NOTICE

The quality of this microfiche is heavily dependent upon the quality of the original thesis submitted for microfilming. Every effort has been made to ensure the highest quality of reproduction possible.

If pages are missing, contact the university which granted the degree.

Some pages may have indistinct print especially if the original pages were typed with a poor typewriter ribbon or if the university sent us an inferior photocopy.

Previously copyrighted materials (journal articles, published tests, etc.) are not filmed.

Reproduction in full or in part of this film is governed by the Canadian Copyright Act, R.S.C. 1970, c. C-30. Please read the authorization forms which accompany this thesis.

THIS DISSERTATION
HAS BEEN MICROFILMED
EXACTLY AS RECEIVED

La qualité de cette microfiche dépend grandement de la qualité de la thèse soumise au microfilmage. Nous avons tout fait pour assurer une qualité supérieure de reproduction.

S'il manque des pages, veuillez communiquer avec l'université qui a conféré le grade.

La qualité d'impression de certaines pages peut laisser à désirer, surtout si les pages originales ont été dactylographiées à l'aide d'un ruban usé ou si l'université nous a fait parvenir une photocopie de qualité inférieure.

Les documents qui font déjà l'objet d'un droit d'auteur (articles de revue, examens publiés, etc.) ne sont pas microfilmés.

La reproduction, même partielle, de ce microfilm est soumise à la Loi canadienne sur le droit d'auteur, SRC 1970, c. C-30. Veuillez prendre connaissance des formules d'autorisation qui accompagnent cette thèse.

LA THÈSE A ÉTÉ
MICROFILMÉE TELLE QUE
NOUS L'AVONS RECEUE
TOPOGRAPHIC FEATURE CODING

by

Bruce D. Wood

A thesis presented to the University of Ottawa in partial fulfillment of the requirements for the degree of Master of Arts in Geography

OTTAWA, Ontario, 1984

ABSTRACT

This research has investigated in detail the scope and breadth of the topographic feature coding problem. Perspectives on the nature, issues and need for an integrated approach to feature classification are discussed. Six existing classifications were critically analysed as part of this research and several ideas were generated to serve as a guideline for the implementation of a code which could facilitate both conventional map production and geographic analysis.

RESUME

La présente recherche consiste en une étude approfondie de la portée et de l'ampleur du problème de codification des détails topographiques. L'auteur traite des perspectives sur la nature, les aspects importants et la nécessité d'une méthode intégrée de classification des détails topographiques. Il effectue également une critique de six classifications actuelles et formule plusieurs nouvelles idées pouvant orienter la mise en œuvre d'un code qui pourrait faciliter la production traditionnelle de cartes et l'analyse géographique.
ACKNOWLEDGEMENTS

A special thanks is made to Professor D. H. Douglas (Thesis Advisor), Department of Geography, University of Ottawa, for his enthusiasm, encouragement and endless helpful discussions during the preparation of this thesis. His continual support during all phases of this research is greatly appreciated.

The author would also like to acknowledge Dr. M. Allam, Dr. E. Siğişierska and Dr. J. M. Zarzycki, Surveys and Mapping Branch, Department of Energy, Mines and Resources, Ottawa; Dr. G. Decker, Geocartographics Centre, Statistics Canada, Ottawa; and Dr. B. Wellar and Dr. R. Roberge, Department of Geography, University of Ottawa, for their helpful comments and advice during this research program.

Financial support for this study was provided by assistantships and teaching responsibilities with the Department of Geography, University of Ottawa, as well as an award from the Ontario Graduate Scholarship Program in 1983–1984. Further support provided by the Department of Geography, University of Ottawa included interactive cartographic software and computing facilities.

To the numerous others who offered professional, emotional and financial support, thank you.

Finally, my son Kai contributed to the study by graciously giving up many of the hours he would have preferred to have spent being with his father. For a two year old it was a significant sacrifice.
## CONTENTS

**ABSTRACT** ................................................................. ii

**ACKNOWLEDGEMENTS** ................................................... iii

<table>
<thead>
<tr>
<th>Chapter</th>
<th>page</th>
</tr>
</thead>
<tbody>
<tr>
<td>I.</td>
<td></td>
</tr>
<tr>
<td>TOPOGRAPHIC FEATURE CODING</td>
<td>1</td>
</tr>
<tr>
<td>Objectives</td>
<td>1</td>
</tr>
<tr>
<td>Research Framework</td>
<td>2</td>
</tr>
<tr>
<td>Geography and the Nature of Computer Mapping</td>
<td>4</td>
</tr>
<tr>
<td>Problems of Spatial Data Exchange</td>
<td>7</td>
</tr>
<tr>
<td>The Growing Problem</td>
<td>10</td>
</tr>
<tr>
<td>Meeting the Challenge</td>
<td>13</td>
</tr>
<tr>
<td>The Feature Coding Problem</td>
<td>16</td>
</tr>
<tr>
<td>II.</td>
<td>18</td>
</tr>
<tr>
<td>CONCEPTS ABOUT TOPOGRAPHY</td>
<td></td>
</tr>
<tr>
<td>Principles for a Topographic Data Base</td>
<td>21</td>
</tr>
<tr>
<td>Conventional Perspective</td>
<td>21</td>
</tr>
<tr>
<td>Digital Topographic Mapping</td>
<td>23</td>
</tr>
<tr>
<td>The Topographic Map</td>
<td>25</td>
</tr>
<tr>
<td>Historical Perspective</td>
<td>26</td>
</tr>
<tr>
<td>Conventional Definitions</td>
<td>28</td>
</tr>
<tr>
<td>The Canadian Topographic Map</td>
<td>29</td>
</tr>
<tr>
<td>NTS Involvement in CIDA Work</td>
<td>30</td>
</tr>
<tr>
<td>NTS Work in Canada</td>
<td>31</td>
</tr>
<tr>
<td>Data Collation and Feature Representation</td>
<td>32</td>
</tr>
<tr>
<td>The UTM Grid</td>
<td>34</td>
</tr>
<tr>
<td>Concepts of Scale</td>
<td>35</td>
</tr>
<tr>
<td>Uses, Collation and Exchange of NTS Data</td>
<td>36</td>
</tr>
<tr>
<td>III.</td>
<td>40</td>
</tr>
<tr>
<td>TAXONOMY AND CLASSIFICATION</td>
<td></td>
</tr>
<tr>
<td>Notational Flexibility and Formal Grammars</td>
<td>43</td>
</tr>
<tr>
<td>Principles of Logical Division and Classification</td>
<td>45</td>
</tr>
<tr>
<td>Further Observations on the Classification Process</td>
<td>46</td>
</tr>
<tr>
<td>A Complex but Relevant Process</td>
<td>47</td>
</tr>
<tr>
<td>Coding Conventions</td>
<td>48</td>
</tr>
<tr>
<td>Coding Geographic Data</td>
<td>50</td>
</tr>
<tr>
<td>The Design of a Coding System</td>
<td>51</td>
</tr>
<tr>
<td>Selected Approaches to Topographic Feature Coding</td>
<td>52</td>
</tr>
<tr>
<td>Implication of Taxonomy to the Feature Coding Problem</td>
<td>55</td>
</tr>
</tbody>
</table>
IV. ANALYSIS OF SIX NATIONAL BASE MAP FEATURE CLASSIFICATIONS

The USGS Attribute Codes for Digital Line Graphs
Description and Analysis
Parameter Coding and Data Integration
Revisions to the USGS Classification Codes
Summary of the USGS Coding System
THE DMA Standard
Description and Analysis
Attribute Lists
Summary of the DMA Catalog
The Australian Standard
Description and Analysis
Summary of the Australian Code
The Ordnance Survey Code
Description and Analysis
Summary of the OS Coding System
Canadian Hydrographic Service Feature Codes
Description and Analysis
Summary of the CHS Feature Coding System
The CCSM Classification
Description and Analysis
The Importance of the Review Task
The Importance of a Uniform Classification
The Problem of Devising a Flexible Code
Retrieval and Ranging
Parameter Coding and Data Integration
The Concept of Scale Independence
Summary of The CCSM Classification

V. SUMMARY

Conceptual Strengths of a Classification System
Guidelines for Development
Inventory
Insertion
Parameter Coding
Separate Storage of Attribute Data
Ranging
Dimensional Classification
Conclusions and Implications for Further Research

BIBLIOGRAPHY
LIST OF TABLES

Table                                                                 page
1. A Classification of Features on Canada’s Topographic Maps              33
2. Some Specific Uses: the 1:50,000 Topographic Map Series               37
3. Distribution of DLG Codes by NMP Category                              60
4. A Sample of USGS Special Codes – Roads and Trails                    63
5. A Sample of USGS Parameter Coding                                     64
6. Revised Major Categories of the USGS Coding Schema                    68
7. Revised Distribution of DLG Codes by NMP Category                     69
8. A Sample of the DMA Feature Attribute/Text Lists                      75
9. Feature Classification in the Australian Standard                     80
10. An Appeal For Higher Orders of Measure                                83
11. Australian Standard – Feature Conglomeration                          84
12. A Sample of Ordnance Survey Feature Codes                            89
13. Samples From the Canadian Classification                              101
Chapter I

TOPOGRAPHIC FEATURE CODING

1.1 OBJECTIVES

The purpose of this thesis is to document in detail the scope and breadth of the topographic feature coding problem. In particular, it calls attention to the numerous serious and expensive pitfalls associated with a precipitous approach to the problem and suggests several research directions toward the development of a uniform classification which should be considered.

This research is important because the whole question of standardization in cartography is becoming increasingly important as more and more cartographic data are being transferred in digital form. Traditional cartographers have been primarily concerned with the standardization of map symbols. Standards for digital cartography require new perspectives, which in turn entail a sound philosophical and methodological approach to the problem of feature classification. This entails therefore, not just criticisms, but explanations of why and how mistakes are being made. A profound academic and research approach is different from a task force approach to the problem.

The role of a national topographic data base is important. The organization of these data and the subsequent coding schema must be flexible enough to accommodate the changing nature of both users and producers of digital map data. A topographic feature classification
could provide the basic framework for the software development required of user friendly information systems capable of both graphic representation and geographic analysis. It might well be the case however that it is quite difficult, if not impossible, to implement a classification that will meet everybody's need for there are many unresolved issues of enormous practical significance, that should be taken into consideration in order to facilitate the many complexities associated with digital mapping. The major issues addressed in this thesis concern the development of a classification system founded upon the concepts of logical consistency, terse and exact feature description, data measurement level and of course, clear scientific thinking.

1.2 RESEARCH FRAMEWORK

The Federal Government of Canada, and other local and provincial levels of government and industry, are in the process of encoding spatial data in digital form. A classification of topographic features has recently been proposed by the Canadian Council of Surveying and Mapping (CCSM) as a national standard. This standard, if adopted, will basically become the interchange linkage for all of the digital data collected by the Topographic Survey of the Surveys and Mapping Branch. Topographic Survey is, and has been in the process of digitizing terrain information since the 1960's. One of its present goals is to digitize map data of the whole country at the scale of 1:50,000. An investigation

---

of this classification is therefore extremely timely. Governments within Canada, of course, are not the only ones at, or through, the threshold of providing map data in digital form. Indeed the whole problem of exchange of spatial information in digital form between governments and between the private and public sectors is looming monstrously on the information processing horizon.

The problems of spatial information technology exchange can be classified into four levels elucidated by Witiuk as the conceptual level, the algorithmic level, the systems level and at the final results level. The topographic feature coding problem, seen in the context as occupying the lowest of the four may be misleading, in that, especially with spatial data, there are a host of conceptual problems related to data which are easy to underestimate, and easy to discount when others complain of having them. This may explain the lack of attention given to data exchange.

At a very basic level there are a host of technical standards regarding EDP mediums, formats, self defining headers and codes that must be developed and adopted, but these problems, although of concern, can be considered to lie largely outside the realm of cartography and within the fields of electronic data processing. Within the scope of cartography, the main concerns are those regarding data structure and the encoding methods for feature representation, and whether the features recorded are fundamentally the geographic phenomena themselves, or recordings of the traditional graphics utilized on paper maps to

---

represent the geography.

1.2.1 Geography and the Nature of Computer Mapping

Throughout their long and interesting history geographers have utilized the cartographic form of expression as a means of communication to provide synthesized and comprehensive perspectives of broad and complex spatial relations. The map as a traditional medium of communication is rapidly undergoing a transition that has been largely centered around the technological innovation of the computer.

Many applications and developments in computer cartography, and the important underlying theoretical research has been undertaken by members of the geographic discipline in an attempt to maintain and improve communication, both within the sub-realms of geography itself, and between various sciences and professions. New perspectives of viewing earthly phenomena through developments in remote sensing, as well as recent developments in geographic information processing, are further contributing to the fundamental measurements, with the aid of the computer, which will greatly increase the perception of the geographer of the future.

Despite the fact that their tools and means through the 1980's are continually developing, the fundamental principle of empirical observation, observation which is measurable, replicable and subsequently verifiable has remained constant. This key concept, in association with its emphasis on the spatial component, provides the basis for the continuum of scientific geographic inquiry. This is particularly true for those who use maps and as a concept is all the more applicable in the era of digital mapping.
In traditional geography the medium for the written word and/or the graphic form of expression was the paper on which the original author documented his research. The usual mode of communication was the language and/or cartographic convention the author chose to follow. In the digital era, particularly as it relates to the graphic form of representation, many of the commonly accepted standards of communication have undergone rapid transition.

Much cartographic innovation evolves directly from the requirements and possibilities of the digital data base. Tobler discusses several of these aspects by stressing the theoretical relationships cartography holds with mathematical geography, geodesy, photogrammetry, remote sensing, computer graphics, geographical matrices, fourier transforms, quantization, coding, map generalization, pattern recognition, generalized spatial partitionings, geographic coding and conversions, map projections and geographic information systems. A basic starting point in realizing these requirements and possibilities is a clear comprehension of the distinction between two conceptual types of maps: the real map (ie. the conventional paper map) and its digital counter-part (the virtual map).

---


Traditional cartographic principles are of course requisite for cartographic training but we must be aware of the questions being raised regarding the fact that a student seeking training in cartography is opting for early obsolescence if he or she does not learn the basics of programming, image manipulation, digitization, cartographic data structures and the available production options made possible by computer assisted cartography. Further, the need for a total integration of cartography within all of geography's topical fields of interest is founded on the fact that "in the hands of a well-trained geographer, cartographic thinking assisted by modern remote sensing and interactive computing methods is without equal as a creative, productive means for geographical analysis and interpretation." The methodological tool box recommended by Tomlinson certainly underscores the importance of developments in computer mapping to the geographic discipline for the measurement, modelling and planning of man's environment.

---


1.2.2 Problems of Spatial Data Exchange

The problems of exchanging spatial data have come to the forefront of spatial information processing only recently, however in the broadest sense there has been a need and an increasing attempt to exchange spatial data for over twenty years...by Robert Sparks, Albert Guttenberg, James Anderson, and Clawson and Stewart, to name but a few.' These early commentators suggested that a uniform land use classification would be an important step in the right direction if the planning profession is to keep pace with population growth and technological development. Two major handicaps they identified were: 1) the incompatible classifications and differing definitions of terms; and 2) the lack of measurement. Further they noted that progress in planning and planning research depends so much on an exchange of empirical findings, it is especially important for planners to have a common language, a language so precise as to leave little room for misunderstanding and that the basic problem in the entire land use field is one of developing a classification of land uses, utilizing definitive criteria for separating the various classes, that will give objective and repeatable results for any area being studied.' Anderson drew particular attention to the need for objectivity in the identification,


classification and mapping of idle land suggesting that sub-categories should be established to best delineate abandoned or idle land in terms of an area's last, or most obvious economic use (e.g. idle crop land, idle pasture land, idle forest land, idle urban land etc.).¹¹

A major work on standardization by Clawson and Stewart found that accurate, meaningful, current data on land use are essential if public agencies and private organizations are to know what is happening, and are to make sound plans for their own future action.¹² Some very pertinent issues they emphasized included the need for: 1) a continuing organization of specialists representing most if not every group concerned with data about land; 2) a comprehensive system based on logical concepts, each on a pure line basis; 3) maximum detail in the basic enumeration stage, with groupings and summaries only at a later stage; 4) flexibility such that data can be used in great detail or summarized into broad groups as desired, and flexibility such that either different groupings of uses, or different groupings of areas can be used for summarization; 5) the availability (and future availability of comparable data for earlier periods) to anyone who needs it; and 6) a minimum of duplication between sources of data, especially duplication in the expensive data-collection and initial statistical processing.¹³


¹² Marion Clawson with Charles L. Stewart, 1965.

¹³ Pure series data are required to avoid ideas of sharply different characteristics (dimensions) being intermingled in data series. It was one of the most significant conclusions of the Standard Land Use Coding Manual: A Standard System for Identifying and Coding Land Use Activities, Urban Renewal Administration, Housing and Home Finance Agency and Bureau of Public Roads, Department of Commerce. (Washington, D.C.: U.S. Government Printing Office, 1965).
Clawson and Stewart also remind us to be realistic about the difficulty of implementing a single general purpose national classification system suggesting that perhaps a second system, such as one based on longitude and latitude, could be accepted by everyone as the second system, to which they would be willing to relate all data enumerated or assembled by them. Further, a basic thesis they brought out at this early date was that to overcome comparability problems over time and space, data on land could be transferred electronically to another point of use very quickly. "Thus it is no longer necessary that each point of data use have its own storehouse of land use data in the form of tabulations, publications, maps etc., but rather, data obtained in one place or by one agency could be stored in a selected place, ready for use by anyone quickly and inexpensively."  

14 Is this not the basis of spatial data transfer in the digital era?

The principles underlying these ideas, although generated in the late 1950's and 1960's are applicable to the requirements that should be facilitated in the collation of a national digital topographic data base in the 1980's. The basic idea is to collect and encode data in a format that facilitates their use beyond the immediate needs of the collecting agency.

In 1970, at the opening address of what is often referred to as the 'First Auto Carto', Tomlinson advocated that "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

14 Marion Clawson with Charles Stewart, 1965, p. 54.
development program to ensure that systems to handle the data will be available." He further stated that 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." In the 1980's a proposed Canadian national standard for the exchange of digital topographic data is being implemented largely under the guise of a bureaucratic task force... the data collectors.14

1.2.3 The Growing Problem

The problems associated with, and the importance of improving our ability for spatial data transfer are accentuated by the fact that more and more government agencies, universities and private industry are making use of the computer for storage, retrieval and analysis of spatial data.17 Often however the unit of analysis has been collected without careful thought as to what other uses the data might be put. Several diseconomies associated with this phenomena include the simple


(but expensive) fact that: 1) often spatial data are not readily
interchangeable; 2) resources are lost in the duplication of data
collection and/or the extended efforts required to transform, replot or
regenerate the data in order to use them; and 3) any real benefits
derived from using the computer (such as timeliness, efficiency and/or
replicability) are lost to the costs associated with outdated,
incomplete, and inaccessible data.\footnote{\label{foot21} Yan, Witiuk and Fisher, 1977; Interministerial Committee on
Geographical Referencing, 1978.}

Many examples of the difficult problems surrounding the efficient
use of digital spatial data are presently being faced by users of the
new technology. For example, Tripe delineates some important issues
concerning the Canadian Hydrographic Service's acquisition of a digital
data base: "Which data are to be stored and in what form? How does one
access data and who should access the data? What quality control
maintenance is required? What form will the output take, digital,
graphic or both? How large is the data base likely to be?"\footnote{\label{foot22} R. L. K. Tripe, "A Critical Review of Automated Hydrography Within
the Canadian Hydrographic Service," Proceedings, Conference of
Commonwealth Surveyors, Overseas Development Administration, Part II,
Paper J1 (1979) pp. 1-10.}

Similar research described by Puderer concerned the interface of
the postal code (six digit alphanumeric code in operation since the
early 1970's) with the area master file (line segment file similar to
the DIME data structure based on block faces).\footnote{\label{foot23} Henry A. Puderer, "Census Geography Staff's Planned Postal Code
Master File - Area Master File Linkage." Paper presented at the
Canadian Association of Geographers Annual Meeting, (University of
Ottawa, June 1982).}
Canada to improve both automated geographic coding and statistical analysis. Motivating factors stimulating this interface included the need for site specific data as well as to improve the transferability of the data base with other geographic information systems.

