

A COMPUTERIZED APPROACH TO INCREASED MAP LIBRARY UTILITY

KATE DONKIN AND MICHAEL GOODCHILD

McMaster University

ABSTRACT. The McMaster Map Library staff, in a search for a cataloguing system for a previously uncatalogued collection of maps, is currently putting into operation an automated system for the classification and retrieval of non-topographic maps. This classification occupies a standard IBM data card of 80 columns and will produce a sorted list for any of 13 fields. Each map requires only one card where 13 cards would be required to carry the same information on a standard card catalogue. Mechanical searching of these cards produces printed, sorted listings of any map under any of the 13 characteristics. This should greatly facilitate the rapid retrieval of maps for use in research and provide printed lists of holdings that could be made available to other libraries. It is hoped that this proposal will stimulate general interest in the problems of map classification and retrieval and any suggestions for the refinement of the system would be welcomed by the authors.

A library catalog is essentially an information retrieval system enabling a book to be located from a knowledge of some characteristic of the book, notably its title, publisher or author. This is conventionally achieved by filing cards according to each of these characteristics. Thus, a library with author, title and subject catalogs would have three copies of each card sorted alphabetically by author, subject and title.

A further means of locating books is conventionally available. The books may be placed on the shelves in a particular logical order. In both the Dewey and Library of Congress systems, some degree of subject sorting is achieved by shelf location. Librarians argue that a subject shelf sorting is preferable to any other as it gives very great savings in user time and energy, grouping together on the shelves books related in subject matter.

An author and title are usually unique to a book. Though at times a book may have to be classified by more than one author, or more than one alias, it is a simple matter to enter the book in the appropriate number of places in an author catalog.

A subject catalog is less simple, however.

Any book is a collection of facts and ideas, and a complete information retrieval system must allow for the classifying of the facts and ideas into the appropriate subject areas beyond the conventional single subject classification. The problem of finding a *single* subject heading for a book is responsible for bottlenecks in library cataloging departments with different catalogers disagreeing regarding the appropriate subject heading. But progress is being made: *key word abstracting* projects being used in legal and medical libraries are equivalent to multiple subject heading classification.

The number of key words abstracted must be a compromise. The larger the number of words abstracted the closer a complete information retrieval system is approached, and the greater is the amount of material which must be sifted in a search for key words. The cost of setting up such a system is high, and the time for a typical search is long. Furthermore, a search for a word will often produce so much output that little may be gained by the user of the system. Consequently, the number of key words must be limited.

With only a few subject headings or key words being entered in a catalog for any

Kate (Moore) Donkin, B.A. (Toronto) is map curator for the Geography Department and Mills Memorial Library at McMaster University. In the Geography Department she is also concluding an M.A. in historical geography. Michael F. Goodchild, B.Sc. (Cambridge) is currently working on a Ph.D. in the Geography Department of McMaster University on karst geomorphology. In that Department he is also an instructor in computer programming. MS submitted March, 1967.

book, a certain level of education is being assumed on the part of the searcher. Level of education initially affects the grouping of facts and ideas under subject headings, and also influences knowing where to look in a subject catalog. For example, consider the word "topographic" as applied to maps. A trained geographer is well acquainted with the meaning of the word "topographic", and a single subject classification is all that is necessary for the vast number of topographic maps in libraries. But to a non-geographer, the term conveys little idea of the information to be found on a map; a multiple subject classification with entries for contours, roads and even windmills would be necessary.

The problem of differential knowledge may be overcome, however, by an initial step before the catalogs are consulted. This step might take one of three forms:

1. A consultation with library staff
2. A book entitled "How to use the subject catalog"
3. A book of synonyms, such as the blue-green pages of a telephone book.

As the number of subject headings and other cataloged characteristics increases, new means of information retrieval are needed. A computerized sorting system can lead to catalogs in book form in seconds, starting with one set of punched cards in any order. For large collections, the entire collection can be searched in seconds for a certain title or subject using magnetic tape.

