Interior. 1969.

67. VITOSHAKIN. Theory of continuum and processing of information. Pergamon Press, New York, U.S.A. 1961.

68. WAUGH, T.C. The choropleth mapping system. Department of Geography, Edinburgh University, Edinburgh, Scotland.

69. WAUGH, T.C. and THOMAS, A.L. Geographic Information - Manipulation and Mapping System (GIMMS) a system description. Department of Geography, Edinburgh University, Edinburgh, Scotland.

70. YOELI, P. 1965. Analytical hillshading. Surveying and Mapping.

71. YOELI, P. 1966. Analytical hillshading and density. Surveying and Mapping.

72. YOELI, P. The mechanization of analytical hillshading. The Cartographic Journal, Glasgow.

LIST OF PARTICIPANTS

Mr. E.L. Amidon
Research Forester
U.S. Department of Agriculture
Forest Service (Pacific Southwest Forest & Range Experiment Station)
1960 Addison Street
P.O. Box 245
Berkeley, California 94701

Mr. Samuel Arms
Bureau of Government Research
University of Oregon
Eugene, Oregon

Mr. W.C. Aumen (92000)
U.S. Army Topographic Command
6500 Brooks Lane
Washington D.C. 20315

Mr. D. Bickmore
Experimental Cartography Unit
Natural Environment Research Council
Royal College of Art
Exhibition Road
London, S.W.7
England

Dr. A.R. Boyle
Department of Electrical Engineering
University of Saskatchewan
Saskatoon, Saskatchewan

Mr. Hugh Calkins
Urban Data Centre, Urban Systems Laboratory
University of Washington
125 More Hall
Seattle, Washington 98105

Mr. B. Farrell Chown, P.Eng.
President
Digital Methods Limited
1316 Carling Avenue
Ottawa 3, Ontario

Professor J.T. Coppock
Department of Geography
University of Edinburgh
High School Yards
Edinburgh 1
Scotland

Mr. K. Dueker
Institute of Urban and Regional Research
University of Iowa
Iowa City, Iowa 52240

Mr. D.T. Edson
Department of the Interior, U.S. Geological Survey
Topographical Division, Branch of Research and Design
1340 Old Chain Bridge Road
McLean, Virginia 22101

Mr. J.J. Foster
Director, Information Systems
Department of Regional Economic Expansion
161 Laurier Avenue West
Ottawa, Ontario

Prof. Dr. H.G. Gierloff-Emden
University of Munchen
Geogr. Institute
Luisenstrasse 37
8 - Munchen 2
West Germany

Mr. J.A. Gilliland
Department of Energy, Mines and Resources
Inland Waters Branch
Room C.42, No. 8 Temporary Building
Ottawa, Ontario.

Dr. Michel Godron
Centre d'Etudes Phytosociologiques et Ecologiques
C.N.R.S.
B.P. 1018
34 Montpellier
France

LIST OF PARTICIPANTS

Mr. W.E. Grabau
U.S. Army Engineer
Waterways Experiment Station
Vicksburg, Mississippi 39181

Mr. Max Guy
Ingenieur Geologue
Institut Francais du Petrole
B.P. 18
1 - 4 Avenue de Bois Preau
92 - Rueil Malmaison
France

Doz. Dr. R. Herrmann
Geographisches Institut
University of Wurzburg
Klinikstrasse 3 II
Wurzburg 87
West Germany

Mr. Robert E. Keith
Bureau of Governmental Research and Study
University of Oregon, Box 3177
Eugene, Oregon

Mr. P. Kingston
IBM Canada Limited
P.O. Box 517, Scotia Square
Halifax, Nova Scotia

Dr. Victor LaGarde
U.S. Army Engineers
Waterways Experiment Station
Vicksburg, Mississippi 39181

Professor Duane Marble
Department of Geography
University of Uppsala
Uppsala, Sweden
(The Transportation Center
(Northwestern University
(Evanston, Illinois 60204

Dr. R.A. McIvor
Defence Research Board
Defence Information Sciences Service
190 O'Connor Street
Ottawa 4, Ontario

LIST OF PARTICIPANTS

Mr. G.A. McKay
Canadian Meteorological Service
315 Bloor Street West
Toronto 181, Ontario

Mr. L.W. Morley
Director, Program Planning Office
Resource Satellites and Remote Airborne Sensing
Department of Energy, Mines and Resources
Ottawa, Ontario

Mr. G. Morton
Systems Dimensions Limited
770 Brookfield
Ottawa, Ontario

Dr. Azriel Rosenfeld
Computer Science Center
University of Maryland
College Park, Maryland 20742

Mr. Jean V. Salmona
Observatoire Economique Mediterraneen
I.N.S.E.E. (Min. of Finance & Economy)
9 rue Vallence
13. Marseille 8^{e}
France