An important argument presented by Harris contends that topographic mapping must provide the digital base to avoid a wasteful and confusing situation where thematic mappers make their own data base.\textsuperscript{21} Avoidance of this situation depends on the actual costs involved in extracting data from Topographic Survey Departments. Some have made the analogy to pulling teeth. In order to efficiently utilize this topographic data base, and avoid the type of diseconomies noted above, the digital topographic data base must be structured so as to allow the transfer to the widest array of users. \textsuperscript{22} In the United States, one example cited by Marble concerning natural resources and the problems of exchanging spatial data is that of a hydrologist developing an integrated model showing the relationship between land use and run off in a river basin. Digital data sets exist for all the ecosystem components that the analyst is interested in. "But try to combine them. Try to take the digital information on land use, water flow, stream quality and precipitation from different agencies, collected from different bases, with different geographic coding schemes, all of which are basically incompatible in one form or another, and try and put them


\textsuperscript{22} This returns us to some of the questions pointed out by R. L. K. Tripe, \textit{Proceedings}. Paper JI, 1979.
1.2.4 Meeting the Challenge

Problems of exchange are increasingly being met by governmental committees established for the purpose. One of the earliest experiments in Canada was the formation of the Inter-Agency Spatial Data Transfer Committee which grew out of the less formal National Capital Geographic Information Processing Group of 1975. This committee, composed of officials from various of the pioneering agencies in geographic information processing, i.e. the Canada Centre for Remote Sensing, the Geocartographics Group at Statistics Canada, the Lands Directorate at Environment Canada, and the Systems and Consulting Group of Agriculture Canada, produced a publication entitled "Standard Format for the Transfer of Geocoded Polygon Data" in 1979. This publication suggested formats and data structures for the encoding of topological files representing polygon network segments, or chains. Overall the proposed standardized format attempts to meet the criteria of machine and language independence, expansibility, generality and self-definition (i.e. files should be self-defining with respect to spatial and descriptive data; e.g. attribute names, map projection, ground resolution).

---


Similar problems of exchange were recognized in Ontario where of sixty-four geographical reference systems in the Government..."it would be difficult if not impossible to transfer data between these systems without major modifications to most of them." One objective of the Committee on Geographical Referencing was to develop standards and specifications which could be incorporated into municipal, provincial and federal information systems in order to expedite the transfer and correlation of geographically referenced data with one of the prime concerns being the establishment of an acceptable grid reference.

The Alberta Government, wishing to avoid the chaos of multiple development of incompatible systems endured by earlier entrants to the field set up the Land Related Information Processing Systems (LRIS) Coordination Project in 1978. This was a committee of six senior managers from the key ministries that had direct interests in land related information. Cooke estimated that by 1981 there were twenty-six operating automated mapping systems within the Government of Alberta with varying degrees of incompatibility between them and that the LRIS Coordination Project was dealing with that with a top down conceptual approach.14

Similar concerns expressed by the (Canadian) Task Force on National Surveys and Mapping (1978) revealed that a basic need at this time is the design and adoption of criteria for the storage,


accessibility and exchange of digital spatial data among federal and provincial departments and agencies, utilities and the private sector. The Task Force noted that the main thrust in research and development is being applied to digital mapping and the application of new technology to national mapping with the most pressing need at present being the formulation and adoption of standards for the communication of data. Five such issues, centered on Topographic Survey were identified: 1) completion and upgrading of the basic National Topographic Survey coverage of Canada; 2) relationship between federal and provincial mapping programs and related activities; 3) relationships with other federal departments and agencies; 4) relationships with industry; and 5) impact of new technology on topographic mapping.

Of course other countries are also facing the same standardization issues and problems in similar ways. The ACSM National Committee for Digital Cartographic Data Standards is a case in point. All of these attempts to solve the data transfer problem by committees illustrate the importance of the subject, and indicate a measure of the intractability of the problems being faced.

---


1.3 The Feature Coding Problem

There are many aspects to the handling and transfer of spatial data. As noted, this thesis attempts only to provide some definitive ideas on how the feature coding problem might be resolved and confines itself to the problems associated with the: 1) development of a classification system of 'real-world features' which is acceptable as a national standard; 2) the logical organization and codification of this information to facilitate its optimal usage in a user-friendly manner; and 3) the provision of further research guidelines as to how much of 'what exists' can we reasonably expect to collate and codify.

The digital description of a cartographic or geographic feature can be considered as having two components, a geometric description and an attribute description. The geometric description has been given rather more attention. Methods for numerically recording points, lines and surfaces have been developed to a high degree of sophistication; topological structuring of data to represent networks, graphs, solids, etc., to enable more efficient searches and manipulations, to compute properties from the geometry and to generate interesting graphics are also well represented in the literature. Geometric description problems seem to offer more interesting challenges. This may be because of the multi-dimensional aspects of geometric data, but also because when a technician or researcher is working with a data-set it often is of one kind of feature. It is a file of contours, or a digital elevation model of temperature, etc., and this fact reduces the importance of its identification since its description will be understood or may be noted with the written description of the file. Of course the importance of
self-describing data grows when one considers mass exchanges amongst
governments and between the private and public sectors and the prospects
of automatic exchanges between machinery.

Problems of devising self-describing codes for objects and their
non-spatial attributes at first seem straightforward and perhaps rather
obvious. If road maps are composed of subsets from fifty-seven different
symbols, simply create a table of fifty-seven items, number them and
include an integer from 1 to 57 with each geometric feature in the
digital file. Sort them alphabetically before numbering them if it is
expected that the list will be examined frequently by eye. Even the
addition of that 58th feature need not cause excessive concern. Add it
on the end, or assign to it a decimal number to insert it into its
proper alphabetical place, or redraft the code, attaching a flag to the
data identifying the specific code list to employ. In fact, do not even
use a numerical code, but rather let the alphabetic words themselves or
mnemonic short forms act as the code. All of these solutions are indeed
employed, but it is apparent that even at this level the problem is
losing its trivial innocence. Contrasting with a road map series, a
topographic series may have from 1000 to 2000 different types of
features. The mere compilation of the list of things appearing on a
whole series for a country the size of Canada is no trivial task.
Chapter II

CONCEPTS ABOUT TOPOGRAPHY

Adequate and up-to-date coverage of topographic maps is a primary requirement of a country to permit efficient administration and to aid in its development. Further, the fundamental data needed to tie resource information systems together are found on topographic maps and it is these data that provide the common reference and foundation upon which all systems must be built to adequately serve today's resource managers. Ottason and Rystedt also note that the topographic map often provides the background for thematic mapping of which the aim is to illustrate geographical aspects of particular themes or topics and the main reasons for introducing computer-aided techniques in small scale topographic mapping at the Land Survey in Sweden is because map consumers need landscape information not only in the shape of maps but also in digital form in order to take advantage of electronic data processing in different planning and computing activities. To further emphasize its importance it should be noted that the topographical map, generally, provides the foundation or the source for the production of

---

27 Harris, 1980, p. 61.


the foundation for the following types of thematic and special maps and overlays: navigational charts (such as aeronautical and hydrographic charts); town and estate plans; route maps (such as road, street, traffic flow and route accessibility maps); atlases; world maps; physical, geological, and land use maps; population and political maps; maps to depict medical statistics and all the statistics of the environment which have a geographical connotation; maps for land registration; and maps for administration and development....at the large scales, topographic maps are used for engineering, administration, and ownership registration. At the smaller scales the basic topographic information can be related to other resource fields such as soils, geology, hydrology and demography.

As in traditional cartography, the topographic map forms the basic building block for thematic and special purpose maps. One contention is that the general purpose topographic feature classification adopted will have major implications on the level of geospatial analysis permitted with the encoded data in the future. This contention stems from three facts: 1) automated cartography is not limited to the automated drafting of maps but extends into the area of cartographic data services from which a wide variety of graphic products, both stored cartographic information as well as derived information, can be produced (displayed) on demand; 2) the topographic map, either directly or after scale and projection changes, provides the foundation for many of the published maps used increasingly by the many

---

32 Harris, 1980, pp. 60-61.

33 Digital Steering Committee, 1980, p. 2.
sciences and professions who demand terrain information; and 3) any feature coding standard, if it is done well, may provide the foundation of other systems created for and by specific users, whether or not they utilize data from a Governmental Survey. This may create opportunities for transfers from non-governmental agencies, such as utility organizations, to the government for generalization to survey needs.

Linders indicates that the development of a system for mass exchanges implies a taxonomy or classification of information with the objective of providing a single logical system for the storage and management of all land mass data. The conceptual framework (model) proposed by Linders suggested that such a system must incorporate detailed information for: 1) feature location; 2) feature taxonomy (the component elements of the data base must be uniquely classified through a data dictionary); 3) attribute data (for further differentiation of features); 4) relational data; and 5) representational information (the exact depiction of each feature element in terms of its graphic components). He suggests that features of the same type should be stored as classes within the file system such that one could, for example, select particular elements of a data base for information retrieval (e.g. NTS, class of feature, specific feature). Rhind adds that in order to be able to handle all the possible retrieval questions based on relationships between different objects in the data base,


topological, geographic, and attribute directories are all necessary.³⁴

2.1 PRINCIPLES FOR A TOPOGRAPHIC DATA BASE

Although the actual digital data base is not the product that many large Governmental cartographic agencies are mandated to produce, the fact that we can utilize digital topographic data for many purposes and kinds of maps should be re-emphasized. The optimal use of the resources invested in the collection of topographic data would be to collate digital data (put less emphasis on paper maps) with the broadest concept of spatial data transfer in mind (by implementing an exemplary standard necessary for the maximum amount of benefit to be acquired through topographic data collection). Collated digital topographic data, compatible in many more ways than is presently being realized (a user should not have to digitize off a NTS paper map if that map has been produced from a digital data base), should facilitate the production of all kinds of maps, be they real or virtual.

2.1.1 Conventional Perspective

From a conventional point of view, much work has gone into the study of standards for symbol representation. Komkov noted that a comparative study of a number of English, American, French, German, Italian, Russian and other general maps shows that the classification of the symbols for the representation of the main details of landscape, relief, water features, forests, populated places, roads and boundaries

coincide on many points. This compatibility between maps of various countries has in large part been facilitated by adoption of standards such as the Greenwich prime meridian, the metric system, geodetic datums, projection system, scales, feature classification, methods of feature representation and proposals for standards (such as the specifications for topographic mapping, recommended for international use proposed by the Institut geographicque national de France at the second U.N. regional cartographic conference for Asia and the far east). Also, because the rules relating geographical concepts and map signs are best understood in topographic mapping, standardization has progressed further and is more acceptable in topographic cartography. Further, topographic maps are especially fit for international standardization because the universal principle of the topographic map as representative of the exact localization and the correct delimitation of areal subjects is accepted.


2.1.2 Digital Topographic Mapping

Two contentions of this research are that: 1) the progressive development of automated cartography requires that Topographic Mapping Departments record the finest elements of information technically feasible in a format that is readily transferable, retrievable and usable; and 2) "only if the best sides of the national cartography of individual countries, concerning the fields of general geographic (topographic), special and thematic mapping, are thoroughly studied, generalized and borrowed will cartographic standardization prove a success."\(^1\) It should also be re-emphasized that "geographical or cartographic data collection is often so expensive that it is justifiable only where multiple uses are possible, irrespective of whether the data are used to produce graphics or non-graphic answers."\(^2\)

Furthermore, "in the traditional concept, the original graphic manuscript and the reproduction materials derived from it are the media containing the accuracy of terrain data. In the digital era, however, the accuracy is contained in the digital topographical data,... the degree of their usefulness depending, to a large extent on the sophistication of the data and file structure."\(^3\)

\(^1\) Komkov, pp. 210-213. Although Komkov's research largely discusses attempts at standardization in the conventional sense, a valid scientific principle is expressed despite the change in technology.

\(^2\) Zarzycki, Harris and Linders, 1975, quoted in Rhind, 1976, pp. 515-516.

The need for the exchange of digital data between various governmental, private and public organizations will continue to increase as more and more of these agencies utilize the computer in their everyday operations. A National Feature Classification Standard should facilitate and act as a framework on which the features and the forefront of research available from the other disciplines in digital form can be integrated (documentation should also be supplied that explicitly states the limitations imposed on the classification which could affect its utility in future use). Linders suggests that a model "classification must be systematic (orderly), hierarchial (reflect the natural association between elements), consistent (without ambiguity), flexible (able to add), efficient (convey a maximum amount of information with a minimum amount of coding), and readily amenable to computer processing."44 Regarding efficiencies in the actual coding, we must be concerned with real, and not false efficiencies. Another byte is a small price to pay for just one more advantage, but if it means that to conserve that byte we must abandon a logical system, that effort is utterly pointless. Most importantly, "the dangers inherent in a precipitate approach to a digital system cannot be over-emphasized, and this is particularly true under the constraint of a general purpose national system."45


2.2 THE TOPOGRAPHIC MAP

The information in a topographic map contains the answer to two questions: where (location) and what (all other information about features). Through the integration of technology with cartographic science, the topographic map, its detail of presentation, and its level of cartographic craftsmanship has achieved high standards of excellence. There has been some question however concerning the information content, and some have noted that this factor, in specific situations, has reduced the subsequent utility of the topographic map in its conventional representation. Keates observes that, for topographic maps, divisions at ground level provide the most evident features, in particular the obstacles to movement and the positions of recognizable entities. He makes the point however, that far from being a picture of the earth, the map is a highly abstracted selection, and if the user is not aware of the implications of this he is likely to fail to deduct correctly the information presented, or be vaguely dissatisfied because he is expecting the map to present other information. Developments in remote sensing technology have led others to comment that in parts of the world with dramatic seasonal differences, especially in areas with a minimum amount of man-made landscape, it is very often pointless to use a conventional topographic map which invariably describes terrain in its


Keates, p. 169.
summer appearance.\textsuperscript{44}

2.2.1 Historical Perspective

In accord with the original meaning of the word 'topography', we should be attempting to portray on the topographic map, those visible natural and manmade geographic (topographic) features which are defined or can be referred to the earth's surface. Some literature suggests however that the word topography, as with much scientific terminology, has been subject to misuse and improper application. Misuse of any concept leads to misunderstanding and it is for this reason that part of this section treats and elaborates on the definition of the topographic concept.

Early work by Campbell and James states that the common misuse of the word 'topography' to refer only to surface features such as relief, landform, and terrain has made the recapture of the original meaning difficult.\textsuperscript{47} In fact Campbell noted that at least 90\% of the persons using the term topography, did so incorrectly.\textsuperscript{50} The extent of this misuse was great enough that it stimulated much discussion concerning the need for a new word to conserve and portray the original meaning of the word topography.\textsuperscript{51} These discussions led to one suggestion that the


\textsuperscript{50} Campbell, 1928, p. 37.

\textsuperscript{51} Preston E. James, "The Terminology of Regional Description," Annals
topographic map, in association with the original meaning of the word topography, refer to a map of relatively large scale which permits the plotting of the smallest relevant area units of human occupancy: such units as soil types, individual fields on a farm; individual clearings in a forest; or in urban studies, the components of the commercial core. The use of the adjective topography, as cited by James and Jones, to describe "those studies carried out on scales large enough to permit the mapping of the specific features of human occupancy" differs greatly from a strictly hypsometric conceptualization of topographic features.

In his comparison of several European topographic maps Piket suggested that a topographic data classification should be designed and produced on a landscape and international basis, rather than by individual country or a national basis. The difficulties in coming to an agreement about what the contents of topographic maps ought to be is an old issue, but one which merits extensive attention as more and more large scale cartographic data are collated for machine processing.


53 Preston E. James and Clarence F. Jones, eds., American Geography, Inventory and Prospect, Published for the Association of American Geographers (Syracuse: Syracuse University Press; 1954).

54 J. J. C. Piket, "Five European Topographic Maps: A Contribution to the Classification of Topographic Maps and Their Relation to Other Map Types," Geografisch Tijdschrift 6, no. 3 (1972):270.
2.2.2 Conventional Definitions

The root of the term topography is the Greek word "topos", meaning place. Literally then, any map representing place is a topographic map. As defined in the Oxford English Dictionary topography refers to the science or practice of describing a particular place, city, town, manor, parish, or tract of land; the accurate and detailed delineation and description of any locality. Similarly, topography is defined in Websters Dictionary as the description of a particular place (as a city, town, manor, parish, or tract of land); and the art or practice of graphic delineation in detail usually on maps or charts of selected natural and man-made features of a place or region especially in a way to show their relative positions and elevations.

There are many kinds of maps and one way of classifying them, as in topographic maps, is on the basis of scale. As defined by the International Cartographic Association a topographic map is "a map whose principal purpose is to portray and identify the features of the earth's surface as faithfully as possible within the limitations imposed by scale." Harris suggests that in some usages, the 'topographic map' has

53 Campbell (1928):37. Piket (1972) notes that the "term topography comes from the Greek geographers, who used the term to describe areas of limited extent, placing it at the same level (ie. next to) as the term chorography (the description of regions), and next to geography (the description of vast areas and of the entire earth)."


57 P. B. Gove et. al., Websters Third New International Dictionary of the English Language Unabridged, (G. + C. Merriam Co. Mitchell, 1971), p. 2411. Also defined under topography in Websters Dictionary is: "the physical or natural features of an object or entity and their structural relationships < statistics which reveal the economic topography of our time - R. D. Mack >."
conveniently been used to refer to certain scales of maps, say, between the scales of 1:10,000 and 1:250,000, the scales larger being called the engineering and cadastral scales and the scales smaller being called the base maps for thematic maps. In the practice of applying computer technology to cartography, he suggests that a topographical map should be taken to mean a cartographic description of a place or environment in terms of the visible features to the extent that the scale of the map permits."

2.3 THE CANADIAN TOPOGRAPHIC MAP

A prerequisite in the analysis and evaluation of the proposed CCSM topographic feature classification involves a good understanding of the topographic map, its design process, uses, and specifications. Further, in designing any classification system one must, at the outset, delineate not only the present users, requirements and objectives but the future potential ones also.

"The Topographical Survey Directorate of the Surveys and Mapping Branch of the Department of Energy, Mines and Resources is the organization responsible for national mapping in accordance with terms of the Resources and Technical Surveys Act, 1966-67, c.R-7, R.S.C. as amended by the Government Organization Act, 1970, c.14 (2nd Supp.), s. 9, R.S.C., and as implemented by the EMR Earth Sciences Services Program." The objective of a topographical service activity is to

---

Footnotes:


*** Harris, 1980, p. 60.
"ensure the availability of information concerning the topographical features of the Canadian landmass and their representation to defined standards." These responsibilities include new topographical mapping, map revision, and maintenance of both the aerial survey data bank and the national topographic data bank (to ensure the availability of topographic data in digital form)."

Aside from the former objective, about 5% of Topographic Survey's resources are accounted for by providing 'technical mapping support, not directly related to NTS mapping,' to: 1) other governmental departments and agencies (such as the Department of Public Works, Department of Indian and Northern Affairs, National Capital Commission, and the Canadian International Development Agency); and 2) to ensure adherence to standards of all aerial photography taken for the Federal Government. The Canadian International Development Agency (CIDA) is cited as being the agency which derives the major amount of support service and a small sub-unit in Topographic Survey has been formed for this purpose alone."

2.3.1 NTS Involvement in CIDA Work

Topographic Survey involvement in CIDA work has been in progress since the late 1950's, and has involved approximately 20-25 projects totalling over $24 million worth of mapping, mostly at the 1:50,000 scale in such countries as Nigeria, Tanzania, Ghana, Kenya, Guyana, Trinidad and other Caribbean countries. "Approximately half of the amount accounts for projects still in progress (as of the late 1970's),


idem, 1978, pp. 44, 59 and 73.
and the demand for support is expected to continue for several years. The individual project sizes are becoming larger, for example: the impending work in Indonesia could run to $15 million for the data base alone, and another $14 million if map compilation and drafting is added. In 1981-82, for example, technical advice and inspection services were provided on Surveying and Mapping projects under the auspices of CIDA in Nigeria, Tanzania, Indonesia, Zimbabwe and the Barbados. Although only accounting for a small percentage of their resources, the transfer of high technology to the third world is important. Amongst other areas of transfer, the clear and logical design of a general purpose topographic feature classification for the handling and transfer of digital spatial data is required if it is to be capable, both at present and in the future, of practical application, or at least to serve as a model, for the developing countries.