Once collection searching becomes a computer operation, other peripheral activities readily follow. Interlibrary searches may be made with considerable time savings by linking data banks. Ordering of materials and interlibrary borrowing may be speeded and simplified.

The McMaster University Map Library is not cataloged. This situation led to a search for the best method of cataloging possible

within our present technology. Some systems have been proposed (Murphy, 1963; Thomas, 1963; Donohue, 1964; Hagan, 1964; Stallings, 1966) but only in one case were any useful proposals made. Stallings proposed a system which draws heavily on the Library of Congress book classification system to develop a subject catalog for maps. We believe that this system is unsatisfactory and that the proposed list of map characteristics is inadequate.

The following proposals are made in the hope that they will stimulate interest and a greater interchange of ideas, and that they will lead towards a system of general applicability. The system described is currently being put into operation at McMaster University.

Classification of a map is designed to fill a standard IBM data card, which can carry 80 characters of information, either numerical or alphabetical. The 80 characters are divided among 13 fields as follows:

DATA ARRAY ON 80-COLUMN CARD

Field	Card Columns	Characteristic
1	1-20	Title
2	21-27	Publisher's code
3	28-32	Latitude of lower right corner
4	33-38	Longitude
5	39-46	Scale / 10.0
6	47-52	Library location
7	53	Type code, e.g. wall map
8	54-59	Political division
9	60-62	Date produced
10	63-74	Three subject codes
11	75	Language
12	76	Projection
13	77-80	Spare user codes

Each characteristic, or each field, is a means of access to a map. Thus, a user requiring a map of a certain subject in a certain area, say goats in Ungava, can either consult the subject catalog under goats (field 10) or the political catalog

ACHRAY	31F13	4545	7730	633	10310	513000938	
ASIA PHYSIOG PROV				200000	5040427000009455300		
ATLANTIC OCEAN		4000	2000	200000	50102	939	
AUSTRALIA STRATEGIC		4000	15400	80000	50410	81000094355055504	
AUSTRALIA VEGETATION				50000	50209	810000956	
BALFOUR TSP				2534	10307	24000095613011201	
BARI	38	4000	1721	2500	30507	5170009344101	
BATTLEFORD	267	5230	10815	3801	20604	25200091523075504	
MANITOBA SOUTH		4900	9200	10137	20501	251000952550461096110	
BELGIUM PAFT FRANCE	19	5100	300	4000	30703	534000917640665016601	
BEVERLEY TSP SOUTH P				10	103003241100		
BROWNFIELD DIST ALB	572B	5215	11000	633	20208	253000957	
BURLINGTON	30M5C	4315	7545	250	10304	249400960	
CALI	NA18		7200	10000	40504	34200094225045100	5
CAMBRIDGE DIST FTPATH				250	30203	511000936580956006502	
CANADA EAST COAST	3663	6300		22176	10103	2100009245811	
CENTRAL AMER ISLANDS		600	6200	57001	40102	33000094459006104	
CHERRY POINT ALBERTA	84D5W	5600	11900	1267	10109	2530009515100220700	
COLUMBIA RIVER BASIN	MS38	5132	11713	316	20406	260000955511051045504	
ELGIN AND KEITH	29	5700	325	633	20501	513000934560041045108	
EMO LAND USE VANCOUVER416MI	4911	12252		250	50302	2600009665601	
EMO NIGHTPOP TORONTO 40BM12	4331	7929		2500	50304	24000096610454010	
ENGLAND SOUTH	11	5030	50	25344	30107	511000934510758086601	
GASPE N NEW BRUNSWICK	14	4706	6100	500000	5010132110009447302		
GEOG FACTORS OF RECREATION				760320	402022100000960530023025701		
HAMILTON MARKET		4300	7900	6336	502012241100950510964067306		
HANNA KINDERSLEY	72NW	5000	10800	50000	20705	9645600	

Figure 1. Print-out of information sorted by title.

under Ungava (field 8). Each of these catalogs consists of lists of maps printed from punched cards sorted by one characteristic. Only one set of cards has to be punched; the catalogs are produced on a computer using a program written by M. Goodchild.