Mr. Owe Salomonsson
Swedish Central Board for Real Estate Data (C.F.D.)
Box 123, S-17223
Sundbyberg, Sweden

Professor E.H. Schlosser
Associate Director
Delaware County Planning Commission
Research and Mapping Division
Court House Annex
Media, Pennsylvania 19063

Mr. A.H. Schmidt
Associate Director
Laboratory for Computer Graphics and Spatial Analysis
Harvard University
115 Memorial Hall
Cambridge, Massachussets 02138

LIST OF PARTICIPANTS

Mr. W.E. Schmidt
U.S. Central Intelligence Agency
Washington, D.C. 20505

Dr. Arie Shachar
Department of Geography
The Hebrew University
Jerusalem, Israel

(School of Architecture and Urban Planning
(University of California, Los Angeles
(Los Angeles 90024, California

Mr. John V. Sharp
U.S. Army Topographic Command
6500 Brooks Lane
Washington, D.C. 20315

Mr. R.W. Shaw
Private Consultant in Data Processing
2221 Circle Place
Ottawa 8, Ontario

Mr. Ronald L. Shelton
Center for Aerial Photographic Studies
Cornell University
453 Hollister Hall
Ithaca, New York 14850

Mr. David F. Sinton
Laboratory for Computer Graphics and Spatial Analysis
Harvard University
134 Memorial Hall
Cambridge, Massachussets 02138

Professor D. Steiner
Division of Environmental Studies
Department of Geography
University of Waterloo
Waterloo, Ontario

Dr. Carl Steinitz
Laboratory for Computer Graphics and Spatial Analysis
Harvard University
121 Memorial Hall
Cambridge, Massachussets 02138

LIST OF PARTICIPANTS

Mr. D. Symons
Chief, Computer Information Systems
National Capital Commission
299 Carling Avenue
Ottawa, Ontario

Mr. Antoine Terjanian
Geography Section
Dominion Bureau of Statistics
No. 8 Temporary Building
Carling Avenue
Ottawa, Ontario

Professor W. Tobler
Department of Geography
University of Michigan
Ann Arbor, Michigan 48104

Mr. R.F. Tomlinson
Office of the Secretariat
Cabinet Committee on Planning and Programs
Government of Nova Scotia
P.O. Box 505
Halifax, Nova Scotia

Mr. Ronald W. Tweedie
Head, Computer Section, Planning Division
New York State Department of Transportation
1220 Washington Avenue
Albany, New York

Dr. R. Van den Driessche
O.R.S.T.O.M./D.G.R.S.T.
70 Route d'Aulnay
93 - Bondy
France

Prof. Dr. B.V. Vinogradov
Laboratory on Air Space Methods in Geography of the Earth
Leningrad University
Leningrad
U.S.S.R.

LIST OF PARTICIPANTS

Mr. T. Waugh
University of Edinburgh
High School Yards
Edinburgh 1
Scotland

Mr. J.I. Weldon
Chief, General Surveys Systems
Dominion Bureau of Statistics
Room 920, General Purpose Building
Tunney's Pasture
Ottawa, Ontario

Dr. B.S. Wellar
Department of Geography
Institute for Social and Environmental Studies
University of Kansas
Lawrence, Kansas 66044

Dr. P. Yoeli
Associate Professor
Faculty of Civil Engineering Studies
Israel Institute of Technology
Technion City
Haifa
Israel

SYMPOSIUM STAFF

Symposium Secretary	Miss L. Blanchard UNESCO/IGU First Symposium on Geographical Information Systems P.O. Box 2800 Ottawa, Ontario
Technical Secretary	R.W. Shaw Multiple Access General Computer Corporation Ltd., 100 Sparks Street Ottawa, Ontario
Administrative Staff	Miss C. Beal Office of the Secretariat Cabinet Committee on Planning and Programmes Government of Nova Scotia
	Miss S. Fraser International Development Research Center Ottawa, Ontario
	Miss D. Morrell Dalhousie University Halifax, Nova Scotia
	Mr. A. Terjanian Geography Section Dominion Bureau of Statistics Ottawa, Ontario

Secretarial Assistance

Mrs. V. Amposah, Miss S. Booth, Miss D. Chenier, Mrs. K. Kennison, Mrs. J. Gleeson, Mrs. M. Irvine, Miss M. McCullen, Mrs. M. Morrison, Mrs. J. Voyce, Mrs. E. Walker, Mrs. C. Weatherdon, Miss L. Whitelocke