2.3.2 MTS Work in Canada

In Canada the three basic MTS scales are the 1:250,000 series, the 1:50,000 series and the 1:25,000 series. The 1:25,000 scale series have since been suspended and no maps at this scale were published in 1981-1982. In addition to these series are produced topographical maps at the scales of 1:125,000, 1:500,000 and 1:1,000,000. The 1:500,000 and 1:1,000,000 are produced largely as base maps for aeronautical charts, the latter being produced as part of a world-wide project supervised by

\[\text{Task Force Report, p. 73; and Energy, Mines and Resources Canada, "Surveys and Mapping 1981-82," Minister of Supply and Services Ottawa, 1982; p. 5.}\]

2.3.3 Data Collation and Feature Representation

Although the technology associated with topographic map construction is becoming increasingly more sophisticated, the underlying principles of data collation, based on aerial photography of the area to be represented, the photogrammetric measurement and drawing of the map detail based on the air photos and the integration of additional information by field survey, with the exception of the amount of information which could be collected for a digital data base, remains fundamentally the same. In supplying this additional information, the field surveyor: 1) determines the exact geographical location and height above sea level of a network of features clearly defined on the airphotos; 2) classifies the road and track surfaces and locates schools, churches and other buildings that are shown on the map by special symbols; and 3) checks the names of places and features and completes any detail that can not be determined from the airphoto.++


### TABLE 1

A Classification of Features on Canada's Topographic Maps

<table>
<thead>
<tr>
<th>Main Division</th>
<th>Examples of Features Included in Each Division</th>
</tr>
</thead>
<tbody>
<tr>
<td>Culture</td>
<td>cities, villages, buildings, railways, highways, land boundaries.</td>
</tr>
<tr>
<td>Water</td>
<td>sea, lakes, rivers, streams, ponds, marshes, swamps, glaciers, snowfields.</td>
</tr>
<tr>
<td>Relief</td>
<td>mountains, hills, valleys, cliffs, slopes, depths.</td>
</tr>
<tr>
<td>Vegetation</td>
<td>wood areas, orchards, vineyards, cleared areas.</td>
</tr>
</tbody>
</table>

**SOURCE:** Adapted from Sebert, 1972, p. 1.

Table 1 provides several examples of features found on Canada's topographic maps and shows how these may be classified into four main divisions. On the 1:50,000 Canadian topographic map series approximately 564 features have been identified by committee no. 1 on National Digital Mapping Standards (1982). In attempting to establish a classification system that was inclusive for all possible topographic features, they approached various external agencies such as the National Capital Commission (NCC), Environment Canada, provincial agencies and utilities and compiled approximately 1558 feature-attribute combinations for the classification.**

**" Sebert, *Every Square Inch*, 1972, p. 1.**

**" CCSN, *Standards, Report of Committee I*, April, 1982."**
2.3.4 The UTM Grid

One of the most important attributes of the topographic map is the use of choice projections, typically orthomorphic (conformal). Canada has adopted the six degree Universal Transverse Mercator (6 degree UTM) projection for its topographic maps. Its adoption was not a straightforward process however for in the 1960's, the Federal Government (ie. Surveys and Mapping) were considering a system based on the Transverse Mercator Projection modified to a three-degree zone (3 degree MTM). This system was also considered by the provinces and while some adopted it, others (and the Federal Government) did not. In the case of Ontario, a consensus of government departments was reached and in 1969, the Surveys Act was amended to provide for systems of coordinate surveys according to the parameters and characteristics of the 3 degree MTM. In 1972 however the actual implementation of the 3 degree MTM was turned down. An interesting point is the fact that

The UTM projection is a conformal projection. The true shape of the features represented within any small area of the map is thus maintained and within a given area the scale of the map is constant in all directions.

When the globe is divided into 60 zones each being 6 degrees longitude 16 out of 60 zones bearing numbers 7-22 cover Canada. Each zone is 100,000 metres in width divided equally by the central meridian. The East boundary (in Canada only) is 0 metres; the West boundary 100,000 metres. Each zone (divided by horizontal and vertical grid lines) is sectioned into basic mapsheets. The horizontal grid lines are associated with the distance from the equator, the southern most part of Canada being almost 4,620 kilometres from the equator.

Ontario Ministry of Natural Resources, The Ontario Geographical Referencing Grid (The UTM Grid System), Surveys and Mapping Branch, Geographical Referencing Section, May 1981, p. 21. It is worthy to mention that following the recommendation of the larger government bodies on the adoption of a proposed standard (in this case the MTM projection system), some municipalities (such as the city of Toronto, or the National Capital Commission) adopted and developed their
several municipalities are achieving an increasingly sophisticated urban
digital mapping system.\footnote{For instance, Vancouver, Burnaby, Calgary, Edmonton and Toronto. For
discussion of information technology in Canadian cities, see for
example: Barry S. Wellar (Editor), \textit{Prospects for Planning: Coming to
Grips with New Realities}, (Ottawa, Canada: Canadian Institute of
Urban and Regional Information Systems Association} (URISA) provide a
fairly comprehensive overview of related developments in urban
digital mapping systems. References specifically relating to Canadian
contributions can be found in K. Lauder and L. Lavallee, \textit{A Canadian
Bibliography of Urban and Regional Information System Activity},
Ministry of State for Urban Affairs, Ottawa, 1976.}

\subsection*{2.3.5 Concepts of Scale}

The term parent or base scale is used to describe the largest
scale of topographic map required to depict the environment adequately.
These parent scales frequently vary between countries. For example in
Britain the parent scale for national topographic mapping is $1:1250$ for
the urban environment, $1:2500$ for the rural environment and $1:10,000$ for
the wilderness areas; in the United States, $1:24,000$ and in Australia
$1:100,000$.\footnote{It's interesting to note that NATO military require $1:50,000$ scale
maps and the United Kingdom Ordnance Survey is hard pressed, and is
working feverishly to produce them from the digital data bases at the
scales noted above.} In Canada, the current federal program aims to provide
national parent scale coverage at the $1:50,000$ scale. The major thrust
of Topographic Survey has thus been directed towards the completion of
the $1:50,000$ scale series topographic maps, revision of existing maps,
with new mapping priorities being directly related to specific
developments such as pipelines, roads, and transmission lines. In 1975
approximately 6,575 sheets of the $1:50,000$ topographic series had been.

\footnote{For instance, Vancouver, Burnaby, Calgary, Edmonton and Toronto. For
discussion of information technology in Canadian cities, see for
example: Barry S. Wellar (Editor), \textit{Prospects for Planning: Coming to
Grips with New Realities}, (Ottawa, Canada: Canadian Institute of
Urban and Regional Information Systems Association} (URISA) provide a
fairly comprehensive overview of related developments in urban
digital mapping systems. References specifically relating to Canadian
contributions can be found in K. Lauder and L. Lavallee, \textit{A Canadian
Bibliography of Urban and Regional Information System Activity},
Ministry of State for Urban Affairs, Ottawa, 1976.}
compiled. When completed, the 1:50,000 series will have 13,150 sheets of which exactly 9,000 had been compiled as of March 31, 1982."

From this basic scale coverage the regional or local family of topographical map series can be derived in a decreasing order of smaller scales. Where the provincial and major municipal governments undertake larger scale mapping (where the requirements justify and resources permit), these new larger scales will become the parent scales.\(^7\) The important point to observe is that unless a coordinated approach is taken to the federal, provincial and municipal programs, serious diseconomies will be encountered and the potential benefits of automation will hardly be realized.\(^7\) The adoption of a flexible, inclusive, and open-ended topographic feature classification for the exchange of spatial data is a fundamental prerequisite of such a coordinated approach in the digital era.

2.3.6 Uses, Collation and Exchange of NTS Data

Table 2 provides an indication of some specific users of the Canadian topographic map series at the 1:50,000 scale. Many of these users are utilizing digital technology and the implications of their needs in the future should be borne in mind as one designs a national topographic feature classification standard. Recent research is also


\(^7\) Harris 1980, p. 81.

\(^7\) Work in the United Kingdom contrasts in this case. The contrast is a legislative union versus a Federal country. In the United Kingdom all mapping is done by the Ordnance Survey.
strongly indicative of the fact that the optimal use of national resources requires that Survey and Mapping Departments supply digital data to the user in their true geometric form regardless of paper graphic production mandates. In fact, the great utility of the data supplied by Topographic Survey is their locational accuracy and the foundation this provides for many phases of digital mapping.

<table>
<thead>
<tr>
<th>TABLE 2</th>
</tr>
</thead>
<tbody>
<tr>
<td>Some Specific Uses: the 1:50,000 Topographic Map Series</td>
</tr>
</tbody>
</table>

| military operations | engineering |
| education | administration |
| resource exploration and development | energy corridor planning |
| recreation | trade and commerce |
| base maps for thematic map series | communications |
| revision of smaller scale maps | geographical studies |
| rural real estate | land planning |
| scientific studies | pollution control |
| total planning of large areas | supply of natural engineering materials |


As a general rule in the exchange of digital data the data as collected should represent the real world as accurately as possible. Graphically edited digital data to fit a specific cartographic product or use of a product may make the data virtually valueless for other applications and may result in a complete recollection. Fortunately

74 Vitek cites this example: "In collecting road data on two parallel roads the operator may recognize that when symbolized the roads will
Topographic Survey has decided to supply 'positionally accurate' data to users. This is a good approach, similar to the model followed in the United Kingdom, where for example, digital data are supplied and sold by the Ordnance Survey in their geometrically correct form.

A criticism launched by the (Canadian) Task Force on National Surveys and Mapping however was that Topographic Survey held the view that their mandate as defined in the legislation appears only to cover the mapping needs for resource development. After hearing from over 200 agencies, the Task Force were "overwhelmed by the broad range and intense interest of users which extend far beyond the confines of resource development as shown by the specific uses listed in Table 2."

In fact, the Task Force disagreed with Topographic Survey's interpretation, and contended that under the Government Organization Act, 1970, c.14 (2nd Suppl.) s. 9, where "technical surveys" are defined as geological, topographical, hydrographic, oceanographic and meteorological surveys, there is no intimation of such a restriction."

The point to realized is that an optimal operational general purpose overlap and therefore moves the road before digitizing. The data fits the product but they no longer represent the real world. One solution is to digitize the road twice, once for general use and a second time as offset for a specific product use. As software becomes available to make post collection corrections interactively, the need for the second digitization would not be required. This problem also occurs where the data are collected showing correct positions, however in later steps, in checking overlays of data, such as the symbolized roads and building, the building locations are moved because of symbol overlaps. The final product data base is valid for the product but not for general use. In this case the original data which shows the correct position must be saved for exchange purposes." Richard L. Vitek, "Data Base Requirements For Geographic Mapping," Data Base Techniques for Pictorial Applications, Lecture Notes in Computer Science, vol. 81, edited by A. Blaser, Springer-Verlag Berlin Heidelberg 1980, pp. 1-25.

topographic feature classification standard mandates holistic and sound system design (a clear picture of potential uses), experimentation (to avoid difficulties which may be encountered in the future) and implementation.
Chapter III

TAXONOMY AND CLASSIFICATION

One of the first realizations in this research, in attempting to achieve a thorough and consistent investigation of the topographic feature coding problem is that it requires an extensive and far-reaching background, and one of the first steps required would be an exercise on the subject of taxonomy and classification.

Through the illustration of classification principles, one transcends various realms of knowledge. Dolby for example distinguishes between classification of the sciences, the logico-mathematical theory of classification, the classification of the objects of study within a particular science (for example, as in the biological classification of living organisms), and also between the philosophical traditions of the sciences from related practical classifications intended for particular purposes (such as library science)."

Taxonomy basically deals with the classification of all living things according to observed natural, or hypothetical, relationships, or both. The idea of taxonomy was first made explicit in the history of Western thought by Aristotle in his ORGANON and METAPHYSICS."

---


ruling principle is that the highest genus is divided by means of
differentiae into subaltern genera, and each of these is then divided
and subdivided until the ultimate species is reached. This principle
has been handed down through the Stoics, Porphyry and the Greek
commentators to Linnaeus, from whom it passed into modern biological
usage.\[^{96}\]

Some of the leading taxonomists such as John Ray (1627–1705) who
followed the Aristotelian method of classification based his systems of
classification on all the structural characteristics (as opposed to
single features) of the observed phenomena.\[^{1}\] Carolus Linnaeus
(1707–1778), usually regarded as the father of modern taxonomy,
attempted a natural classification by initiating a standard hierarchy of
class, order, genus, and species. Utilizing the principles of the
Linnaean approach, the taxonomist first recognizes natural groups and
then decides on the rank that should be assigned them. The Linnaean
system was the first to use binomial nomenclature consistently, provided
rules for assigning names to plants and animals, and also provided
workable keys that made it possible to identify plants and animals.\[^{2}\]
One approach to modern biological classification is to identify or
recognize taxa through the utilization of standardized keys and to
assign these taxa to a particular rank or natural grouping. The

\[^{96}\] A. L. Peck, "Introduction," Aristotle: Historia Animalium,
pp. v–viii.

Westfall.

\[^{2}\] Encyclopedia Britannica, 15th ed., s.v. "Classification Biological:
morphological basis of such classifications distinguishes between closely related species of plants and animals on feature differences such as colour, size, proportion, presence or absence of cellular differentiation, anatomy etc.\textsuperscript{33}

Grigg observes that classification, defined as the grouping of objects (elements) into classes (sets) on the basis of common properties or relations, is a necessary preliminary in most sciences, and it is often argued that the state of classification is a measure of the maturity of a science. Since the time of Aristotle philosophers have discussed the logical procedures designed to prevent internal inconsistencies from arising within a classification scheme.\textsuperscript{44}

Armand notes that Aristotle, the founder of scientific logic, and other great thinkers such as Avicenna, Bacon, Descartes, von Leibnitz, and Lomonosov, saw in logic a means for uncovering and proving the truth. Although classifications can be built on various principles (morphologic, generic, temporal, spatial, quantitative, etc.) all must follow certain general and unalterable laws of logic: 1) the sum of classes must be equal to the scope of the classified generic concept; 2) only one classificatory criterion should be used within any one level of classification; and 3) a classification must not skip logical levels.\textsuperscript{55}


Lee defines classification as a fundamental method of science which seeks to introduce simplicity and order formulating a scheme of mutually exclusive and collectively exhaustive categories based on the important characteristics of the things concerned and on the actual relations between them. Lee's definition corresponds closely with the logico-mathematical theory of classification, which in turn coincides with what, in the mathematical theory of sets, is called a partition, "a division of a set of objects into subsets is a partition if and only if: 1) no two subsets have any elements in common; and 2) all of the sets together contain all of the members of the partitioned set."

3.1 NOTATIONAL FLEXIBILITY AND FORMAL GRAMMARS

To the rules that a successful classification must be mutually exclusive and collectively exhaustive, Wynar adds that the flexibility of the notation is of first importance if the classification scheme is to be current. The classification scheme must also employ terminology that is clear and descriptive, with consistent meaning for both the user and classifier. Although most traditional classification schemes are based on a logical division of the universe of knowledge, by contrast, computer-based classification systems are empirical and descriptive, attempting to develop thesauri with but one thing in common: a set of descriptors well suited to manipulation.

---


One system which integrates both logical consistency and notational flexibility was proposed by Ranganathan. His analytico-synthetic scheme aims at analyzing the subject matter into constituent elements (facets), and extends the traditional rules of logical division in that categories are free to combine with each other. Five fundamental characteristics (facets) are recognized: PERSONALITY (P), MATTER (M), ENERGY (E), SPACE (S), and TIME (T). The symbol in brackets (facet indicator) is used to differentiate which characteristic is being referenced. A series of tables, classified according to the rules of logical division is associated with each facet (dimension), and the extension of the colon system to any subject matter (basic class) forms the basis of 'facet-analysis' theory. The system is based on a seemingly simple principle: list facets and combine them, following the identification of a basic class, or subject matter, in the order (P-M-E-S-T). Repeat of a facet indicator identifies that a second level (characteristic) of this facet (dimension) is repeated.

Used in the context of geographical mapping Buichguens et al. state that it is appropriate to use information retrieval languages of the facet-classification - and descriptor type, largely because these languages reflect logical and semantic correlation in characterizing data. Their factorial subsystem, in progress since 1975, requires a


** Description, principles, and procedures of the faceted classification can be found in B. C. Vickery, Faceted Classification (a guide to construction and use of special schemes prepared for the Classification Research Group), London, Aslib, 1960; and B. C. Vickery, Classification and Indexing in Science, 3rd ed., London: Butterworth, 1975.
rich and flexible information retrieval language presented in thesauri. In their thesaurus, semantic characteristics of features are given as pairs, consisting of a characteristic (property) and an associated value. The authors mention that set theory and predicate logic are used to compile the formulae of retrieval prescriptions and elementary information about the feature and its characteristics. They further state that the most important result of their research is the provision of complete, accurate and uniform cartographic information.

3.2 PRINCIPLES OF LOGICAL DIVISION AND CLASSIFICATION

Realizing the importance of logical and consistent classification criteria, ten principles that should be taken into consideration when designing a classification system, as proposed by Grigg are cited below:

1. classifications should be designed for a specific purpose;
2. objects which differ in kind will not easily fit into the same classification;
3. classifications must be changed as more knowledge is gained about the objects under study;
4. the differentiating characteristics should be properties of the objects classed;
5. in logical division the division should be exhaustive;
6. in logical division and classification the species or classes should exclude each other;
7. in division, the division should proceed as far as possible upon one principle;

8. the principle of division must be important for the purpose of the classification;

9. properties which are used to divide or classify in the higher categories must be more important than those used in the lower categories;

10. the logical consistency of the hierarchy will only be maintained if rules five through nine are observed.  

3.3 FURTHER OBSERVATIONS ON THE CLASSIFICATION PROCESS

Grigg further suggests that a general purpose classification attempts to express the actual order or system of things classified by correlating the greatest amount of knowledge about the objects under study. He also remarks that despite some taxonomists who assume that if a differentiating characteristic with a large number of accessory characteristics is found then this classification will be the correct classification and will serve all purposes equally well, modern logicians and taxonomists would agree that there can be no natural classification in the sense that there is one and only one classification which will serve all purposes.  

Ellen adds that in practice no classification can be formulated in terms of all natural discontinuities, but classifications can be ordered according to their degree of departure from the philosophically ideal. He thus suggested

--


seven indicators to evaluate the internal logic and ordering of a classification:

1. variability;
2. arbitrariness;
3. expression of inclusiveness;
4. anomaly;
5. structural complexity;
6. terminology, nomenclature and taxonomy; and
7. integration in semantic fields.''

3.4 **A COMPLEX BUT RELEVANT PROCESS**

Classification is obviously very complex and based on logical principles which necessitate much conceptual foresight in both design and implementation. Problems inherent in any classification system vary in extreme depending on the initial objectives of that science or discipline. These problems are more manifest in a general purpose classification because of the attempt for broader applications. The choice of adhering to a particular classification structure really revolves around the fundamental question of classification purpose. If we attempt, as in the context of this thesis, to review and propose ideas on a general purpose topographic feature classification to serve as a national standard for the handling and transfer of digital spatial data, then we must be prepared to evaluate selected systems on the basis of the requisites that this concept entails. The task is not trivial and should be dealt with accordingly. Harvey makes several pertinent

"Ellen, 1979."
remarks concerning the fact that the geographic literature is replete with complex classifications, many of which appear to have been devised with no particular purpose, and many which have never been used for anything. He further states that..."if we fail to identify our purpose we must expect our purpose to be governed - perhaps insidiously - by the classification system we adopt...Let us ensure that we do not allow the classifications developed to prejudice our future and let us also ensure that we make an efficient use of them in pursuing a deeper understanding of the phenomena we are concerned with.""

3.5 CODING CONVENTIONS

Whereas classification is the grouping of objects on the basis of common properties or relations to help elucidate a particular problem, coding is simply a systematic method of labelling and recording the groups of data for processing purposes. It is basically an abbreviated or alias format for information exchange which assumes nothing more, at least initially, than an information source of symbols which can be uniquely recognized."'


As is well known there are many kinds of coding systems that have been standardized, all of which must be expanded as the need arises, and all of which, from a historical perspective are always undergoing change. Standard labels (i.e. file headers) are types of codes that facilitate and are utilized in the exchange of digital data. ASCII and EBCDIC codes are well known bit numbering codes, both valuable coding systems, largely because of the fact that they are standardized. Since computers frequently work in bytes, which are defined to be 8 bits each, hexadecimal codes (a numbering system allowing all logical extensions of binary arithmetic) are better adapted to machine architecture than the more common decimal system of numbers, and have evolved to be frequently used for computer processing." The use of hexadecimal digits allows for the full saturation, hence exploitation, of the code's capability in EDP." Morse code, a variable-length coding system designed for speed of transmission of information, facilitates its objective by assigning the most frequently used letters (i.e. the letters i and e) the simplest (shortest) code and the infrequently used letters (such as j), a longer code. The International Standard Book Number (ISBN) is an example of a widely used fixed-length coding system which publishers assign to uniquely identify their texts based on fields (logical units) of country, publisher, book number and a sum check (error detecting) digit.