Figure 1 shows a set of maps coded according to the card layout pattern, and Figures 2 and 3 show maps after sorting for political division and scale.

The advantage of using a computer rather than a card sorter is illustrated in Figure 2. In the case of political division, and with several other fields, the information on the card is in code form. By using a code file, illustrated in Figure 4, the names can be replaced before printing, making the output more readable.

The program was written in Fortran IV, and is suitable for use in any computing centre possessing a Fortran compiler. The program requires a verbal specification of the field to be sorted, the particular file of maps to be examined, a code file if neces-

sary, and will produce a list sorted for any of the 13 fields, in the case of field 1 sorting alphabetically.

By allocating codes in a particular fashion, maps with related characteristics can be placed in logical proximity in the catalogs. Canada is allocated the political code 200000, and all subdivisions of Canada begin with a 2 with, for example, 240000 indicating Ontario and all subdivisions of Ontario having the same two first digits. In the political division catalog, the list appears in this order. If the searching for a map is by magnetic tape rather than catalog, a call to 241100 would produce a list of the maps coded 241100, then 241000, then 240000 and finally 200000. The sample of code file in Figure 4 is allocated in this manner; there is no consistent areal unit for a particular digit.

The subject codes are allocated similarly. The first three digits are arbitrary, for example, 2340 is urban land use. Successive degrees of detail in the subject are given increasing values in the final digit with 2341 denoting urban land use divided