Acknowledgements of assistance have been made elsewhere in this work to the Chairmen, Speakers and Rapporteurs, whose efforts created a substantial part of the content of the Symposium. It is appropriate that their efforts are further noted here and that an expression of most grateful appreciation is made in these acknowledgements. Throughout the process of bringing the participants together in Ottawa, through the administration of the Symposium itself and during the essential follow-up process that has seen the initiation of the Working Group structure, I have been aided by the Symposium Staff which is listed separately in this report. It is invidious to single out particular names within a small staff when each individual contributed so much personal time and effort to the many activities included, but I would be incurring their joint displeasure if I did not extend particular thanks to the Symposium Secretary, Miss L. Blanchard, for twelve months of excellent work on behalf of both UNESCO and the IGU Commission on Geographical Data Sensing and Processing. Added to the efforts of the Symposium Staff was the major support of the staff on the Canadian Commission for UNESCO. The guidance and administrative ability of David Bartlett, Jacques-Victor Morin and Olga Jergen made the Symposium a possibility.

In the physical infrastructure of a symposium there are many elements that either have to be purchased at considerable expense, or provided through the generosity and support of individuals in institutions and companies. At this point it is proper to record the sincere appreciation of the Symposium to both the individuals concerned, and the support of:

Bell Canada
Benson & Hedges Limited
Camera House, Ottawa
Canadian Broadcasting Corporation
Chateau Laurier Hotel, Ottawa
J.M. Hill & Company Ltd., Ottawa
I.B.M. Canada Limited
Multiple Access General Computer Corporation Ltd.
The Royal Bank of Canada
Systems Dimensions Limited, Ottawa
University of Ottawa
Xerox of Canada Limited
Office of the Secretariat, Cabinet Committee on Planning and Programs, Government of Nova Scotia

The Symposium benefited from expositions of system capability arranged by John Weldon of the Dominion Bureau of Statistics,

Government of Canada, and by John Foster of the Department of Regional Economic Expansion. The work and administrative arrangements involved placed a heavy burden on both of these members of the Symposium Committee and their efforts added significantly to the successof the meeting.

In the process of editing and production of this report, I have had voluntary and extensive assistance and invaluable advice from several persons. Prominent among these are Michael Asner, Peter Jacobs and Wilma Haworth of the Office of the Secretariat, Government of Nova Scotia. In particular, I wish to acknowledge the many post-symposium discussions and contributions leading to the establishment of the Working Groups and the development of the concepts expressed in the Summary of Proceedings by Prof. Waldo Tobler, University of Michigan and member of the IGU Commission on Geographical Data Sensing and Processing, and Peter Kingston of I.B.M. Canada Ltd., Halifax. The Symposium itself was essentially a starting point, and the main work that has to be done is now resident in the Working Groups. The full membership of the Working Groups is acknowledged elsewhere in this report. Particular appreciation should, however, be made of the responsibilities accepted by the Chairman, Prof. D. Steiner, Dr. A.R. Boyle, Dr. V. LaGarde, Mr. D.T. Edson and Prof. D.F. Marble.

Finally, I wish to express my appreciation to Prof. Ray Boyle of the University of Saskatchewan, member of the Symposium Committee, for his continual support in the work of the Commission and in particular for his assistance in the entire publishing process of this volume.

IGU COMMISSION ON GEOGRAPHICAL DATA SENSING AND PROCESSING:

Chairman:	Mr. R.F. Tomlinson Office of the Secretariat Cabinet Committee on Planning Programs Government of Nova Scotia, Canada
Executive Secretary:	Professor D. Steiner Department of Geography and Planning University of Waterloo Waterloo, Ontario, Canada
	Dr. A. Shachar Department of Geography The Hebrew University of Jerusalem, Israel
	Dr. D.S. Simonett Department of Geography University of Sydney, Australia
	Professor W. Tobler Department of Geography University of Michigan Ann Arbor, Michigan, U.S.A.
	Professor B.V. Vinogradov Head, Laboratory on Air-Space Methods In Geography of the Earth Leningrad University, U.S.S.R.

SYMPOSIUM COMMITTEE:

Chairman:	Mr. R.F. Tomlinson
Members:	Professor A.R. Boyle Department of Electrical Engineering University of Saskatchewan Saskatoon, Saskatchewan
	Mr. J. Foster Director Information Systems Department of Regional Economic Expansion Government of Canada, Ottawa
	Professor D. Steiner
	Mr. J.I. Weldon Chief, General Surveys Systems Dominion Bureau of Statistics Ottawa
Technical Secretary:	Mr. R. Shaw Multiple Access General Computer Corporation Ltd. 100 Sparks Street Ottawa
Administrative Secretary: Conference Correspondence:	Miss L. Blanchard UNESCO/IGU First Symposium on Geographical Information Systems P.O. Box 2800 Ottawa, Canada