" Internally it is all binary arithmetic. Representationally, we use hexadecimal codes as they fit into bytes quite nicely.

" The ASCII (American Standard Code for Information Interchange) system uses 7 binary digits (with the eighth digit serving as a parity check) to represent 128 possible alphabetic, numeric and assorted other symbols. EBCDIC (Extended Binary Coded Decimal Interchange Code) codes allow 256 possible symbol combinations (many of which are currently unassigned).
The standardized \( x, y \) coordinate system is also a type of coding system that could be represented by an infinity of ways but, to make it understood, the right hand rule of \( x, y, z \) was adopted to allow users to get onto more interesting things.

3.5.1 Coding Geographic Data

There are many ways of coding data so as to use them. A symbol on the map is a code, as are Standard Land Use Classification codes, Standard Industrial Classification codes, transportation codes etc. The distinguishing feature between the many ways of coding the data depends upon the purpose for which we intend to use these data.

The coding of a topographic feature classification is no different in principle, than any other spatial coding scheme, such as, for example, geocodes, except that geocodes typically deal exclusively with areal data (regions). Otherwise the same two types of information are provided: "one which tells something about the place, and one, such as coordinates, which describes the geographical location and relationships of the place."

The term geo-referencing (geographical referencing) is commonly applied to a geocoding system which, cross-connected and enhanced with geographic coordinates such as latitude and longitude or the UTM Grid (two of the most common and best known reference systems in the world) attempts to provide the working framework for a land data base.

---

management system. The Ontario Basic Mapping Program (O.B.M.), based on the UTM 6 degree grid, is one such system that is presently being developed to provide a complete topographical information base, in digital form, at the 1:10,000 (Southern Ontario), 1:20,000 (Northern Ontario) as well as the 1:2,000 scale mapping (under a cost-sharing agreement with municipalities). There exists much commonality between a geo-referencing system and a topographic feature classification system. The two should not be viewed as mutually exclusive, but rather as points on a continuum in the classification and coding process.

3.5.2 The Design of a Coding System

The most important principle to ensure in the assignment of a coding schema to a classification of data is that we do not artificially reduce its intended order and flexibility. Within this framework the geographer may wish to investigate: 1) a mnemonic code (i.e. how language can be coded based on the abbreviation of words, such as by the elimination of vowels, but yet retaining interpretative association with the original words); 2) the question of computer storage requirements and the concepts of bits per byte, bytes per word and numeric versus alphanumeric savings; 3) the list data structures and their feasibility for implementation; and 4) cross referencing in the classification allowing groupings across properties (something rather standard in List Processing). Of course, of equal importance in a coding system, is the need to facilitate ranging capabilities. Although it is possible to include the properties of a logical subset (a facility required for the retrieval and display of ranges of data) with either alphanumeric or
numeric codes, and although greater flexibility is offered by the alphanumerics set, computing is made easier with a numeric set. Tom suggests that a gapped system which employs a decimated option (also internally gapped) is a possible approach and that "considering the parameters of storage, data structure and spatial data processing, a numeric code is advantageous over an alphanumeric code because generally decimated codes are more efficient for storage and data structures."\(^{101}\)

3.5.3 Selected Approaches to Topographic Feature Coding

Hanke describes a code consisting of the four fields of information, N1, N2, N3, N4 where: N1 represents information on the drafting function; N2 represents information on the computation function; N3 represents information on approximate or actual value of coordinates; and N4 represents information on the Pen Function (begin or end of line, special search function).\(^{102}\)

Christ describes a 6-digit code used in the German Federal Republic which supplies three fields of information: 1) a unique feature is identified; 2) within a designated category; at 3) the scale at which the feature was first digitized. For example, the code 504002 identifies creeks greater than 5 metres wide (002) of category hydrography (4)


first encoded at the 1:50,000 scale.\textsuperscript{103} Weber describes a similar coding logic (minus the information on scale) where the code 3401 (for example) represents a traffic way (3) for cars (4) and finally a motor highway (01).\textsuperscript{104} In April 1975, by an agreement between the surveying administrations of the West-German Länder entitled the "Standard Data Format for the Exchange of Digital Cartographic Data", the feature codes of all classes of objects appearing on official topographic maps at scales of 1:5,000 to 1:1,000,000 were standardized by a four-digit decimal classification inducing a hierarchical order of the object classes. Weber observed however, that no data had been exchanged in this format by 1979.\textsuperscript{105}

An approach described by Bertschinger for an interactive system in Switzerland provides information according to the categories of object code (point or line), object type (for example, inhabitants), and object values (for example, 100,000). Any number of object codes, as well as \textit{any number of descriptors} (types, values) are allowed per feature. The basic element (feature) of the data base (the object/edge) is identified as being either a line such as a river (8110:li), or as a point feature such as a church (1142:sp).\textsuperscript{106}

\textsuperscript{103} Fred Christ, "Digitizing Digitizer Editing and Graphic Output of Topographic Map Data," ed. Erno Csati. op. cit., 1974, pp. 39-44.


Another approach used on a small cartographic system at the Department of Geodesy at the University of Technology, Delft consists of three digits (fields) of 'ABC' where for a point object, A must always be 0, B must be between 0 and 8, and if the feature is a control point, B must be 9. \(^{10}\) C can be any number. In coding point objects, a detail point can be coded by a number between 0 and 89 and a control point by a number between 90-99. In the coding of line features, A must always be 1 for straight lines, and 2 for curves. B represents the nature of the line to indicate whether the line is fully drawn, dotted or dashed and can be any number between 0 and 9. If C = 1, then the point is a begin point, else if C = 2, then the point is the end point of the line, else if C = 3, then the line object belongs to more than one model, map or photograph, and is used as a special feature for join line searches (i.e. searching two points with C=3 and the same A and B having the same coordinates).

An 8-digit code initiated in the early 1970's by the Surveys and Mapping Branch, Canada is based on a classification arranged into ten categories of features, each of which is identified by the first digit in the code. The second and third digits specify a particular kind of feature within these categories and the remaining digits serve to describe particular aspects of a feature. For example, the code 00300250 identifies a depression contour at 250 metres elevation, the code 23400000 identifies a standard gauge railroad line, single track, while the code 23420000 identifies a standard gauge railroad line, abandoned. Kihl notes that as of 1975, 160 maps from the 1:50,000 scale National

Topographic Series have been digitized and that a pilot sheet at the
1:100,000 scale is being derived from the data stored in four 1:50,000
scale cartographic files.¹⁰⁸

3.6 IMPLICATION OF TAXONOMY TO THE FEATURE CODING PROBLEM

As noted the digital description of a cartographic or geographic
feature can be considered as having two components, a geometric
description and an attribute description. It is the attribute
description which, at the user level, allows us to identify and select
features. The sophistication of our retrieval mechanism depends not only
on the logical organization or grouping of the features but also on the
number of different non-spatial attributes associated with each feature.

Taxonomy implies the ordering of information into a hierarchy to
develop understandings of phenomena, but in the context of cartographic
or geographic features, to allow groupings on the basis of the criteria
used to form the hierarchy. A tree structured ordering, for instance,
could be made to allow a multitude of alias names or ordered groupings
to result in the same classification. A digitizer operator may identify

¹⁰⁸ T. H. Kahl, "Automated Production - 1:50,000 Scale Topographic
current (1984) coding system is arranged into ten categories:
0)contours; 1)control information; 2)roads; 3)communication;
4)buildings; 5)marine features natural; 6)vegetation and
hypsography; 7)other cultural features; 8)marine features man-made;
and 9)text and labels. Each feature is tagged with a five-digit
numeric code to identify the geometric data (with each data point
represented as a complete x,y,z element). Data is stored as
"cartographic spaghetti." A six digit alphanumeric code is further
tagged to each feature based on fields of information including
level (1 of 63 levels), weight (line thickness, colour), line style
(e.g. solid line), automated closure factors, and type of feature.
The centre line is followed for most features except outlines to
scale such as larger buildings. The attribute data (feature code) is
stored with the geometric data.
a feature with the words "thermal power station," while another might say "atomic electricity plant." Logical algorithms could be constructed to automatically recognize both as being the same thing placing them into: "building, industrial, electric power, thermal, nuclear". The coding system should be such to allow such algorithms to be written. On the retrieval side, like objects should be sequentially groupable on the basis of their identification or their properties. Beyond being possible to call all "jails and penitentiaries," it should be possible to call all "two to six lane roads," obviating the need to enter "three", "four", and "five". Furthermore, groupings across properties should be possible, such as a call for all "abandoned" things. These capabilities would be prerequisites for a system that would automatically filter and generalize data to prepare a file for maps at different scales. "Extract all information on transportation suitable for a map at the 1:4 million scale."
Chapter IV

ANALYSIS OF SIX NATIONAL BASE MAP FEATURE CLASSIFICATIONS

Six classifications, now in use, or in the proposal stage, were analysed with the objective to identify an optimal code that was inclusive, flexible and open-ended, that could possibly be used in a general system, or at least to identify principles towards the development of one. Review, analysis and evaluation of these systems placed emphasis on identifying general strengths as well as pinpointing weaknesses and faults. Particular attention was paid to: 1) logical consistency; 2) terse and exact feature description; 3) data measurement level; 4) multi-dimensional aspects; 5) ranging facilities; and 6) parameter coding.

4.1 THE USGS ATTRIBUTE CODES FOR DIGITAL LINE GRAPHS

One responsibility of the USGS, National Mapping Division, is to manage the National Mapping Program (NMP) in preparing and making available multi-purpose maps and geographic data to users on a national basis. The "Attribute Codes for Digital Line Graphs" were initiated with the formation of the Digital Application Team in 1977 with the purpose of creating and integrating two basic computer file formats, Digital Line Graphs (DLG's), and Digital Elevation Models (DEM's), to facilitate this activity.¹⁹ Each DLG file, initially corresponding to a standard
1:24,000 topographic map (with each category of map information stored in a separate file or sub-file), consists of three levels. Level I contains raw digital data suitable for low to moderate quality plotting or display. Level II files have been edited and contain extensive graphic attribute codes. Level III files are designed for integration with GIS's requiring fully topologically structured data sets.

Designed to accommodate all categories of planimetric data represented on conventional line maps, the specific purpose of the DLG coding system is to achieve fully topologically structured data sets (capable of fulfilling the needs of both automated cartography and geographic analysis) where "all lines are explicitly linked to nodes (ends of lines or line junctures) and all areas are explicitly coded to the bounding arcs (lines) such that all nodes, lines, and areas are linked into a completely defined digital line graph that can be interpreted in a computer environment without resorting to human assistance." 118

The primary source for the DLG coding schema are those features listed in the USGS Technical Instruction and Symbol Specifications for topographic maps at the 1:24,000 and 1:100,000 scales. Small-scale DLG data bases digitized from the 1:2,000,000 scale sectional maps of The National Atlas of the United States are also produced and distributed. A third program to produce DLG data from the 1:1,000,000 scale map series is now being initiated in co-operation with the Bureau of the Census.

4.1.1 Description and Analysis

The 1980 coding schema was organized according to eleven base categories, six of which were incomplete: reference systems, hydrography (wetlands and coastal features), surface cover, non-vegetative features, survey control and markers, geographic names, and orthophotographic imagery. The non-base categories supply peripheral information such as the 33 7-digit codes associated with the U.S. Public Land Survey (which identify the section, township, range, origin of survey by principle meridian etc.), the 133 4-digit codes used to identify State Plane Coordinate Zones, and the 20 2-digit codes for identifying UTM Coordinate Zones.

Table 3 shows that the completed base categories of the USGS system are further arranged according to the node, area, line, and point (degenerate line) characteristic of each feature classified. Cross-classifying nodes as attributes of geographic features serves the functional purpose of identifying the beginning and end points of linear features (and/or some significant point along a linear feature, such as an intersection) and thus changes in feature classification. This concerns for example, the machine recognition, within the data base, of features coded as the upper origin of a stream, change in stream classification or status, stream road intersection, shoreline road intersection, road-railroad intersection, structure over railroad, canal or stream junction, or pipeline intersection. Further applications are quickly found in detecting either ends of bridges, tunnels, dead end streets, or features along a linear feature, such as a gate. Nodes are

### TABLE 3

**Distribution of DLG Codes by NMP Category**

<table>
<thead>
<tr>
<th>Base Category</th>
<th>Organization of USGS Attribute Codes</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Node</td>
</tr>
<tr>
<td>Hypsography</td>
<td></td>
</tr>
<tr>
<td>Contour</td>
<td></td>
</tr>
<tr>
<td>Hydrography</td>
<td></td>
</tr>
<tr>
<td>Streams</td>
<td>32</td>
</tr>
<tr>
<td>Water Bodies</td>
<td>5</td>
</tr>
<tr>
<td>Boundaries</td>
<td>2</td>
</tr>
<tr>
<td>Transportation</td>
<td></td>
</tr>
<tr>
<td>Roads and Trails</td>
<td>27</td>
</tr>
<tr>
<td>Railroads</td>
<td>22</td>
</tr>
<tr>
<td>Water Navigation</td>
<td>3</td>
</tr>
<tr>
<td>Transmission/Pipelines</td>
<td>25</td>
</tr>
<tr>
<td>Significant Cultural Features</td>
<td>2</td>
</tr>
<tr>
<td></td>
<td>118</td>
</tr>
</tbody>
</table>


Further used to identify selected features such as a power station, gas field or refinery, which logically occur as the end (or beginning) point of linear features such as transmission lines or pipelines. The provision and labelling of the nodes within the data base (network of base map features) as part of the classification codes is unique to the USGS feature coding schema.

The dimensional classification of features on the basis of their lineal geometric property is logical in that a road, stream, shore of
island, trail etc., are most frequently encoded as a single line. Many line features are encoded simply as the centre-line of the 'real world feature' they are representative of, and are thus accompanied by attributes to their other dimensions (such as width). The topological significance of this coding ties in with the concept of a completely defined digital line graph. For example, the two networks of hydrography and transportation account for 85 and 97% of the feature codes categorized as lines and nodes respectively.

Classifying features according to their areal significance is founded on the principle of polygon recognition, for example, boundaries or water-bodies. In certain instances however, the dimensional classification of features on the basis of area raises the question of consistency. At varying scales is not a point an area and vice-versa? In the category of significant cultural features (approximately 28% of the feature codes), the geometric difference between features coded under area attributes such as large church (140.0102), zoo (140.0135), fairground (140.0138), orphanage (140.0112), and those features classified under point attributes such as church (140.0302), fairground (140.0341) or rodeo grounds (140.0341) is unfounded unless the whole data base is scale specific.  

4.1.1.1 Parameter Coding and Data Integration

The 450 base category features of the 1980 schema are identified by a 7-digit numeric code divided by a period into a 3 digit major code and a 4 digit minor code. The first two digits of the major code define

\[111\] This redundancy has been eliminated in later revisions of the USGS "Attribute Codes" (revised Feb. 1982, Sept. 1983, and Oct. 1983).
the map category (and are unique for the category). The third digit is used in two ways: 1) if zero, the minor code following identifies a particular feature for that category; 2) if non-zero, the code has a special parameter interpretation. As a parameter, the third digit of the major code and all four digits of the minor code are interpreted for their numeric value, each of which may have an alias interpretation.

Table 4 illustrates only a small sample of parameter codes which identify features (within a specified base category) utilizing a parameter 'digit' (the symbol N) to describe a particular quality from a list of possible attributes. The rational and advantage of this type of coding is that it provides a series of relational information measures such as accessibility (unrestricted, limited, controlled), temporal status (proposed, abandoned) etc., to selected features throughout the coding schema. However, several problems could develop if the mechanism is not carefully implemented: 1) certain feature attribute combinations could be precluded because the list of possible attributes is mixing concepts (accessibility, type of road, ownership and temporal status); 2) the number of possible alternatives (designated by the special symbol 'N') is limited to ten; and 3) the parameter (descriptive attribute) is buried within the code itself and may make it difficult to access and share data.¹¹³

¹¹³ This latter point was raised by David Godfrey and Ernest Chang in The Telidon Book, pp. 179-188, Victoria B.C.: Press Forcepic (1981) largely on the basis of research reported by E. F. Codd, "A Relational Model of Data for Large Shared Data Banks," Communications of the Association for Computing Machinery 13, no. 6 (June 1970):377-387.
<table>
<thead>
<tr>
<th>Code</th>
<th>Meaning</th>
</tr>
</thead>
<tbody>
<tr>
<td>100.N201</td>
<td>primary route, hard surface (undivided)</td>
</tr>
<tr>
<td>100.N202</td>
<td>primary route, hard surface (divided, 25' or less)</td>
</tr>
<tr>
<td>100.N203</td>
<td>primary route, hard surface (divided, 25' or more)</td>
</tr>
<tr>
<td>100.N204</td>
<td>primary route, hard surface (one way traffic)</td>
</tr>
<tr>
<td>100.N205</td>
<td>secondary route, hard surface (one way traffic)</td>
</tr>
<tr>
<td>100.N206</td>
<td>secondary route, hard surface</td>
</tr>
<tr>
<td>100.N207</td>
<td>improved light duty</td>
</tr>
<tr>
<td>100.N208</td>
<td>unimproved dirt</td>
</tr>
<tr>
<td>100.N209</td>
<td>trail</td>
</tr>
<tr>
<td>100.N210</td>
<td>4-wheel drive vehicle trail</td>
</tr>
<tr>
<td>100.N211</td>
<td>urban street</td>
</tr>
<tr>
<td>100.N212</td>
<td>foot trail</td>
</tr>
<tr>
<td>100.N213</td>
<td>bridle trail</td>
</tr>
<tr>
<td>100.N214</td>
<td>pack trail</td>
</tr>
<tr>
<td>100.N215</td>
<td>historical trail</td>
</tr>
<tr>
<td>100.N216</td>
<td>bicycle trail</td>
</tr>
<tr>
<td>100.N217</td>
<td>primary route, hard surface (interchange road)</td>
</tr>
<tr>
<td>100.N218</td>
<td>secondary route, hard surface (interchange road)</td>
</tr>
<tr>
<td>100.N219</td>
<td>improved light duty interchange road</td>
</tr>
<tr>
<td>100.N240</td>
<td>ferry crossing</td>
</tr>
<tr>
<td>100.N241</td>
<td>road through parking area</td>
</tr>
<tr>
<td>100.N250</td>
<td>perimeter of parking area</td>
</tr>
<tr>
<td>100.N293</td>
<td>road or trail subject to inundation</td>
</tr>
<tr>
<td>100.N294</td>
<td>road or trail on dam</td>
</tr>
<tr>
<td>100.N295</td>
<td>road or trail on bridge</td>
</tr>
<tr>
<td>100.N296</td>
<td>road or trail on levee</td>
</tr>
<tr>
<td>100.N297</td>
<td>road or trail tunnel under ground</td>
</tr>
<tr>
<td>100.N298</td>
<td>road or trail tunnel under water</td>
</tr>
<tr>
<td>100.N299</td>
<td>road or trail under construction</td>
</tr>
</tbody>
</table>

where:
N=0 unrestricted access
N=1 limited access
N=2 toll road
N=3 privately operated or controlled public access
N=4 proposed road
N=5 abandoned road

<table>
<thead>
<tr>
<th>Code</th>
<th>Category</th>
<th>Feature and/or Symbol</th>
<th>Comment</th>
</tr>
</thead>
</table>
| 03N.--- Hydrography | stream | E.W.S. | N=3 for feet  
                         |                       | N=4 for metres           |
| 035.--- Hydrography | stream |                    | river mile, right justified    |
| 039.--- Hydrography | stream | C.F./S. | enter (1)             |
| 091.00-- Boundaries |        |          | enter state FIPS code in two digits right justified |
| 092.00-- Boundaries |        |          | enter county or county equivalent FIPS code in three digits right justified |
| 101.00-- Transportation |          |          | number of lanes, right justified |
| 102.00-- Transportation |          |          | interstate route number, right justified |
| 106.00-- Transportation |          |          | county route number, right justified |
| 125.0-- Transportation |          |          | water navigation approx. width of canal or channel if discernable from source material right justified |
| 12N.--- Transportation |          |          | water navigation approx. depth of canal or channel (right justified). N=1 for feet, 2 for metres, 3 for fathoms |

**Abbreviations:**
- E.W.S. - elevation of water surface;
- C.F./S - coincident feature or symbol;
- (1) - feature category in blanks right justified;

Table 5 contains only a small sample of codes which further illustrate the relational capabilities facilitated by parameter coding. This includes the capability for ratio data measurement (for example, codes 03N-----, 12N----), coincident coding (for example code 039-----), and integration with FIPS codes (091.00--, and 092.0--

Concerning coincident coding, the mechanism is similarly used for every category such as boundaries (099), roads and trails (109), railroads (119), etc., however it can be seen that it is not possible to consistently identify the coincident features beyond the category level to any particular feature type. For example, in the case where a stream is coincident with a boundary, the single 7-digit coincident code would not inform the user of the actual boundary type. Further, the major code for other hydrographic features such as rivers, irrigation channels or canals, and ditches is '030' also (1980 major code).