FIELD	TITLE AND CODE	LAT	LONG	SCALE	LIB	LBEN	TYP	PBLIT	DATE	SUBJECT	CØDES	LG	PJ	SP1
1007/1	HANCOCK	-0.00	-0.00	100	100	10300	3	241100	1	-0	-0	-0	-0	-0
1007/1	EMERSON SOUTH P	-0.00	-0.00	100	100	10300	3	241100	1	-0	-0	-0	-0	-0
1007/1	WATERBURY INSURANCE	-0.00	-0.00	100	100	10300	3	241100	1	-0	-0	-0	-0	-0
2500/1	BURLINGTON	43.15	75.45	600	600	50405	3	249400	1950	5600	-0	-0	-0	-0
2500/1	CAMBRIDGE DIST FIPATH	43.15	75.45	2500	2500	30304	3	511000	1936	5809	6502	-0	-0	-0
2500/1	EMERSON USE VANCORVER	49.11	122.52	2500	2500	50302	3	260000	1966	5110	5504	-0	-0	-0
31160/1	COLUMBIA RIVER BASIN	45.45	117.13	3160	3160	20406	3	260000	1955	5110	5504	-0	-0	-0
63300/1	ACHRAY FELD DIST ALB	51.32	117.30	6330	6330	10310	3	253000	1938	5600	5104	5108	-0	-0
63300/1	BROWNFIELD DIST ALB	52.15	110.00	6330	6330	20501	3	253000	1937	5600	5104	5108	-0	-0
126000/1	ELGIN AND KEITH	57.00	113.25	12600	12600	10807	3	253000	1937	5600	5104	5108	-0	-0
250000/1	CHERRY POINT ALBERTA	46.00	113.00	250000	250000	10807	3	253000	1937	5600	5104	5108	-0	-0
250000/1	BAI MICHIPIP	49.90	116.24	250000	250000	10807	3	253000	1937	5600	5104	5108	-0	-0
250000/1	MIDLESEX COUNTY	43.00	75.00	250000	250000	10807	3	253000	1937	5600	5104	5108	-0	-0
380010/1	BATTLEBØRD	-0.00	-0.00	38010	38010	10309	3	240000	1937	7045	1201	-0	-0	-0
380010/1	BEIJLUM PART FRANCE	52.30	108.15	38010	38010	20404	3	240000	1911	2307	5504	6601	-0	-0
500000/1	KATHLEEN LAKE	51.00	127.00	500000	500000	30703	3	253000	1915	5107	6601	-0	-0	-0
500000/1	KØYANG	-37.40	127.00	500000	500000	20806	3	271000	1924	5107	6601	-0	-0	-0
500000/1	MT REVELSTØKE NAT PARK	51.00	111.40	500000	500000	20806	3	271000	1924	5107	6601	-0	-0	-0
500000/1	PELLE CLARA BUEN VIS	41.48	84.30	500000	500000	20301	3	240000	1927	5304	5505	-0	-0	-0
500000/1	SANTA TØN MARKET	21.48	95.12	500000	500000	20301	3	240000	1927	5304	5505	-0	-0	-0
633600/1	SAMUEL	51.30	73.40	633600	633600	50301	3	241100	1940	5109	6400	7306	-0	-0
633600/1	TIGHNISH	51.30	73.40	633600	633600	50301	3	241100	1940	5109	6400	7306	-0	-0
1000000/1	CALL	46.45	67.00	1000000	1000000	10807	3	213000	1945	5109	-0	-0	-0	-0
1000000/1	LANDAU AD ISAR	-0.00	72.00	1000000	1000000	40504	3	342000	1942	5204	-0	-0	-0	-0
1013700/1	MANITØBA SOUTH	49.15	112.20	1013700	1013700	30403	3	535000	1937	5102	6110	-0	-0	-0
1250000/1	SCOTTSBØRG	42.00	86.00	1250000	1250000	40206	3	251000	1922	5504	6109	-0	-0	-0
1261760/1	NIAGARA PENINSULA	42.00	79.00	1261760	1261760	40206	3	251000	1922	5504	6109	-0	-0	-0
2500000/1	CANADA EAST CBAST	36.63	82.00	2500000	2500000	10307	3	134000	1924	5107	2501	-0	-0	-0
2500000/1	PHENIX CITY	32.40	96.00	2500000	2500000	10307	3	134000	1924	5107	2501	-0	-0	-0
2500000/1	FISHINA SOUTH	50.30	20.50	2500000	2500000	30704	3	210000	1924	5107	5804	-0	-0	-0
2500000/1	ENGADIN KØTØENAY	50.30	20.50	2500000	2500000	30704	3	210000	1924	5107	5804	-0	-0	-0
3001400/1	PLOCHER CREEK	49.00	114.00	3001400	3001400	20307	3	260000	1931	4205	6601	-0	-0	-0
3001400/1	AUSTRALIA VEGETATION	49.00	114.00	3001400	3001400	20307	3	260000	1931	4205	6601	-0	-0	-0
5000000/1	HANNA KINDERSLEY	50.00	108.00	5000000	5000000	20705	3	810000	1916	6109	-0	-0	-0	-0
5068800/1	HALLØRCA	38.00	6.20	5068800	5068800	30304	3	520000	1944	5600	-0	-0	-0	-0
5700100/1	CENTRAL AMER ISLANDS	40.00	154.00	5700100	5700100	40102	3	300000	1940	5101	2507	-0	-0	-0
8000000/1	LABRISA	37.30	11.00	8000000	8000000	50410	3	300000	1944	5900	6104	-0	-0	-0
10000000/1	LIBREVILLE	20.00	11.00	10000000	10000000	30502	3	340000	1943	5505	5504	-0	-0	-0
10000000/1	PALEA	24.00	12.00	10000000	10000000	30502	3	340000	1943	5505	5504	-0	-0	-0
20000000/1	ASIA PHYSIC OCEAN	-0.00	124.00	20000000	20000000	50107	3	700000	1942	5106	2507	-0	-0	-0
20000000/1	ASIA RIVER	-0.00	124.00	20000000	20000000	50107	3	700000	1942	5106	2507	-0	-0	-0
20000000/1	ATLANTIC OCEAN	-0.00	200.00	20000000	20000000	50404	2	270000	1945	5900	2201	2504	-0	-0
20000000/1	ATLANTIC OCEAN	-0.00	200.00	20000000	20000000	50404	2	270000	1945	5900	2201	2504	-0	-0
25000000/1	SCANDINAVIA	58.00	20.00	25000000	25000000	30301	3	550000	1944	5101	4101	-0	-0	-0
25000000/1	WEST SIBERIA	52.00	88.00	25000000	25000000	50501	3	600000	1945	6104	6603	6401	-0	-0
40000000/1	LAKE ØNTARØ	43.00	75.00	40000000	40000000	50507	3	248000	1923	5201	5510	-0	-0	-0
40000000/1	LAKE ØNTARØ	43.00	75.00	40000000	40000000	50507	3	248000	1923	5201	5510	-0	-0	-0
50000000/1	GASPE N NEW BRUNSWICK	47.06	61.00	50000000	50000000	50101	3	211000	1944	7302	-0	-0	-0	-0
50000000/1	RAINFALL EUR MEAN AV	30.00	50.00	50000000	50000000	30101	3	300000	1920	5101	-0	-0	-0	-0
76032000/1	GEØG FACTØRS RECREATION	46.00	34.00	76032000	76032000	30102	2	500000	1923	5101	2302	5701	-0	-0
90000000/1	PHYSIØGRAPHIC EUROPE	46.00	34.00	90000000	90000000	30102	2	500000	1923	5101	2302	5701	-0	-0