Multiple feature coding was implemented in the USGS system to deal with the situation where more than one classification code is required to accurately identify and/or describe a feature (up to 99 codes may be assigned to any one feature). For example, an abandoned, narrow gauge, double track railway would be coded as 180.0201 (railroad) 180.0603 (abandoned) 180.0606 (narrow gauge) 181.0002 (parameter code for number of tracks, right justified). If the railway was coincident with another feature, the coincident code 189.00-- would be tagged on the end of the series of codes. Three small issues are raised with the efficiency of this coding mechanism: 1) although workable, a lengthy conglomeration of codes is often required to identify a feature because the whole 7-digit feature code is repeated for every attribute which is
applicable; 2) the same attributes receive different classification
codes under different map categories. For example, the code for under
construction is 170.0603 with roads and trails, 180.0608 with railroads,
and 190.0601 with pipelines and transmission lines; and 3) regardless of
the capability of multiple codes to identify more than one coincident
feature for any one geometric description, it can be seen that there are
problems in the consistent identification of the coincident features
beyond the category level (i.e. specific delineation to particular
feature types).

An important concept of the USGS classification is its
integration with a wider network of spatial data. Table 5 shows two
examples of parameter codes which incorporate Federal Information
Processing Standard (FIPS) 55 codes, as identifying links in the
encoding of boundaries (codes 091.00-- and 092.0--). This simple, but
highly important concept provides cross reference linkages between FIPS
55 codes and the Bureau of the Census. Further linkages are planned from
the Bureau of the Census to the USGS Geographical Names Information
System. The concept provides a valid research foundation to investigate
these possibilities for other coding systems, such as an integration of
Statistics Canada's Standard Geographical Classification with that of
the proposed CCSM topographic feature classification.

4.1.1.2 Revisions to the USGS Classification Codes

Table 6 shows the revised major categories of the USGS coding
system. The classification has been undergoing extensive revision, the
latest versions examined being the February 1982, September 1983 and
October 1983 Draft Reports. The biggest structural changes involve the addition of new feature codes to, and the elimination of redundant codes in, almost all categories, with a particular emphasis on: 1) the re-organization of the base categories hydrography, transportation systems, and cultural features; and 2) the introduction of the concept of general purpose, and descriptive classification codes. Codes for surface cover have also been implemented based on polygon structures. More explicitly stated in these revisions is the fact that the purpose of the coding system is to: 1) describe real world features and their inter-relations rather than their cartographic representation; 2) provide every standard cartographic symbol with a unique attribute code number; and 3) support a variable number of attributes assigned to each feature (a user may add information) such that one, two or more feature codes may be used for a complete description of a feature. If multiple feature codes are used, their ordering is not significant. Provided also are full textual descriptions for each classification code consisting of feature definition, its associated cartographic symbol (if one exists), and instructions for use of the code concerning the different applications to different scale sources (symbol number references), associated and/or coincidence codes, and scribing scales (if different from standard size symbols). Other changes include the elimination of the category orthophotographic imagery, and the expansion of the small scale codes to two categories.

Examination of each category in detail reveals many further changes such as the code revisions in the category of transportation, their application to both the 1:24,000 and 1:100,000 scale maps, the
### TABLE 6

Revised Major Categories of the USGS Coding Schema

<table>
<thead>
<tr>
<th>Code</th>
<th>Base Category</th>
</tr>
</thead>
<tbody>
<tr>
<td>020.</td>
<td>Hypsography</td>
</tr>
<tr>
<td>050.</td>
<td>Combined Hydrograph</td>
</tr>
<tr>
<td>070.</td>
<td>Surface Cover</td>
</tr>
<tr>
<td>080.</td>
<td>Non-vegetative Features</td>
</tr>
<tr>
<td>090.</td>
<td>Boundaries</td>
</tr>
<tr>
<td>150.</td>
<td>Survey Control and Markers</td>
</tr>
<tr>
<td>170.</td>
<td>Transportation Systems</td>
</tr>
<tr>
<td>180.</td>
<td>Roads and Trails</td>
</tr>
<tr>
<td>190.</td>
<td>Railroads</td>
</tr>
<tr>
<td></td>
<td>Pipelines, Transmission Lines, and</td>
</tr>
<tr>
<td></td>
<td>Miscellaneous Transportation Features</td>
</tr>
<tr>
<td>200.</td>
<td>Significant Man-made Structures</td>
</tr>
<tr>
<td>300.</td>
<td>U.S. Public Land Survey **</td>
</tr>
</tbody>
</table>

**Note:** Major codes for non-base categories of map information.

**SOURCE:** USGS "Attribute Codes," revised Feb./82, Sept./83, Oct./83.

The implementation of codes which describe features scribed to scale (the ratio between pen width for plotting and the actual feature dimensions), or the substitution of the numeric equivalent 01 to 26 for the letters a to z (such as used for road signs and highway route number descriptors, a point which illustrates well the USGS tendency to avoid alphabetic characters in their codes).
### TABLE 7

Revised Distribution of DLG Codes by NMP Category

<table>
<thead>
<tr>
<th>Base Category</th>
<th>Number of Classification Codes</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>(1)  (2)  (3)  (4)  (5)  (6)  (7)</td>
</tr>
<tr>
<td>Hypsography</td>
<td>0     0     5     1     4     0     2</td>
</tr>
<tr>
<td>Combined Hydrography</td>
<td>5     18    6     6     24    17    4</td>
</tr>
<tr>
<td>Surface Cover</td>
<td>0     4     1     0     0     0     3</td>
</tr>
<tr>
<td>Boundaries</td>
<td>2     2     4     1     0     0     4</td>
</tr>
<tr>
<td>Transportation Systems</td>
<td></td>
</tr>
<tr>
<td>Roads and Trails</td>
<td>5     0     18    0     3     29    9</td>
</tr>
<tr>
<td>Railroads</td>
<td>2     0     8     0     3     12    4</td>
</tr>
<tr>
<td>Pipelines, Transmission Lines,</td>
<td></td>
</tr>
<tr>
<td>Misc. Transportation Features</td>
<td>3     0     4     0     8     5     3</td>
</tr>
<tr>
<td>Significant Man-made Structures</td>
<td>2     33    7     7     41    18    1</td>
</tr>
<tr>
<td>Total Number of Classification Codes</td>
<td>19  57  53  15  83  82  30</td>
</tr>
</tbody>
</table>

Notes to identify kind of attribute code:
- (1) Node, (2) Area, (3) Line, (4) Point (degenerate line),

SOURCE: USGS "Attribute Codes," revised Feb./82, Sept./83, Oct./83.

Table 7 illustrates the revised distribution of the Digital Line Graph Codes for the six base categories completed as of October, 1983. A quick comparison with Table 3 reveals a phenomenal amount of change in the distribution, organization, and number of codes assigned to each base category. The introduction and/or reclassification of 165 codes as either general purpose and/or descriptive attributes accounts for much of this change.
General purpose classification codes (minor code range 400-499) apply to multiple feature types and are assigned to features regardless of their geometric configuration. For example, a traffic circle, a railroad station, a power station, or a hydro-electric plant, are classified and receive the same feature code regardless of whether the feature is a point or an area. Previously all features in the USGS classification (Table 3) were classified according to their cartographic data type. This change in organization indicates that the geometric configuration of a feature may not be essential to a feature coding classification.

Descriptive classification codes (minor code range 600-699) are based on the concept, although it is not fully implemented and improvements are warranted, of lists representative of uni-dimensional feature characteristics. Dimensions such as temporal status (proposed, under construction, abandoned, dismantled), access (controlled, restricted), ownership (private, public), site (elevated, under ground) etc., are being implemented into the USGS coding schema through the descriptive classification codes. Each 7-digit descriptive classification code is to be tagged on to a feature as described in the section on multiple feature coding above.

4.1.2 Summary of the USGS Coding System

The designers of the U.S. Geological Survey Attribute Codes for Digital Line Graphs have incorporated one or two advanced concepts and have provided an important model to examine. Parameter, coincidence, and multiple feature codes, are built into the classification. The
topological structuring associated with the dimensional classification and encoding of the base map features is an attempt to more closely model spatial phenomena for machine processing. Further, the classification provides cross linkages with a wider network of Government Agencies and users through the clever use of parameter coding. The fact that the classification is in use but subject to continual re-development, indicates, not only the intractability of the feature coding problem, but also, the experimental nature of the subject.

These re-developments are closely tied with recent work in the area of cartographic data standards and although several weak points were established in this analysis, progressive changes are quickly being implemented. The direction reflective in many of these revisions could be best summarized by listing the six goals identified by the ACSM working group on cartographic features: 1) independence of cartographic representation and scale; 2) universal in nature; 3) logically structured; 4) consist of a single class with multiple attributes; 5) explicitly defined; and 6) derived from basic topographic and hydrographic feature-sets.14

4.2 **THE DMA STANDARD**

The DMA Standard Catalog is designed primarily for use by their three Production Centres: the Aerospace Centre, the Hydrographic Centre, and the Topographic Centre.\(^{113}\) The principal source for features included in the catalog were those listed in the cartographic symbology specifications of each centre, and were supplemented by features in the NATO STANAG symbol book and a draft of the Australian Standard for the Exchange of Topographic Data on Magnetic Tape.\(^{114}\) Although special purpose in the sense that their "primary mission is to provide map, chart and geodetic data to the U.S. Military Service," the DMA's large organizational size, in association with a long history of digital applications, merits their catalog well worthy of study.\(^{115}\) The DMA Topographic Center has been involved in digital mapping applications since 1951 with the production of digital surface matrices starting in 1964."\(^{116}\)


\(^{114}\) The DMA, "Product Specifications for Digital Land Mass System (DLMS) Database," DMA, St. Louis Air Force Station, MO., 1977, (revised 1980) "contain the most recent developments by the DMA to digitally code landscape and cultural features." Written communication, J. L. Sawhobk for Charles D. Hall, Acting Deputy Director, Programs, Production and Operations, Defence Mapping Agency, Hydrographic and Topographic Center, Washington, D.C. 20315, September 10, 1982. Although containing a large number of features of the 1977 Catalog, these specifications illustrate a different data collation methodology revolving around the Feature Analysis Data Table (FADT).

4.2.1 Description and Analysis

The 1977 catalog is based on a four-level hierarchical scheme of: 1) 10 categories; 2) 34 sub-categories; 3) 97 classes; and 4) 753 features as well as the extensive use of attribute lists. In comparison to the major categories of the USGS (revised) coding schema, the DMA system is more comprehensive in that it includes major categories of map information concerned with marine dangers, maritime features and navigation aids, aeronautical features and navigation aids, and related aero/maritime features and navigation aids. Similar to the approach taken by the USGS Codes for DLG's, the sub-categories for categories 3 through 9 are delineated by the area, lineal and/or point characteristic of the feature classified. This dimensional classification of geographic objects was not applied to the first three categories (the 'culture' feature types).

The assignment of a unique four digit numeric code for every feature in the catalog permits computer searches of digital files on four-digit descriptions rather than word look-up. Proceeding from the most to least significant digit, each digit in the code relates specifically to a particular level of the classification hierarchy. For example, in the code 1023, the position of the digits 1, 0, 2 and 3 identify the category (transportation), sub-category (railroad), class (standard gauge) and feature (single track) respectively. The specification of the complete four digit code provides retrieval of the individual feature. One rational for searching the four digit code,

rather than word look-up, manifests itself most upon examination of the explicit nature of the coded features, many of which are but a lengthy conglomerate of modifying concepts. An example of this imbalance is provided by code 3746, which identifies a sunken wreck, dangerous to surface navigation (less than 11 fathoms over wreck). Logically the feature is a sunken wreck. Dangerous to surface navigation is a modifying concept, as is less than 11 fathoms over wreck.

4.2.1.1 Attribute Lists

The DMA code contains a master list of 116 attributes that can be applied universally to any of the ten major feature categories. Also, there are approximately 52 classes of attributes (whose cumulative total from all ten categories equals 875) which are applicable only to specific categories (a summary perspective of these lists is provided in Table 8). There are several important advantages related to the use of these attribute lists: 1) they provide an operational basis for the assignment of multiple attributes per feature;\textsuperscript{11} 2) they facilitate all the possible combinations required for standard maps and charts; and 3) providing that they are constructed on uniform criteria, they facilitate compactness, flexibility and utility. For example, in the first category of the DMA catalog (facilities and populated places), the mathematical possibility of 12,180 geographic descriptor combinations could be encoded for each feature (105 attributes X 116 attributes of the master list). Since up to 99 attributes can be applied

\textsuperscript{11} Up to 99 attributes can be assigned to any one feature. This is indicated by a prefix in the feature code header (2 bytes) which precedes the attribute code, as described in the DMA Format, 1977, p. 9.
TABLE 8
A Sample of the DMA Feature Attribute/Text Lists

<table>
<thead>
<tr>
<th>class/type composition</th>
<th>accuracy alignment</th>
</tr>
</thead>
<tbody>
<tr>
<td>appearance</td>
<td>soundings</td>
</tr>
<tr>
<td>danger area/natural</td>
<td>tides and currents</td>
</tr>
<tr>
<td>danger area/manmade</td>
<td>accessories</td>
</tr>
<tr>
<td>colour</td>
<td>shapes</td>
</tr>
<tr>
<td>foreshore flats</td>
<td>terms</td>
</tr>
<tr>
<td>hard composition (loose)</td>
<td>topmark</td>
</tr>
<tr>
<td>hard composition (marine life)</td>
<td>utility</td>
</tr>
<tr>
<td>hard composition (porous)</td>
<td>lights (text)</td>
</tr>
<tr>
<td>hard composition (solid)</td>
<td>marine stations and land mark (text)</td>
</tr>
<tr>
<td>marine terms</td>
<td>flashing frequency/characteristics</td>
</tr>
<tr>
<td>soft composition (inorganic)</td>
<td>strips and bands</td>
</tr>
<tr>
<td>soft composition (marine life)</td>
<td>frequency</td>
</tr>
<tr>
<td>docking/anchorage facilities</td>
<td>radio stations</td>
</tr>
<tr>
<td>permanently frozen hydrography</td>
<td>radio and radar stations (text)</td>
</tr>
<tr>
<td>shoreline (landform composition)</td>
<td>controlled demarcation</td>
</tr>
<tr>
<td>navigation dangers/restriction</td>
<td>audio communication types</td>
</tr>
<tr>
<td>limits/hazards</td>
<td>accessories</td>
</tr>
</tbody>
</table>


Simultaneously to every feature, the number of possible permutations is greatly increased. In transportation (the second of the three culture categories) 119 attributes (plus the 116 possibilities afforded by the master attribute/text list) can be applied to all 42 features of the category. It is worthy to note that the proposed CCSM classification requires 221 class, category, feature, attribute code combinations to represent 221 possible feature variations for the category of
transportation (without including water-ways). The comparison is made only to stress the importance of facilitating combinations of geographic descriptors in a coding system.

However, several problems could develop if the list mechanism is not carefully implemented: 1) in many cases, the same attribute receives different codes. For example, the 3-digit code for the attribute 'wood' is identified at various places throughout the catalog as 092, 119, 105, 047 and 035; and 2) the choice of alphabetically ordering the attributes and retrieving them on the 3-digit code derived from this sort often precludes retrieving ranges of ordinally related data. For the cartographer this puts restrictions on the type of ranging facilities afforded.

Related to the above are several smaller issues which would prevent the complete integration of the DMA system as a model coding system: 1) because features (like attributes) are numerically coded according to alphabetization, anomalies result (as do problems of facilitating a flexible, consistent and orderly mechanism for insertion) as new features are added. For example, a karez (4349) is positioned between a stream (4348) and wadies (4350), as is (in the category of vegetation) the attribute grapes (020) being placed between garden (019) and general (021); 2) related to the insertion problem, further expansion of new items is limited to 3, 9, and 12 features for the classes of unique waterway features, open water, and hydrographic effected land; and 3) at several places in the catalog there is a mixing of levels of the classification. Standard gauge is categorized as a class, yet it is also coded as an attribute. Certain other features are
coded both as features and as attributes, for example, a river, garden, firebreak, hedgerow or shelterbelt, and in selected cases the same features are assigned different 4-digit codes (example siding, windmill) because they occur under different categories.

4.2.2 Summary of the DMA Catalog

The DMA Catalog is a comprehensive system that meets the requirements of three mapping centres. The extensive use of attribute lists as a type of parameter coding provides evidence of the feasibility of integrating relational aspects with a hierarchical structure, an approach relevant both to the design of a classification based on the concept of word look-up, and also to one which facilitates the use of combinations of geographical attributes per feature. One of the keys to such a system is terseness and consistency in feature description. Several issues to avoid were also identified: 1) problems of insertion (associated with the single integer character representing features); 2) problems of the alphabetical sort (anomalous placement of new items, the logical grouping of items, and the effect on ranging facilities); 3) inconsistency in feature description (terseness versus conglomerations of modifying concepts); 4) mixing of levels; and 5) double coding.
4.3 THE AUSTRALIAN STANDARD

The Australian Standard (1981) was constructed with the purpose of specifying a file structure, and formats, including a feature coding system, to be used for the interchange of digital mapping and charting data. It has been cited as the first standard on digital mapping and charting data exchange, and is intended to provide a means of exchange based on individual records for each feature, without any attempt to define structures or relationships with the data, regardless of the various scales, differing methods, equipment, and organization from which the original data are collated. A wide array of scientific, industrial and governmental organizations were officially represented on the committee entrusted with its preparation.

4.3.1 Description and Analysis

One of the most important advantages of the Australian standard is the fact that the coding system implemented does not alter the structural organization achieved through its classification mechanism, a format which, although less comprehensive for the variety of feature types and information provided, models closely the DMA (1977) system.

---

120 Standards Association of Australia, "Interchange of Feature Coded Digital Mapping Data," Australian Standard AS 2482-1981, Standards Association of Australia, 1981. Inquiries concerning the development of the Standard were made to several Australian Mapping Agencies in August, 1982: the Division of National Mapping, Queanbeyan; the Division of National Mapping, Belconnen; and the Central Mapping Authority, Bathurst. The 1981 Standard was purchased from the Standards Association of Australia in late October 1982.


122 Concerning EDP mediums, formats, and codes, reference is made in the Standard (p. 2) to seven existing Australian Standards.
The 1981 standard consists of 657 features classified according to four categories (culture, hydrography, relief and vegetation), 22 sub-categories (feature type classes), and a large number of small (consisting of only a few items) but (for a large part of the standard) logically related groups of features constructed according to differentiating criteria which closely approximate the concept of pure data series. Because the relationship between the code assigned and the feature identified is not alphabetical (unlike both the CCSM and the DMA Standards), the use of 'logical' ranging facilities is provided.

This coding structure allows the user retrieval, by the re-grouping, at any level of detail. For example, all the features of a particular sub-category or all the features of a particular class (such as accommodation or residential within the sub-category of buildings), or any particular range of features can be specified by using a range of codes. Although a hierarchical structure could well be used in the initial organization of the data, the implementation of the coding structure treats the data as one level. Integration of this concept with any topographic classification system, through experimentation, could result in some significant research findings.

Table 9 illustrates the major categories and sub-categories (feature type classes) of the standard. The total number of features in each major category (indicated in parentheses) illustrates well the concentration on culture (54 % of the classification) and hydrography (26 %). Relief, and vegetation, only account for a total of 10 % of the feature codes. Although slightly unbalanced, and definitely not as comprehensive as the DMA coding system for integrating the needs of
three mapping centres, the Australian system does facilitate
aeronautical (aviation facilities, and aeronautical navigation features)
and hydrographic (foreshore and offshore features, bathymetry, and
navigational features related to hydrography) mapping. Contrary to both
the USGS and the DMA feature classifications, the Australian standard
does not utilize the dimensional classification of geographical objects
as a framework or mechanism for their classification, rather it informs
the user of the nature of the feature (be it point, line, or area to the
right or left side of defined boundary), as part of the feature header
record in the exchange format.
One problem in devising a classification is that of dealing with left-over or miscellaneous feature types. In the Australian standard, the divisive logic of certain categories, such as 'other cultural features' (the ninth sub-category of culture), is not consistently applied. A tower, pole, manhole, telephone box or a landmark area boundary are classified as area and other features. A footpath is grouped with a series of features such as title boundary (road) as a cadastral feature. The linear features of this sub-category are indeed such, but are no different from many other features of the standard which have been classified on different principles such as land use, economic activity, surface characteristic, or proposed function.

Each feature is identified by an 8-digit (integer) feature code. The first 4 digits are the feature identifier. The four digit modifier code, although unspecified by the standard, can be used to further describe a feature, providing a list of these codes and their meanings is user-supplied with the data. As implemented, up to three 8-digit codes can be used to fully describe any particular feature (hence, our multiple feature coding facility). Multiple feature coding to facilitate the usage of combinations of geographic descriptors is most efficiently integrated into a classification when the data are organized according to groups of similar characteristics. Roads, for example, are grouped logically into several classes: 1) primary, secondary, minor, vehicular track, and other; 2) by surface characteristic; 3) proposed function; 4) features related to roads; 5) accessibility; and 6) number of lanes. Although it is limited to three 8-digit codes to identify any

123 See the feature record header segment, character position 16, Table 2, p. 7 in the Australian Standard 2482-1981.
particular feature, the Australian standard does facilitate combinations of characteristics.

There are only a few explicit instances of parameter coding in the Australian standard. All data digitized before coding are assigned the feature code 0000. Modifiers to features in the sub-category of mapping control data such as trig station, or bench mark contain the height of the point in metres. For contours the feature modifier contains their elevation. For relief features, height may be recorded as a feature modifier. For spot heights or depths these values are given as the third coordinate value in a set of coordinate values. However, the fact that feature modifiers must be user defined is both an asset and a deficiency: a deficiency because it reduces their importance as far as a national data base is concerned, an asset because it permits the integration of existing classifications. For example, in the Australian standard a factory, refinery, or a mine receive only one four digit code. If one wanted to be more specific about the type of economic activity associated with these and many other features, the existing Standard Industrial Classification would fit in very nicely. Some very ideal characteristics of a uniform classification, such as terseness, exact definition, and flexibility are thus afforded by this facility.

One disadvantage of the Australian standard is the lack of an exact metric value as part of the feature attribute value. For instance, Table 10 shows that fifteen codes are used to identify the size of a population centre. A feasible alternative is that one feature code could be used, with the exact numerical value attached. Any generalization

124 Quantities or numbers should not be coded since this introduces additional translation and a loss of preciseness. This may be
TABLE 10
An Appeal For Higher Orders of Measure

<table>
<thead>
<tr>
<th>CODE</th>
<th>MEANING</th>
</tr>
</thead>
<tbody>
<tr>
<td>0002</td>
<td>General Population Centre</td>
</tr>
<tr>
<td>0003</td>
<td>Capital City</td>
</tr>
<tr>
<td>0004</td>
<td>Population 500 000 and over</td>
</tr>
<tr>
<td>0005</td>
<td>100 000 – 499 999</td>
</tr>
<tr>
<td>0006</td>
<td>75 000 – 99 999</td>
</tr>
<tr>
<td>0007</td>
<td>50 000 – 74 999</td>
</tr>
<tr>
<td>0008</td>
<td>25 000 – 49 999</td>
</tr>
<tr>
<td>0009</td>
<td>20 000 – 24 999</td>
</tr>
<tr>
<td>0010</td>
<td>15 000 – 19 999</td>
</tr>
<tr>
<td>0011</td>
<td>10 000 – 14 999</td>
</tr>
<tr>
<td>0012</td>
<td>5 000 – 9 999</td>
</tr>
<tr>
<td>0013</td>
<td>2 500 – 4 999</td>
</tr>
<tr>
<td>0014</td>
<td>2 000 – 2 499</td>
</tr>
<tr>
<td>0015</td>
<td>1 000 – 1 999</td>
</tr>
<tr>
<td>0016</td>
<td>500 – 999</td>
</tr>
<tr>
<td>0017</td>
<td>200 – 499</td>
</tr>
<tr>
<td>0018</td>
<td>Under 200</td>
</tr>
</tbody>
</table>


(i.e. groupings and summaries) could then be done by the user (at a later stage) according to whatever ranges of population size were to be investigated. Utilization of exact metric (attribute) values would also allow the elimination of a whole series of 'special' codes (now required) to describe variations in feature characteristics.

A second weakness concerning the implementation of the standard is that often modifying concepts (although not officially recognized as such), are provided under the guise of one code. The features shown in

desirable for purposes of categorization, but statistical value is lost since the actual numbers can not be derived once they are coded.
<table>
<thead>
<tr>
<th>CODE</th>
<th>MEANING</th>
</tr>
</thead>
<tbody>
<tr>
<td>1052</td>
<td>Hotel, Hotel, Inn</td>
</tr>
<tr>
<td>1252</td>
<td>Orphanage, Home for Children, Child Care, Nursery</td>
</tr>
<tr>
<td>1449</td>
<td>Detention Centre, Gaol, Prison, Penitentiary</td>
</tr>
<tr>
<td>1551</td>
<td>Cathedral, Chapel, Church, Convent, Monastery, Mosque, Synagogue, Temple</td>
</tr>
<tr>
<td>2306</td>
<td>Railway Lines (disused, abandoned, under construction, old formation)</td>
</tr>
<tr>
<td>2307</td>
<td>Ski-lift, Chair Lift, Cable Car</td>
</tr>
<tr>
<td>2504</td>
<td>Underpass or Overpass</td>
</tr>
<tr>
<td>2511</td>
<td>Elevated Linear Feature</td>
</tr>
<tr>
<td>4404</td>
<td>Perennial Watercourse: River, Creek, Brook, Stream, Gully, etc.</td>
</tr>
<tr>
<td>5104</td>
<td>Cliff, Escarpment, Breakway, Precipice</td>
</tr>
<tr>
<td>5105</td>
<td>Landslide, Landslip, Eroded Banks</td>
</tr>
<tr>
<td>5107</td>
<td>Joint, Fault Line, Fissure, Intrusion</td>
</tr>
<tr>
<td>5110</td>
<td>Ridge, Rim, Razor Back</td>
</tr>
<tr>
<td>5113</td>
<td>Rock, Boulder</td>
</tr>
<tr>
<td>5116</td>
<td>A Small Relief Feature</td>
</tr>
<tr>
<td>6502</td>
<td>Orchard, Vineyard, Plantation (other than pine)</td>
</tr>
<tr>
<td>6506</td>
<td>Cropland (clearing cultivated), Pasture</td>
</tr>
</tbody>
</table>


Table 11 document a small set of examples. Many of these codes need to be expanded to permit detailed identification of the exact nature of the feature. One interested in the study of the earth sciences may find contention with the conglomeration of codes 5104, 5106, 5107, or 5110. Code 5116 (defined as a small relief feature) is an example of a feature which could be better described. A short-term improvement to the
standard would be to explicitly provide a mechanism to consistently identify a feature from the list of possible types and modifying concepts associated with these codes. No such mechanism, other than a user-supplied list of attribute modifiers, is provided.

There are 24 feature codes assigned to the category of vegetation. These features are further divided into two sub-categories: 1) natural vegetation and 2) cultivated vegetation. Both forest and scrub are differentiated on the basis of dense, medium and scattered. The use of six feature codes to describe two features (ie. forest, and scrub) is questionable. Dense, medium and scattered are actually attributes representing one characteristic of vegetation growth. Code 6014 (secondary growth) is equally applicable to the former 2 features but is precluded from identification by the use of separate feature codes. One could combine code 6014 with the code for a dense scrub for example, resulting in the feature code 60146006 - secondary growth dense scrub. The ambiguity resulting in deciding, as a user, how data should be retrieved, on what criteria, and what specific feature code defeats the purposes of a national standard. Code 6010, grassland (savannah, heath) does not differentiate well the major kinds of savannah formations. The implementation of a list of standardized attributes would overcome this problem. In Table 11, a similar loss of information is encouraged by the lumping together of codes 6502, and 6506. These two codes should be expanded to represent five distinct features, but there is no room for expansion at the feature level (ie. codes 6503, 6504, and 6505 already exist). The ability to insert decimal codes on the least significant digit would facilitate this expansion and all these former examples provide some practical ground for the use of list processing.
4.3.2 Summary of the Australian Code

In comparison with its 1981 counter-part, the 1974 version of the standard consisted of 155 features categorized according to the same four categories of culture, hydrography, relief and vegetation. Typically, expansion involved the addition of new features and/or sub-categories of features, and although not recognized as such within the standard, the addition of feature modifiers, each assigned a unique feature code. The largest increases relate to buildings (by activity), transportation (type of roads, surface characteristics, proposed function, related features), and the recognition of new classes and sub-classes of hydrographic features. The 502 new additions stress the importance of being able to add new items to a classification. Although based on the same concept of the eight digit feature header, the four digit feature codes attached to all the features in the 1974 version were completely renumbered.

Without explicitly stating so, the standard is operating under the assumption of scale-independence. Concepts of derived generalization and feature representation are not mentioned.

It is evident that, for a large part of the 1981 standard, an attempt has been made to classify and code (thus preserving any semantic ordering achieved through the classification effort) logically related groups of features and properties. The implementation of the coding structure treats the data as one level and does not alter the structural organization achieved through its classification mechanism.

\[\text{Standard for the Exchange of Topographic Data on Magnetic Tape, December 1974.}\]
4.4 THE ORDNANCE SURVEY CODE

The fourth system reviewed was the feature coding schema utilized by the Ordnance Survey, an organization which has had a long history as the central survey and mapping organization in Great Britain. In 1972, the OS introduced a pilot production system for large-scale digital mapping with the primary purpose of applying automated methods to reproduce field survey material in an acceptable graphic form (i.e., the paper map). By 1978 approximately 8000 map sheets at the basic survey scales 1:1250, 1:2500 and 1:10,000 were coded in digital form and between 1978 and 1982 a further 9,533 of these maps (bringing the total to 17,533) were digitized. As of September 1983, a total of 21,845 map sheets had been digitized. Although the storage and encoding of these graphic digitizations largely limits use of the digital topographic data to non-analytical purposes, a study of the OS feature coding system is important for it entails discussion of problems concerning digitization, feature tagging, representation, and generalization.


4.4.1 Description and Analysis

Table 12 illustrates only a small sample of Ordnance Survey feature codes. The complete coding system consists of 113 items grouped into various categories such as buildings, boundaries, utilities, centre lines, and graphic symbols. The first feature in the list receives the code 1, subsequent feature codes are incremented by 1. Several of these categories are not consistently coded (for example, codes 14/22 identify a group of road and railway features, yet codes 83/102 identify the same group of roads and railways, but this time as the centre lines), which leads the author to believe that for OS purpose's, the logical classification of real world features is not of prime importance, but rather, the basis of the coding system is.

Initiated simultaneously with the pilot production project in the early 1970's, the foundation of the OS feature coding system resulted from analysis of the specifications for line and symbol detail at the basic mapping scales with the goal of constructing a list which included, not only all the features shown on the OS maps, but also every possible way of depiction. However, its implementation without ensuring that the digital data was capable of producing adequate cartographic representation led to problems in generalization and symbolization, largely due to the unplanned need of the data to serve as a basis for: 1) selected thematic mapping; and 2) the development of derived digital maps. These problems arise from the fact that, for each scale of graphical map output, there may be different data requirements.

---

### TABLE 12
A Sample of Ordnance Survey Feature Codes

<table>
<thead>
<tr>
<th>Code</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Building - Public</td>
</tr>
<tr>
<td>2</td>
<td>Building - Minor</td>
</tr>
<tr>
<td>6</td>
<td>Archway Symbol</td>
</tr>
<tr>
<td>7</td>
<td>Boundary - Parish or Community</td>
</tr>
<tr>
<td>14</td>
<td>Railway - Narrow gauge</td>
</tr>
<tr>
<td>15</td>
<td>Railway - Standard gauge</td>
</tr>
<tr>
<td>22</td>
<td>Road - Centre line</td>
</tr>
<tr>
<td>24</td>
<td>Minor Control Points</td>
</tr>
<tr>
<td>29</td>
<td>Road fence, wall etc. (casing definitive)</td>
</tr>
<tr>
<td>30</td>
<td>Fence, Wall etc. - Non-Road</td>
</tr>
<tr>
<td>31</td>
<td>Road Pecks (casing definitive)</td>
</tr>
<tr>
<td>32</td>
<td>Surveyed Pecks (banks, baulks, made paths, driveways etc.)</td>
</tr>
<tr>
<td>57</td>
<td>Point Features - Dot</td>
</tr>
<tr>
<td>59</td>
<td>Bank of Double River/Stream</td>
</tr>
<tr>
<td>64</td>
<td>Single Stream</td>
</tr>
<tr>
<td>80</td>
<td>Track - Centre line</td>
</tr>
<tr>
<td>97</td>
<td>Centre Line Minor Single Carriageway &lt; 4m wide</td>
</tr>
<tr>
<td>98</td>
<td>Centre Line Minor Other Roads</td>
</tr>
<tr>
<td>100</td>
<td>Centre Line Railway - Multiple Track</td>
</tr>
<tr>
<td>288</td>
<td>European Constituency Boundary</td>
</tr>
<tr>
<td>314</td>
<td>Other Building (Below Upper Level of Communication)</td>
</tr>
</tbody>
</table>

**SOURCE:** "Digital Data - Feature Codes and Description," 1980.

to achieve such a representation. The obvious example is the road and track centre-lines which do not appear on the survey document and are thus not required for map reproduction at survey scale, but are essential if one is to produce generalized conventionalized roads at derived scales.\(^\text{130}\) Several further related problems were identified subsequent to code implementation: 1) inadequacies associated with digitizing and holding boundaries as independent lines for the

\(^{130}\) Bell, 1978, p. 8.
derivation of small scale maps; and 2) the need to re-assess and consider in greater detail solutions to the coding for depiction of certain linear features such as paths, tracks, road detail lines, and boundary property lines, so as to be able to output complete and unaltered networks. Because the whole concept in OS basic mapping is plotting to scale (with generalization being limited to the omission of minor detail), both sides of many of the latter features (which often vary in width) are digitized. Yet when displayed graphically, particularly at scales approaching 1:10,000 and smaller, problems were encountered. This led to the introduction of new feature codes, such as road, railroad and track centre-lines. These features, plus others such as building delineation and text, previously coded with particular line thicknesses, point size and pecking styles, were incompatible with the needs of derived 1:25,000 scale digital maps, and their codes had to be altered. Bell noted that any sensible feature code list makes provisions for amending the list but it is not enough to identify these requirements as they are encountered. Every time the feature code list was changed the banked data were put in error and usually the only way of correcting them was to go back to the digitizing table.

The OS coding schema does not include any parameter coding such as pointers to associated features, left and right coding, or scale dependencies. It is also restricted to one attribute per feature.

131 Gibson, 1975, pp. 121-132.
Text is added in an elementary way and there is no logical link to the feature to which it applies. Unsurveyed information, such as slopes and vegetation, is added manually (there is no area handling facility) and is not included in the digital data.  

Certain codes do have special meanings, such as code 2 (minor building) which is classified as such simply on size, the code being incorporated to allow omissions of tiny buildings at 1:10,000 and smaller scales. Coincident coding is provided for from a purely graphic perspective, in that rules are provided for the digitizing of features (such as minor buildings) sharing complete or partial sides with another feature, (eg. a fence). Although these rules exist, sliver errors do occur as a result of the selective plotting of only one of the coincident features.

Digital map data at the basic survey scales accompanied by a program to plot the data (including full specifications of the map data and a sample set of parameters) are available to customers on magnetic tape. Selection (or rejection) of the data to be plotted is by feature code (or optionally by OS serial number) used in association with a graphic 'feature description table' for specifying the type of output required (eg. solid or broken lines, line width or colour, form of symbol). Map data are charged per map unit but there is no charge for the plot program (which is retained in the ownership of OS).

Earlier versions of the code explicitly contained much of the attribute information in the code itself, such as the 5-digit feature

---

134 Fraser, 1982, p. 29; idem, 1984, p. 1.

codes discussed by both Williams, and Gardiner-Hill. For example, the code 40800 identified a stream, the code 81124 a building. In the code 81124, we know that it is a name (8) of class building (1) in Gill Sans Roman type (1) with a letter size of 24 point type (24). This type of feature descriptor information is now provided in the exchange record format of the 1980 OS codes, and is implicitly associated with the selected feature code.

4.4.2 Summary of the OS Coding System

Although the OS has pioneered the application of digital techniques to large-scale map production, several issues have evolved which require major policy and investment decisions concerning whether, how fast, and in what way to provide a complete digital topographic data base for Great Britain. One such issue revolves around the emphasis on the replication of the conventional map product, the growing need for more topologically structured data (especially in the longer term), and the requirements of major users in terms of data organization. The call for a scale-free data base is another issue. Although the OS Review Committee recognized that the long term objective should be a scale-free data base, concern was raised to the fact that "a scale-free data base has to be derived from accurate ungeneralized data which can only be provided by digitizing the basic survey scale, and the smaller scales,

---


such as the 1:50,000 scale, being manually generalized, cannot be used
for this." Bell's research suggests the need for a family of scales, an
observation which reflects on the utility of a scale-independent data
base for cartographic representation.

More recently, a set of 222,566 paper topographic maps (54,365 at
1:1250, 158,020 at 1:2500; and 10,181 at 1:10000), has been completed
for Great Britain and will form the basis of a digital data base.
Several new approaches to digital data collation, structuring, and
research are also under development.\textsuperscript{13}\textsuperscript{1} The addition of feature codes to
linework is still a very major component of the whole digitization task
and related developments include the introduction of a multiple feature
code facility, and looking further into raster technology, hierarchical
approaches to automated feature recognition through the categorizing and
assignment of simple groups of feature codes (such as roads, buildings,
or waterways) for easier validation by an operator afterward.\textsuperscript{13}\textsuperscript{1} The
latter research is based on the concept of building complex raster
structures by assigning feature codes to the cells instead of just on or
off, and Adams proposes that raster-based cartographic data will offer
serious alternatives to existing methods when that time comes.

\textsuperscript{13}\textsuperscript{1} David Rhind, Timothy Adams, Simon E. G. Fraser, and M. Elston,
"Towards A National Digital Topographic Data Base: Experiments in
Map Digitising, Parallel Processing and the Detection of Change,"
Canadian National Committee for Auto-Carto Six, Automated
Cartography: International Perspectives on Achievements and
426-437. Earlier work generating statistical data for planning
purposes concerning the projected size and characteristics of a
national digital topographic data bank can be found in Adams and

\textsuperscript{13}\textsuperscript{1} Fraser, 1982, pp. 29-34; idem, 1984, p. 2; Adams, 1982, pp. 117-126.
4.5 CANADIAN HYDROGRAPHIC SERVICE FEATURE CODES

One of the first Canadian Governmental agencies supporting research and development in automated cartography, chart production is now viewed by CHS management as being competitive with conventional cartographic methods. Their main products are the navigational charts (natural resources and fisheries charts account for about 10% of the CHS output) at scales ranging from 1:300 - 1:500 to 1:2500 - 1:5000 (harbour sites); 1:10,000, 1:20,000, 1:25,000 and 1:50,000 (large scale navigational charts of the coastal waters); 1:80,000 to 1:500,000 (smaller scale charts); and 1:1,000,000 to 1:5,000,000 (the smallest scale, general purpose charts). Similar to the purpose of the OS feature coding system, the CHS feature codes are designed purely for paper chart production and are based on a listing of all hydrographic symbols and linework currently appearing on CHS charts. 144

4.5.1 Description and Analysis

Initiated in the early 1970's, the first hydrographic feature classification consisted of approximately 150 features. Based on a taxonomy of categories of land, water, depth contours etc., each alphanumeric code was a mnemonic for the word(s) describing the feature and of four characters in length. Maintenance of the classification required continual updating and although initially intended to be tree structured, two problems manifested themselves and the classification became more complex and less logical over time: 1) the divisive logic of

144 Canadian Hydrographic Service, "Feature Code Manual," Cartographic Research, Fisheries and Oceans Canada, Ottawa, latest version April 5, 1983. Several versions of the coding manual were provided courtesy of Mr. George Czartoryski, CHS, Carto-Research, Ottawa.
the classification became questionable when faced with the recurring
dilemma of choosing to which exact category a particular item belonged;
and 2) the classifiers themselves soon began to run out of logical
abbreviations and often found that the logical mnemonic for a new item
was already in use. Because the classification was implemented with an
incomplete list of items, both of these problems became more acute.