Figure 3. Print-out of information sorted by scale.

192000ALASKA
 200000CANADA
 210000MARITIMES
 211000NEW BRUNSWICK
 212000NOVA SCOTIA
 213000PRINCE EDWARD ISLAND
 220000NEWFOUNDLAND
 221000NEWFOUNDLAND ISLAND
 222000LABRADOR
 230000PROVINCE OF QUEBEC
 240000ONTARIO
 241000NIAGARA REGION
 241100WENTWORTH COUNTY
 241200LINCOLN COUNTY ONTARIO
 241300WELLAND COUNTY ONTARIO
 241400HALDIMAND COUNTY ONTARIO
 241500BRANT COUNTY ONTARIO
 242000LAKE ERIE REGION
 242100OXFORD COUNTY ONTARIO
 242200NORFOLK COUNTY ONTARIO
 242300ELGIN COUNTY ONTARIO
 242400MIDDLESEX COUNTY ONTARIO
 242500KENT COUNTY ONTARIO
 242600ESSEX COUNTY ONTARIO
 242700LAMBTON COUNTY ONTARIO
 243000LAKEHEAD AND NW ONTARIO
 243100THUNDER BAY DISTRICT
 243200RAINY RIVER DISTRICT
 243300KENORA DISTRICT ONTARIO
 243400PATRICIA PORTION KENORA
 244000UPPER GRAND RIVER REGION

Figure 4. Print-out of Codes for Political Divisions. (Part)

REFERENCES

- Donohue, J. C.
 1964: Proceedings of the American Documentation Institute, pp. 137-140.
- Hagan, C. B.
 1964: An Information Retrieval System for Maps, UCLA Map Library.
- Murphy, M.
 1963: Special Libraries, Volume 54, No. 9, pp. 563-567.
- Stallings, D. L.
 1966: S.L.A. Bulletin, Geography and Map Division, p. 5.
- Thomas, K. A.
 1963: S.L.A. Bulletin, Geography and Map Division, No. 54, pp. 8-12.

into major categories such as retail and residential, and 2343 indicating urban land use with retail establishments differentiated.