As of November 1982, there were approximately 500 codes in the
classification. On January 1, 1983 approximately 150 new items were
added to the category of aids to navigation with the Canadian adoption
of the standard buoy system. A large part of the feature code manual had
to be redrafted and approximately 100 old codes were deleted. In
retrospect, the updating and modification of the original classification
of 150 items has been somewhat crisis orientated in approach and
emphasizes the importance of compiling a complete list of all the items
to be represented in the final system, if a conceptual framework
providing uniform criteria for their differentiation is to be adhered
to. Many items appearing on hydrographic charts are not so easily
categorized, and problems in maintaining the logical consistency of the
hierarchy would of course be accentuated by, as examination of the coding
manual illustrates, the different types of items being classified, some
of which are real world features, some are symbols to represent these
features, and some are just symbols.

The CHS automated system is orientated to drafting and is working
well in the production environment. There is in this approach however a
duplication of resources (Federal Government resources for example) in
the data collection surrounding the land-water interface. This is no
trivial point for Canada's coastline is large indeed. Further, the system as presently oriented, is locked in, and will continue to be so, as far as any spatial data transfer is concerned between other potential data base users such as Environment Canada, Topographic Survey, Aeronautical Charts, or any other agency or user in the future. For instance, data transfer between the NTS and CHS concerning the geographic features common to both mapping agencies at the land-water interface is limited for several reasons: 1) the lack of standardization in registration system, encoding technique, different datums (example, Great Lakes) and different definitions of shoreline water levels; 2) CHS requires data description with their feature classification concerning data accuracy, quality and time of collection; and 3) feature codes and formats are not compatible.

Concerning the actual production of navigation charts, symbolism (i.e. feature representation) is accomplished by two methods: 1) flash disk (largest available disk on the market has 96 symbols of which practically, 75 is a good maximum); and 2) call generated (by x, y coordinates). There are four colours to chart overlay: blue, black, magenta and brown. If the CHS feature classification had to be started over, and if the mandate was to be limited to chart replication, then the most logical structure of the feature classification would be the most convenient for production based on final chart overlay, and the mnemonic choice to represent features by category would be, for example, blu, blk, brn, mg. Within this highest level, features could be logically structured according to land, water, contour etc.141

141 It is important to realize that the case against a national (or international) standard in favour of one that is "most convenient
Encoded (hydrographic) spatial data are stored by charts and can be retrieved by: 1) feature code; 2) type of data (e.g. sounding, line, point); 3) source (charts being composed of more than one source by section); and 4) through latitude and longitude but within chart framework (data are recorded to geographic values). Because the CHS orientation is to replicate the paper chart however, several manipulative and processing capabilities associated with digital mapping and the collation of a digital data base are forfeited, for instance: 1) there is a limitation on the capability for retrieval and display of classes of features (there is no requirement at present) such that there is a two character limitation for identifying the feature class one wants retrieved; 2) generalization capabilities are essentially through the assignment of a unique feature code for each feature at its required scale (i.e. data are scale dependent); and 3) parameter, coincidence and multiple feature coding capabilities are not developed. For example, suppose at a unique geographic location there is a rock underwater, a buoy to mark the rock and a water depth above the rock. A coordinate is digitized and a feature code for underwater rock recorded. The same x,y coordinate is again digitized, but this time the feature code for a buoy is recorded. Next the same x,y coordinate is digitized and the feature code for the water depth above the rock is recorded. The capability exists in the data base (for this rather lengthy approach) but in production this encoding is messy and a new code is required in the data structure to indicate the real position of the features.

"for production" results in a structure which impedes data interchange.
4.5.2 Summary of the CHS Feature Coding System

One approach to a new hydrographic feature classification would be to provide a cross table of real world features (versus their graphic expression) from many hydrographic feature specifications and classifications. This type of comparative study would be useful to many agencies for several reasons: 1) its methodology and results would probably make a valuable contribution to the progress towards compatible feature classification for spatial data exchange purposes; 2) it would provide a basis for a second phase of research needing to be explored: that of determining the practical opportunities, limitations and problems associated with an integrated NTS - CHS feature classification system. Such an investigation could consider the needs of other data bases that are presently being established, such as the Beaufort Sea data base, by agencies like Environment Canada; and 3) the study could also concentrate on, or provide a basis for identifying the future role and needs of hydrographic data for their usage in real time cartographic information systems, such as the concerns and needs discussed by Eaton, Anderson and Evangelatos, concerns which are very different in orientation and purpose, from those of constructing a feature classification system for paper chart production only.\(^2\)

---

4.6 THE CECM CLASSIFICATION

The Task Force on Surveying and Mapping (1978) recommended that "as early as possible, Topographic Survey should create mechanisms for the formulation of standards for digital mapping, including the storage and retrieval of digital map data, involving federal, provincial and municipal government, the universities and the private sector." Under the auspices of the Canadian Council of Surveying and Mapping (CCSM) three committees were set up and began meeting in April 1979. Technical committee no. 1 (whose proposed classification forms the basis of this analysis) was assigned the task to develop a standardized classification for the encoding of topographic features. Technical committee no. 2 had the task of standardizing accuracy and resolution parameters such as determining what parameters to go on tape and to explain how these parameters were arrived at (basically to ensure the same ruler of measurement). Committee no. 3, concerned with electronic data processing (EDP), had the task of developing standards for how to write data onto tape. Over 700 copies of the three Draft Reports were distributed to approximately 22 countries in early July, 1982:

The committees themselves were composed of representatives from various federal, provincial and private industry interests. Municipal interests were the responsibility of provincial representatives. The secretariat for the three committees was to maintain a neutral role. Sixty percent of the meetings were held in Ottawa with the remaining

---


forty percent being held at various locations across the country. It is worthy to note at this point that although the classification is under the auspices and guidance of Topographic Survey, the responsibility for setting up a national classification/coding system for the exchange of digital spatial topographic data rests on the Digital Mapping Standards Committees. Also worthy to note is: 1) contrary to Recommendation 3.6 of the Task Force Report, universities were not involved with the Digital Mapping Standards Committees; and 2) the validity of municipal interests being represented adequately by Provincial representation on the three committees is questionable.\textsuperscript{144}

4.6.1 Description and Analysis

Table 13 provides an overview of the proposed CCSM classification. It consists of 1556 fixed-length 10-digit alphanumeric codes, hierarchically arranged according to the four levels of class, category, feature and attribute. Level I, identified by the first alphabetic character in the code, consists of ten classes (designated

# TABLE 13

## Samples From the Canadian Classification

<table>
<thead>
<tr>
<th>Code</th>
<th>Class (level I)</th>
<th>Category (level II)</th>
<th>Feature (level III)</th>
<th>Attribute (Level IV)</th>
</tr>
</thead>
<tbody>
<tr>
<td>GA 00050 000</td>
<td>Hydrography *</td>
<td>Aboideau</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>DB 00100 000</td>
<td>Road/Railway Accessway</td>
<td>Accessway(gravel)</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>AL 24350 000</td>
<td>D.A. **</td>
<td>Riding Academy</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>AR 24400 000</td>
<td>D.A. ***</td>
<td>Right of way</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>GA 24450 000</td>
<td>Hydrography *</td>
<td>River/stream</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>GA 24450 110</td>
<td>Hydrography *</td>
<td>River/stream braided</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>GA 24450 120</td>
<td>Hydrography *</td>
<td>River/stream disappearing</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>GA 24450 130</td>
<td>Hydrography *</td>
<td>River/stream indefinite</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>GA 24450 140</td>
<td>Hydrography *</td>
<td>River/stream intermittent</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>GA 24450 150</td>
<td>Hydrography *</td>
<td>River/stream mainly dry</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>GA 24450 160</td>
<td>Hydrography *</td>
<td>River/stream perennial</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>DA 24500 000</td>
<td>Road/Railway Roadway Rd. gr. div.</td>
<td>-</td>
<td>-</td>
<td></td>
</tr>
<tr>
<td>DA 24500 110</td>
<td>Road/Railway Roadway Rd. gr. div. 1 lane/way</td>
<td>-</td>
<td>-</td>
<td></td>
</tr>
<tr>
<td>DA 24500 120</td>
<td>Road/Railway Roadway Rd. gr. div. 2 lane/way</td>
<td>-</td>
<td>-</td>
<td></td>
</tr>
<tr>
<td>DA 24500 130</td>
<td>Road/Railway Roadway Rd. gr. div. 3 lane/way</td>
<td>-</td>
<td>-</td>
<td></td>
</tr>
<tr>
<td>DA 24500 140</td>
<td>Road/Railway Roadway Rd. gr. div. proposed</td>
<td>-</td>
<td>-</td>
<td></td>
</tr>
<tr>
<td>DA 24500 150</td>
<td>Road/Railway Roadway Rd. gr. div. UC</td>
<td>-</td>
<td>-</td>
<td></td>
</tr>
<tr>
<td>DA 24500 160</td>
<td>Road/Railway Roadway Rd. gr. div. UC 1 lane/way</td>
<td>-</td>
<td>-</td>
<td></td>
</tr>
<tr>
<td>DA 24500 170</td>
<td>Road/Railway Roadway Rd. gr. div. UC 2 lane/way</td>
<td>-</td>
<td>-</td>
<td></td>
</tr>
<tr>
<td>DA 24500 180</td>
<td>Road/Railway Roadway Rd. gr. div. UC 3 lane/way</td>
<td>-</td>
<td>-</td>
<td></td>
</tr>
<tr>
<td>DA 24550 000</td>
<td>Road/Railway Roadway Rd. gr. und.</td>
<td>-</td>
<td>-</td>
<td></td>
</tr>
<tr>
<td>DA 24550 110</td>
<td>Road/Railway Roadway Rd. gr. und. UC 1 lane</td>
<td>-</td>
<td>-</td>
<td></td>
</tr>
<tr>
<td>JA 33100 000</td>
<td>Landcover Woodland Wooded area</td>
<td>-</td>
<td>-</td>
<td></td>
</tr>
<tr>
<td>JA 33100 110</td>
<td>Landcover Woodland Wooded area coniferous</td>
<td>-</td>
<td>-</td>
<td></td>
</tr>
<tr>
<td>JA 33100 120</td>
<td>Landcover Woodland Wooded area deciduous</td>
<td>-</td>
<td>-</td>
<td></td>
</tr>
<tr>
<td>JA 33100 130</td>
<td>Landcover Woodland Wooded area mixed</td>
<td>-</td>
<td>-</td>
<td></td>
</tr>
<tr>
<td>GE 33150 000</td>
<td>Hydrography RHF. Wreck</td>
<td>-</td>
<td>-</td>
<td></td>
</tr>
<tr>
<td>AB 33200 170</td>
<td>D.A. Commercial Yard storage</td>
<td>-</td>
<td>-</td>
<td></td>
</tr>
</tbody>
</table>

**Abbreviations:**

* Water course or related feature.
** Recreation, cultural, historical or ornamental.
*** Administrative, political or cadastral.
D.A. - Designated Area.
RHF. - Related Hydrographic Feature.
UC - under construction.

**SOURCE:** CCSM, Standards, Draft Report of Committee 1, 1982.
area, building, structure, roadway and railway, utility, delimiter, hydrography, hypsography, land cover, textual or node information. In Table 13, the class hydrography is identified by the letter 'G' in the code.

Level II, identified by the second alphabetic character in the code, consists of a total of 68 categories throughout the classification, but is limited to 26 categories within any particular class (A-Z). This limitation could cause problems in the future expansion of the classification concerning the class of designated area (15 categories at present), building (14 categories at present), and structure (10 categories at present). Other classes could potentially encounter problems but it is the three classes mentioned that appear most susceptible considering their broad cultural interpretation and the associated generalization of each category. In Table 13, the category recreation, cultural, historical or ornamental is identified by the L in the code (one category of class designated area).

Level III consists of a comprehensive list of 664 topographic features which are each identified by a unique five digit feature code. The first feature coded is an aboideau (00005). Features are arranged in alphabetical order and each feature receives a code incremented by a value of 50 (thereby facilitating the opportunity for insertion of 49 additional items). If by chance however the committee missed 50 features between two consecutive features then the rectification would be to differentiate at the attribute level between 2 consecutive features (if possible). It should be pointed out that in order to achieve these 49 additions, these new features would have to arrive in the order that they are to be inserted.
The attributes of a feature are used to differentiate between, and describe the characteristics of, a particular feature. Despite the fact that the CCSM Draft Report states that the classification has the flexibility of allowing 999 additional descriptors to the code at the attribute level per feature, this research revealed that this is only selectively the case because although technically correct, the logical interpretation is misleading. Assuming that the arrangement of the attributes in the classification is based on a logical relationship to the features, the classification for many of the attributes is limited to the addition of only 9 additional items. One example illustrating this point is the case of wooded areas (see Table 13). There are only 9 (or 19) possible attribute values that can be assigned to deciduous wooded areas because of the upper and lower bounds of codes for coniferous and mixed wooded areas. Further, in order to achieve the additional descriptors, these new descriptors have to arrive in the order that they are to be inserted.

4.6.1.1 The Importance of the Review Task

Contrary to Recommendation 3.6 of the Task Force Report, universities and municipal interests were not adequately represented. That is, in a fully multi-dimensional approach, as originally suggested by the Task Force Report, all the related players and ideas should be incorporated in the initial design of a classification exercise. The committee's first task, "to examine existing classifications (topographic features, geography, geology, forestry, pedology, land use and ownership, etc.) to discern patterns of compatibility, areas of
difficulty, etc.," in itself, is a huge task.

4.6.1.2 The Importance of a Uniform Classification

In terms of reference, the committee's second task was to "investigate the feasibility and practicability of a single uniform scale-free national classification system for topographic features." The CCSM standard has provided a comprehensive classification of geographic features, yet at several places in the classification objects which differ in kind have been intermingled in data series. Many features are arranged into categories and classes on mixed criteria, intermingling land use, activity, ownership, function (boundaries, corner, node), natural qualities (eg. appearance - distorted surface), and even, in selected cases, graphic representation (flow arrow, hachured area, relief shaded area). Exact and definitive criteria need to be uniformly applied throughout the CCSM classification system (see the principles elucidated by Grigg in Chapter III). Some features, such as factories, are broken down by specific product. Other features are simply coded as built-up areas, deciduous woodland areas, or barren lands. The class of designated area (37 % of the CCSM codes) for example is defined as "an area set aside for a particular use or purpose." This is an inconsistent application of a 'set theory' approach applied on the basis of spatial partitioning. Every feature in the classification is defined by some particular use or purpose. If consistently applied, the class of designated area would provide a logical framework for the whole CCSM

---


classification.

As proposed, the logical consistency of the classification breaks down. For example, several categories of the class of land cover logically belong to the class designated area — agriculture (features such as crop land, orchards, pasture, rangeland). An ecological area, a forest reserve, a wildlife sanctuary or a wilderness area are features found classified under designated area — conservation. Yet, the very components of these designated areas, such as a natural clearing, reforested area, wooded area, brush, scrub, and tundra are found classified under various categories of land cover.

4.6.1.3 The Problem of Devising a Flexible Code

The third task of the committee was to devise a coding scheme for the classification which possessed the following qualities: 1) the flexibility to add and/or delete features and/or attributes; 2) the flexibility to regroup; 3) the ability to incorporate user-defined attributes; and 4) ease of use.

Cited throughout this research are various remarks concerning the limitations imposed on many of the 'ideal' qualities requisite of a uniform feature classification. One such limitation, imposed by the Canadian classification, exists in the category of transportation at the attribute level. Information concerning the status of proposed roads (such as one-way, two-way, divided etc.) can not be incorporated within the hierarchical structure of the classification.¹⁴

¹⁴ Allam noted that the committee debated for half a day and concluded that such information could not be revealed because of political liability (i.e. if a map was produced designating particular plans for the construction of a transportation link prior to a court or
There are several interesting inferences to be derived from this limitation: 1) the data associated with the standardized classification are less valuable for they do not provide information concerning the status of the proposed road; 147 and 2) if the classification code was structured to provide this information, and it was vital to withhold for legal-political reasons, it could be withheld. 148 More important than the above however, is the class of coding problems that this one example represents with respect to the fixed-length limits of the Canadian standard. The concept of open-endedness is not really valid if the code is limited to the four levels as presently structured.

The explanation for the above problem relates to one of the ruling principles of taxonomy handed down to us through Aristotle: in logical division the division should be exhaustive. 151 Many attributes in the CCM classification are 'clumped' together as conglomerations as these branches are terminated because the code limits the classification to four levels. Analysis reveals that logically there are more than four levels to the CCM classification. In the


147 This information is pertinent to researchers, planners and possibly neighborhood interest groups.

148 Conversely, one questions whether the proposed classification, as it is now structured, is flexible enough to incorporate this information in the future if policy attitudes towards the release of public information were to become more lenient, or compulsory, as some interpretations of Freedom of Information Acts maintain or suggest.

151 The main problem is that trees in a uniform classification cannot just terminate. They have to go all the way until the division (branching) is logically exhausted.
category of transportation, there are three levels of attributes alone. It is not the choice of attributes in the CCSM classification that is the problem, it is the manner in which several levels of subaltern genera have been restricted from achieving their full expression. These subaltern genera (levels of attributes) have been identified in the CCSM classification, the internal inconsistency arises as a result of terminating the divisive process and preventing the ultimate species from being reached.

4.6.1.4 Retrieval and Ranging

An advantage of the hierarchical structure of the CCSM system is that whole classes (such as buildings, hydrography etc.) of features may be retrieved. This advantage extends to the next level in that whole categories of features (such as agricultural buildings, industrial buildings etc.) within any particular class may be retrieved.

A disadvantage, which may well be due not to a weakness of the taxonomy, but rather to its implementation, occurs at the feature level. That is, within any one category (level 2) of the CCSM classification, only two possible queries are possible at the feature level: 1) display the whole category of features; or 2) display one feature. Anything in between is typically precluded because any request to display a range of features on the basis of their 5-digit codes typically transgresses many categories. For instance, the range of codes from 24400/25150 identifies the following features: a right of way, rivers, roads, road number symbols, rocks, rock outcrop, roller coasters, roundhouses, and rows of trees. A classification should be based on logical principles to
provide a sound foundation for software development. The use of arbitrarily assigned alphanumerical code does not provide this basis. The use of word look-up tables facilitated by a command scanner based on the development and implementation of list data structures could provide a better logical relationship between data selected and the terseness of required input.\footnote{It is interesting to note that Topographic Survey have their own system (correlation tables) for digital encoding and retrieval and the onus is put on the user to develop, in house, correlation tables to meet one's respective interests and requirements for ease of access and retrieval of spatial data. This aspect of the proposed CCSTM classification requires further textual explanation and qualification on the part of its designers.}

Problems associated with the use of ranging facilities in the identification of like classes of features on the basis of their attributes are related to the problems (discussed above) concerning the conglomeration of attributes, and with double coding (discussed below).