The type code is used to indicate the form of the map; 1 means a wall map and 2 means a raised relief map, etc. The language code signifies the language in which the map was published. The projection code is used for maps with notable or unusual projections. Spare user codes are indicators of interest to a particular researcher, or of relevance to a particular course; they are not applicable to more than one library. The date of a map is taken to be the date at which the information given was collected.

At McMaster University, the coding of non-topographic maps is being done first. Topographic sheets constitute the majority of the collection but pose the smallest retrieval problem.

Listings of the codes so far allocated, and Fortran listings of the sorting program may be obtained from the authors.

RÉSUMÉ. Le personnel de la cartothèque McMaster, à la recherche d'une méthode de classement pour une collection de cartes qu n'avaient jamais été cataloguées, est à mettre sur pied un système automatisé de classement et de sélection de cartes topographiques. Ce classement utilise de fiches mécanographiques IBM à 80 colonnes et produira une liste répertoriée pour n'importe lequel de 13 domaines. Chaque carte n'exige qu'une seule fiche alors qu'il en faudrait 13 pour inscrire les mêmes renseignements sur des fiches de catalogue ordinaire. Le traitement

mécanographique de ces fiches produit des répertoires imprimés et ordonnés sous chacun des 13 titres. Cette méthode devrait faciliter grandement la sélection rapide de cartes pour les travaux de recherche et permet d'établir des listes imprimées des cartes existantes à l'intention d'autres bibliothèques. Les auteurs espèrent que cette initiative stimulera l'intérêt général dans les problèmes de classement et de sélection des cartes et ils accueilleront avec plaisir toutes suggestions visant à perfectionner le système.

ZUSAMMENFASSUNG. Die Kartenbibliothekare der McMaster-Universität sind zur Zeit dabei, ein automatisiertes System für das Einordnen und Auffinden nichttopographischer Karten zu prüfen, um ein Katalogsystem für eine unkatalogisierte Kartensammlung zu finden. Die vorgesehene Klassifizierung wird auf eine 80 spaltige Normallochkarte von IBM übertragen und resultiert in einer sortierten Aufstellung unter 13 verschiedenen Themen. Für jede Landkarte braucht man in diesem Fall nur jeweils eine Lochkarte, während man für dieselben Informationen in einer normalen Kartei 13 Karteikarten benötigt. Durch automatische Sortierung der Lochkarten erhält man eine gedruckte Liste aller Landkarten, die unter eines der 13 ausgewählten Themen fallen, und das schnelle Auffinden von Kartenmaterial für Forschungszwecke wird erleichtert. Das System liefert ausserdem gedruckte Listen des gesamten Kartenbestands, die auch für andere Kartensammlungen von Interesse sind. Es ist zu hoffen, dass der Versuch das allgemeine Interesse an den Problemen der Kartenklassifizierung und -einordnung weckt. Vorschläge zur Verbesserung des Systems nimmt der Verfasser gern entgegen.

RESUMEN — El personal de la Mapoteca McMaster, en su afán por encontrar un sistema adecuado para catalogar una colección no catalogada anteriormente de mapas, está poniendo en operación actualmente un sistema automatizado para la clasificación y recuperación de mapas no topográficos. Tal clasificación ocupa una tarjeta IBM standard de 80 columnas y producirá una lista clasificada por 13 campos. Cada mapa requiere solamente una tarjeta, en tanto que para un catálogo standard se necesitarían 13 tarjetas para llevar la misma información. La elaboración mecánica de estas tarjetas produce listas impresas, clasificadas, de cualquier mapa, bajo cualquiera de las 13 características. Esto seguramente facilitará en gran medida la rápida obtención de mapas para su uso en la investigación, y listas impresas de los mapas que pueden proporcionarse a otras mapotecas. Se espera que esta proposición estimulará el interés general en los problemas de la clasificación de mapas y su recuperación, y los autores agradecerán, además, cualquier sugerencia tendiente a refinar el sistema.