4.6.1.5 Parameter Coding and Data Integration

If the classification is to be used only to translate to and from the standardized code for spatial data exchange purposes, more textual explanation of how the translation is to be accomplished is required. Certain aspects of the translation for data exchange may not be as straightforward as the proposed document seems to suggest. Loss of information may be the mediating factor.\footnote{In comparing a set of features between the USGS, DNA, and Australian classifications, one finds great difficulty in compatibility (ie. matching definitions and terms). There are only approximately 60 features which can be translated directly from the Australian to the Canadian classification and much fewer between all four classifications. Because of their terse descriptions, features like a church, hospital, post-office, quarry, ruin, and sea-wall directly translate. But even features like a road and/or a river are...}
Several weaknesses in the implementation of the CCSM taxonomy are associated with the concept of double coding, that is: 1) many features, such as canneries, colleges, electrical power generating stations and substations, factories, hospitals, jails, and mills, receive different 5-digit codes because they occur under the class of designated areas, and (as point features) under some other class (typically buildings); 2) the attributes associated with each of these features, frequently receive different 3-digit attribute codes regardless of the fact that they have the exact same meaning; and 3) this double coding of identical attributes extends throughout the whole CCSM classification every time some characteristic of a feature changes (for example, the attributes asbestos, columbium, copper, feldspar, gold, iron ore, or lead and zinc, associated with the feature 'mine' receive different attribute codes depending on whether the mine is an underground mine or an open pit mine).\footnote{List data structures are an important concept as are logically ordered lists of features and geographic attributes. There could be many lists; such as a symbol list, a temporal dimension list, an activity list, etc., all of which would go under the guise of property lists. Several valid lists could operate in parallel, and that implies that mechanisms to point relations between these lists could be devised, perhaps expressed in the notation of Francois Bouille as a hypergraph based data structure. Francois Bouille, "Hypergraphs and Cartographic Data Structures: The HBDS System," ICA 9th International Conference on Cartography, College Park Maryland, 1976, pp. 1-18. The important point is to ensure that the criteria used in constructing these lists is consistent and entails the exact definition of terms and highest possible data measurement levels.}

A further weakness in implementation of the CCSM taxonomy is the manner in which it has incorporated a whole gamut of items extracted difficult to match exactly because of the modifying concepts which are frequently attached as part of the feature itself.
from the Standard Industrial Classification (S.I.C.). Utilization of the S.I.C. should not be limited to the selection of an arbitrary set of items, re-organized alphabetically, and implemented in a national standard which claims to have the flexibility to incorporate new items as the need arises. The items are there, acknowledge them and facilitate their use in the event that in the future, a merging of the two standards became a viable part of a more detailed data collation process. One extra coding digit would increase dramatically the utility of the CCSM standard in this direction. This is not to say that the integration of the S.I.C. is to be easily accomplished, or even that it should be integrated in the CCSM standard, but rather, if it is to be integrated, this fact should be referenced in the CCSM standard and a mechanism should be devised to incorporate the S.I.C. codes in their entirety, without destroying the order that the S.I.C. has struggled to achieve. This order is all but lost in the CCSM coding structure. The possible 999 attribute codes per feature associated with the

155 Standard Industrial Classification, Ottawa: Statistics Canada, Standards Division, 1980. The highest level of aggregation in the S.I.C. is the division. Each division represents a broad group of economic activity (agriculture, forestry, mining, manufacturing, etc.) arranged into major groups (2-digit level), industrial groups (3-digit level), and industry class (4-digit level) on the criteria of economic activity or type of product produced (i.e. "groups of producing units engaged in similar types of activity in relation to similar goods and services").

'flexibility' of the CCSM standard, although technically correct, is logically false.

The advantages of parameter coding for facilitating data relationships (coincident coding, multiple feature coding, and attribute lists) and integration (with existing classifications) were identified in the USGS, DMA, and Australian systems, yet Canada's proposed classification contains only 2 parameter codes 001 (abandoned), and 999 (other). The possibility of providing, or at least facilitating these strengths into the CCSM National standard should be further explored. In the USGS system for instance, parameter coding ties in boundaries with the Federal Information Processing Standards (FIPS) at the two and three digit level by two parameter codes. Could not the proposed CCSM standard link up with the Standard Geographical Classification in a similar fashion? The Canadian Hydrographic Service shares a considerable overlap with NTS mapping along the coastal zone. A facility was provided by the DMA standard to integrate data needs. A basic foundation of cartographic features could be provided by collating all the features found on topographic and hydrographic charts.

4.6.1.6 The Concept of Scale Independence

The proposed CCSM classification could be implemented regardless of its capability and utility to facilitate derived generalization. This is not to say how the process of 'derived generalization' can be accommodated in the proposed standard, rather it is to say that derived generalization is a very important concept. If the classification is intended to be scale-independent, what specifications are incorporated
that ensure adequate cartographic representation at various scales? A scale-independent classification (universality) is an important concept. Problems which may arise concerning its usage for derived mapping and feature depiction (as reported in the OS system) have been ignored in the CCSM classification. These problems are important and at least should be identified for they may provide the necessary foundation for the second phase of cartographic feature code development once the basic feature classification has been designed.

4.6.2 Summary of The CCSM Classification

The proposed CCSM topographic feature classification was designed for the specific purpose of providing a uniform (ie. flexible, inclusive, and open-ended) general purpose classification to facilitate the efficient storage, retrieval, and exchange of digital map data for use by federal, provincial, and municipal levels of government, the universities and the private sector. Co-requisite with the need to eliminate the expensive duplication between digital mapping organizations, the standard was constructed under the precedent that "...digital data are digital data, scale-independent and completely independent of any cartographic purpose, collated to provide fundamental measurement. The data associated with the CCSM standard are positional data. No concepts of generalization are considered because this is not the purpose of the data."137

137 Interview with Dr. J. M. Zarzycki, Head of Development Section, Surveys and Mapping, Ottawa, August 11, 1983.
The CCSM classification is a geographic classification of a large number of real world features and their properties, which, despite its many short-comings, could be viewed as the most comprehensive geographical feature classification system in the world today. No classification will be perfect for all situations, but several of the short-comings of the CCSM standard, because of their importance, have been briefly pointed out.

The major advantage of the Canadian draft of topographic features is that it has provided a solid foundation of basic topographic features for a national standard. This basis could be expanded by integrating the strengths delineated from analysis of the five other coding systems. This foundation of basic geographic features for a national digital data base would then be even stronger. Development should occur prior to any assignment of codes and symbols and should be based on sound theoretical principles.
Chapter V

SUMMARY

The purpose of this thesis was to document in detail the scope and breadth of the topographic feature coding problem. In particular, it has identified important aspects of this subject by calling attention to the numerous pitfalls associated with a precipitous approach. Six topographic feature classifications have been examined in this research and several non-trivial problems have been identified with each. Placing the feature coding problem within the context of digital mapping as a whole, it is fair to say that improvements to any of these systems can be directly associated with "attempts to reduce the cost of the system, automate more stages of the system, stimulate and encourage the use of digital data with subsequent long-term benefits to the nation, work towards a scale-free data base, improve the source, reliability and information content of the data base and reach out to the widest possible networks of users."114

Clearly, a sound theoretical perspective on the feature coding problem requires a continuum of dedicated research activities. This entails not only a stimulated intellectual process concerning logic and classification, but the integration of the former with a series of continual pilot projects in the field of interactive graphics and software development.

These ideas are founded on several basic principles:

1. topographic data, in digital form, provided they are stored and encoded in appropriate ways, offer great potential for use with other spatial data in computer-based manipulations;

2. we must not forget for a minute the great advantages digital mapping has over conventional mapping, that is: "flexibility to produce maps regardless of scale and sheet boundaries; the selection and display of features; the combination of digital data with other digital networks; variation of map boundaries; and the simultaneous deletion, update and revision of obsolete data, base map information and any map derived from it;" and

3. as McEwen and Calkins suggest, "the initial coding must contain all the relationships necessary for automated plotting."^1^6^0^

5.1 CONCEPTUAL STRENGTHS OF A CLASSIFICATION SYSTEM

Based on an analysis of the major issues surrounding the feature coding problem, it is recommended that a general purpose topographic feature classification standard should provide for or facilitate, at the minimum (see Table 14 for a summary of these characteristics):

1. terseness and exactness in feature description;

2. retrieval by class, category, feature and/or attribute;

3. the use of ranging facilities in the identification of like classes of features on the basis of the objects and/or their properties;

4. the highest possible data measurement level;

5. the capability to infinitely insert features to the classification as future requirements arise; and

6. parameter coding (integration of existing classifications, multiple geographic descriptors, and coincident coding).

---


16'o McEwen and Calkins, 1981.
TABLE 14
A Comparison of Six Major Classification Systems for Topographic Data Bases

<table>
<thead>
<tr>
<th>SYSTEM</th>
<th>SELECTED CHARACTERISTIC</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>T</td>
</tr>
<tr>
<td></td>
<td>E</td>
</tr>
<tr>
<td></td>
<td>S</td>
</tr>
<tr>
<td></td>
<td>E</td>
</tr>
<tr>
<td></td>
<td>N</td>
</tr>
<tr>
<td></td>
<td>V</td>
</tr>
<tr>
<td></td>
<td>S</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>U.S.G.S.</th>
<th>2</th>
<th>2</th>
<th>2</th>
<th>3</th>
<th>3</th>
<th>3</th>
<th>2</th>
<th>3</th>
</tr>
</thead>
<tbody>
<tr>
<td>D.M.A.</td>
<td>2</td>
<td>2</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>3</td>
<td>3</td>
<td>1</td>
</tr>
<tr>
<td>-----------</td>
<td>---</td>
<td>---</td>
<td>---</td>
<td>---</td>
<td>---</td>
<td>---</td>
<td>---</td>
<td>---</td>
</tr>
<tr>
<td>AUSTRALIAN</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>1</td>
<td>3</td>
<td>2</td>
<td>1</td>
</tr>
<tr>
<td>-----------</td>
<td>---</td>
<td>---</td>
<td>---</td>
<td>---</td>
<td>---</td>
<td>---</td>
<td>---</td>
<td>---</td>
</tr>
<tr>
<td>ORDNANCE SURVEY</td>
<td>2</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>-----------</td>
<td>---</td>
<td>---</td>
<td>---</td>
<td>---</td>
<td>---</td>
<td>---</td>
<td>---</td>
<td>---</td>
</tr>
<tr>
<td>C.H.S.</td>
<td>2</td>
<td>1</td>
<td>1</td>
<td>2</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>-----------</td>
<td>---</td>
<td>---</td>
<td>---</td>
<td>---</td>
<td>---</td>
<td>---</td>
<td>---</td>
<td>---</td>
</tr>
<tr>
<td>C.C.S.M.</td>
<td>2</td>
<td>2</td>
<td>1</td>
<td>1</td>
<td>2</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
</tbody>
</table>

NOTES: Systems are ranked as they compare to each other.
1 - LOW
2 - MEDIUM
3 - HIGH
These recommendations should be incorporated in any national standard, if the standard is to represent the forefront of research in automated cartography, a fundamental and necessary requisite in preparing users of digital spatial data for the many problems and challenges which the future holds.

It is fair to say that there is an enormous amount of research relating to the feature coding problem which remains to be done. The intractability of the problem is only presently being realized. Simple solutions to some non-trivial problems are only scratching the surface of deeper issues. The different approaches to the feature classifications analysed in this research have in large part been a response to specific mandates intended to meet internal data processing needs. The net effect of this from a global perspective, is that "everybody is building their own data dictionaries."\(^{141}\)

Analysis of these systems has illustrated some simple and recurring errors in feature classification, such as:

1. the assignment of 'unique' codes (double coding) for the same features and attributes;

2. the mixing of levels of the classification where features are coded as both features and attributes; and

3. the reduction of data measurement level, commonly occurring through the lack of explicit metric dimensions and the alphabetical arrangement of features and feature attributes.

The solution to bigger problems, such as more efficient approaches to data classification, requires dedicated research and development. Careful and sensible ordering of both feature and attribute lists would serve as a beginning. An efficient classification would facilitate many, if not all possible relationships.

If the features and their attributes are mutually exclusive of one another, both within classes and between them, every possible relationship of the encoded data to be investigated is facilitated. In the proposed Canadian standard, only those relationships that are established "a priori" can be investigated. The DMA catalog, although inherently hierarchical, provides evidence of the feasibility of integrating relational aspects and the potential that an integrated approach could operationalize. The advantages of an integrated approach should be thoroughly investigated. The need for flexibility such as that afforded by the assigning of multiple attributes, or that entailed in the concept of multi-dimensional arrays demands a high price in the initial efforts required in the establishment of a classification system, but may well be worth the rewards in terms of the long term viability of the data base.

5.1.1 Guidelines for Development

The provision of some definitive ideas on how to solve the problem of adopting a uniform classification of topographic features for the exchange of digital spatial data point to two distinct procedures: 1) the ordering of information (be it through logical division and/or classification); and 2) the coding to facilitate its practical usage.
5.1.1.1 Inventory

Comprehensive feature lists should be constructed on the basis of 'pure data series'. This factor alone will influence greatly what can and cannot be done in the way of software development and data base construction. After compiling a list of all the features to be represented in a classification system, "one must define the union of attributes for each set of common features," with the goal of achieving the highest level of numeric ordering (measurement scale) possible in an attempt to model and preserve any existing order. The implementation and facilitation of combinations of attribute values to identify precisely the attribute descriptors will result in such a classification.

5.1.1.2 Insertion

The capability to facilitate change in a classification is often dominated by the constraints of an arbitrarily imposed coding structure. It is reasonable enough to admit that the only truly exhaustive classification system is the alphabetical arrangement and any other organizational criteria, although less arbitrary and attempting to be more meaningful, put restrictions on flexibility and expansibility if they are not logically and uniformly applied. The CCSM classification is constructed on logical principles but it is necessary to construct a complete dictionary of features based on uniform criteria before getting into the coding problem, if the best approximation of the ideal system is ever to be realized. The application of a coding structure prior to

---

achieving the 'best approximation' or even a reasonable resemblance ties the system into degrees of inflexibility, and one is left struggling with a flexible code, when the first process requires a flexible classification. Once the best approximation is established, we should realize that any coding system, in its ideal form, would allow nothing less than the exhaustiveness permitted by the alphabetical ordering of words. A topographic feature classification system should facilitate the capability to insert on the least significant digit of a code. If we allow for the addition of lesser significant digits, it is possible to insert, basically infinitely. This idea is utilized in, for example, the Wylbur editing system, and for upwards compatibility of resolution, with the Telidon system.

5.1.1.3 Parameter Coding

Parameter coding, as the name implies, concerns the implementation in the coding structure of a logical association which allows the codes to acquire special meanings. In the general sense this allows measurable quantities to be assigned to some characteristic of a feature, the characteristic being identified by one or more parameter digits. Relational flexibility is acquired through this coding because the parameters, identified through the classing mechanism, are being constantly moved around. Two special cases of parameter coding are multiple feature coding and coincidence coding.

Multiple feature coding deals with the situation where more than one attribute (aspatial geographic description) is required to accurately identify or describe a feature. We should avoid putting
restrictions on the way attributes (descriptors) may interact and be used to identify an object (feature). Two of the systems reviewed (the OS and the CHS) do not facilitate multiple feature coding at all. The CCSM code makes it very difficult, if not impossible.

Coincidence coding deals with the problem of correctly tagging a feature code to a geometric description which, for all logical purposes, corresponds to two or more real world features. Instead of encoding the geometric description (for example, a line) twice, a coincident feature code(s) can be applied to the one geometric description. Polygons (in the USGS system) use coincident codes to identify corresponding features, such as the portion of a woodland surface cover which is coincident with swampland. Coincidence coding occurs also at a purely graphic level, such as in large scale mapping, where a geometric description of a feature has complete or partial significance with the graphic representation of one or more other features (the OS system). Rules and/or mechanisms should be supplied (they do exist in the USGS System) to cover coding of coincident features.

The rational for coincident coding is to achieve 'user-controlled non-redundant' storage of coordinate data and calls for the separate storage of the geometric feature description recorded once from as many aspatial descriptions that are required to identify as many features that the geometric description may be part of, or representative of.

---

143 N. Bertschinger, Proceedings, 1975, pp. 199-220. For instance, if a coordinate string belongs to a river as well as a border, the attribute information of both features are stored, yet the coordinates need only be stored once.
5.1.1.4 Separate Storage of Attribute Data

Attributes should be stored in a separate file and managed separately from the geometric data. "Store non-spatial data non-spatially keeping the cartographic data in one format, and then with a unique key have the non-spatial portion of the data base (the attribute data) referenced."144 This principle is in use by the CHS Gomads System, the Canada Geographic Information System, is one of the issues presently being addressed by the ACSM effort, and has been cited extensively in the literature.145

5.1.1.5 Ranging

Ranging facilities are most efficiently integrated into a feature coding system when we preserve the higher order achieved through our classification effort. At the feature level this allows retrieval of logically related groups (ranges) of features (one advantage of the Australian standard). Further, in the ideal system redundancy should be limited, not only as far as possible concerning duplication of features and descriptors but also in the coding of these items. For instance, if every unique attribute receives a unique code, ranging across feature classes on the basis of similar attributes is permitted.


5.1.1.6 Dimensional Classification

The question is raised if it is necessary, or even desirable to structure a classification of geographical objects (real world features) on the basis of their point, line or area spatial properties. Certainly the value of incorporating topological information in cartographic data bases has been demonstrated to be more than worth the effort of doing so. However, research indicates that the topological model may be insufficient to guarantee the integrity and utility of large cartographic data bases. Cox, Aldred, and Rhind have suggested that the dimensional classification of geographical objects on the basis of points, lines and areas, although common and having the merit of simplicity, "fails to provide a completely satisfactory framework for geographical data processing, in part because it implies that there are three immutable, qualitatively different and unrelated classes of phenomena."

For the purpose of constructing a classification of geographic features, the geometrically encoded data should be self-defining with respect to their area, point or lineal significance at the various scales at which they will be utilized. It might be redundant, and in

---


fact impose limitations on the relational utility of the data base to attempt otherwise. It is suggested that it is not necessary to structure a 'geographic classification' of features on the basis of point, line and areal dimensionality.

5.2 CONCLUSIONS AND IMPLICATIONS FOR FURTHER RESEARCH

This thesis has provided: 1) a detailed documentation of the scope and breadth of the topographic feature coding problem; and 2) a timely and thought provoking overview and analysis of six major classification systems for topographic data bases.

This research is important because the whole question of standardization in cartography is important. Traditional cartographers have been primarily concerned with the standardization of map symbols. Standards for digital cartography require new perspectives, which in turn entail a sound philosophical and methodological approach to the problem of feature classification.

This entails therefore, not just criticisms, but explanations of why and how mistakes are being made. Examination of the problem has revealed that formation of and interaction within a continuing body of individuals representing a spectrum of map making, map using, and cartographic research disciplines is necessary if cartographic feature classification is to find acceptance and be maintained. It is also evident, as a companion finding, that resolution of the problem is dependent upon a scientifically valid context in which to pursue research and applications initiatives.
BIBLIOGRAPHY


Piket, J. J. C. "Five European Topographic Maps: A Contribution to the Classification of Topographic Maps and Their Relation to Other Map Types." Geografisch Tijdschrift 6, no. 3 (1972):266-275.


White, Marvin. "A Survey of the Mathematics of Maps." ACSM, 
Proceedings, Auto Carto IV, International Symposium on Computer-

________. "Tribulations of Automated Cartography and How Mathematics 
Helps," Canadian National Committee for Auto-Carto Six, Automated 
Cartography: International Perspectives on Achievements and 
408-418.

Williams, Ernest P. J. "The Ordnance Survey System for Pre-Editing and 

Williams, Owen W. "Technology Transfer - The Next Twenty-five Years." 

Witiuk, S. "Issues in Exchanging Spatial Systems Technology." In 
Proceedings of Workshop on Current Issues in Geographic Data 
Processing, pp. 79-91. Edited by D. R. F. Taylor, Carleton 
University, Ottawa. August 1976.

Wood, Bruce. "Round Table Discussion - Report 5," Proceedings of the 
Land Related Information Systems: Municipal-Provincial Integration, 
The Interministerial Committee on Geographical Referencing, Ministry 

Wood, Bruce D. and Douglas, David H. "Cartographic Feature Coding," in: 
Barry S. Wellar, Editor, Automated Cartography: International 
Perspectives on Achievements and Challenges, vol. 2, Ottawa, Steering 
Committee for the Sixth International Symposium on Automated 
Cartography, October 1983, pp. 616-626.

Yan, Joel; Witiuk, Sid; and Fisher, Terry. "Exchanging Spatial Systems 
Technology: Related Issues and a Case Study." In Proceedings of the 
Second Annual Symposium Workshop of the National Capital Geographic 
University of Ottawa, April 1977.

Yoeli, P. "The Mechanization of Analytical Hill Shading." Cartographic 


________. "Status and Future of Computer-Assisted Mapping in the NTS 
Program," National Advisory Committee on Control Surveys and Mapping, 
Proceedings of the Conference on the State and Utility of Computer-

Zarzycki, J. M., Harris L. J. and Linders, J.-G. "Topographic-
Cartographic Data Base and Automated Cartography in Canada." 
Proceedings of the Commonwealth Survey Officer's Conference. Ministry 