# REGIONAL NUMERICAL UNION CATALOG ON COMPUTER OUTPUT MICROFICHE

William E. McGRATH: Director of Libraries; and Donald SIMON: Systems Analyst and Computer Programmer, University of Southwestern Louisiana Library, Lafayette, Louisiana.

A union catalog of 1,100,000 books on computer output microfiche (COM) in twenty-one Louisiana libraries is described. The catalog, called LNR for Louisiana Numerical Register, consists not of bibliographic information, but primarily of the LC card number and letter codes for the libraries holding the book. The computer programs, the data bank, and output are described. The programs provide the capability for listing over two million entries. Also described are the statistical tabulations which are a by-product of the system and which provide a rich source for analysis.

Twenty-one Louisiana libraries have produced on Computer Output Microfiche (COM) a Union Catalog containing locations for 1,100,000 books. About 150,000 of these are current acquisitions (books acquired in the last two years); the rest are volumes in the retrospective collections of ten of the twenty-one libraries. The *Numerical Register of Books in Louisiana Libraries*, as the catalog is now entitled, is the second step toward what is hoped will be a comprehensive current and retrospective list of over two million volumes, the estimated holdings of the participating libraries. The first was a conventionally printed *Register* of 550,000 books, issued in 1971 and distributed to fifty Louisiana libraries.

The new *Register* is not a bibliography. It includes no bibliographic information. It is a location device for books whose bibliographic information is already known and includes nothing that is not also listed by the Library of Congress. The title was deliberately chosen to distinguish it from an older bibliographic *Louisiana Union Catalog*. All books listed in the *Register* are those having a Library of Congress (LC) card number; indeed the LC card number is the entry. The term "numerical" was chosen because we anticipate using other numbers besides the LC number—e.g., the Mansell number, and the International Standard Book Number (ISBN).

The LC card number is the most widely used book number we now have. This fact is put to good use by the Library of Congress in its own NUC— Register of Additional Locations. There are other LC number indexes, but they are not union catalogs. (The Mansell number, of course, will be very useful when publication of the NUC—Pre-1956 Imprints is complete.)

Many more titles can be represented on a page by number codes than by complete bibliographic data, at a ratio of perhaps 600 to 9. Unit costs are, therefore, much less. The first edition (1971) containing 550,000 volumes was produced for an estimated total cost of \$22,600—\$8,600 grant plus \$14,000 absorbed. One hundred copies of the *Register* were printed in hard copy form with approximate overall unit costs for keypunching, computer, travel, salaries, and printing, as follows:

	In terms of actual expenditures (grant funds)	In terms of total funds, expended plus absorbed
Per title entry	$2.5\phi$	$6.0\phi$
Per volume entry	$1.6\phi$	$3.8\phi$

The second edition (November 1972) contains over 1,100,000 volumes and in terms of the second grant, was produced on Computer Output Microfiche for an estimated total cost of \$31,200, i.e., \$10,000 grant plus \$21,200 absorbed. (Reproduction costs for the COM are negligible. For an original copy of 5 fiche, containing all 1,100,000 volumes, we were charged \$25 by a commercial firm, and for extra copies, \$3 each. Copies for distribution will be sold at a slightly higher price.) Unit costs for the COM edition are:

	In terms of	In terms of total
	actual expenditures	funds, second grant
	(second grant funds)	expenditures plus absorbed
Per title entry	$1.8\phi$	$5.6\phi$
Per volume entry	.9¢	$2.8\phi$

Unit costs computed on the basis of total costs to date suggest that they remain relatively constant from cumulation to cumulation.

The concept of a numerical register is not new. The idea was discussed at length in a proposal by Harry Dewey (1) almost a generation ago in which he espoused all the essential ideas, and again in 1965 by Louis Schreiber (2). Both argued that if the bibliographic data including the LC card number were already in hand, one could then merely look up the number in a numerical union catalog to determine a location. Goldstein and others (3) have also studied what they called the "Schreiber catalog" and have produced a sample computer printout of LC numbers. Computer output microfiche, on the other hand, was not anticipated in the original concept. It has made reproduction and distribution cheap, fast, and eminently feasible. The history of the *Register* and its rationale have been discussed more fully by McGrath (4).

# PROGRAMS COMPRISING THE UNION CATALOG SYSTEM

The Union Catalog data record is shown in Table 1. The first three fields are the familiar LC card number, and the fourth, the library location.

#### Table 1. The Data Record

(1)	(2)	(3)	(4)
ALPHA	Year or	Serial	Library
series	numeric	number	
	series	within	
		numeric	
		series	
Agr	69	2354	С

(1) Alpha series prefix – this data field may contain from 1 to 4 alphabetic characters denoting a special series.

(2) Numeric series prefix – this data field may contain 1 or 2 digits.

(3) Serial number – this data field may contain up to 6 numeric digits.

(4) Alphabetic library designation code – this field contains a preassigned alphabetic code (up to 26) designating the participating library.

The three programs which use this data record and comprise the Union Catalog System are shown in Figure 1 and described below.

# LNREDT PROGRAM

LNREDT is an editing program which examines all card input data to determine whether they are acceptable or not.

Each data field as shown above is examined as follows:

Field 1 for the presence and rejection of nonalphabetic characters, and also to determine if the alphabetic code is a member of the accepted set of codes obtained from the Library of Congress; the accepted records are transferred after checking all fields to a magnetic tape file for subsequent use; rejected data records are printed and visually scanned for the source of error;

Fields 2 and 3 for the presence and rejection of nonnumeric characters; Field 4 to determine if alphabetic.

# LNRSRT PROGRAM

LNRSRT sorts all records on the above mentioned tape file. The major sort key is the numeric prefix, Field 2. The minor sort keys in order of the sort sequence are:

Field 1—the alphabetic special series indicator;

Field 3—the book serial number;

Field 4—the library code designation.

# LNRLST PROGRAM

LNRLST is the main program which uses the sorted data tape to:

220 Journal of Library Automation Vol. 5/4 December, 1972



Fig. 1. Flow Chart of the Programs Comprising the Register System.

- a. create a single record for each unique LC number containing the library code designation of each library having this particular book;b. produce a listing of the above records in LC card number order;
- c. generate records of unique titles in combinations of libraries owning the titles;
- d. enter into a memory matrix the combinations of libraries created in part (c); combinations are then counted; each time a combination is encountered, the matrix is searched for a match; if a match is found, the corresponding matrix position is incremented by one; if no match is found, a new matrix position is created with the new combination and the corresponding count initialized to one; this routine also provides for a total count of each library's contributions plus a grand total of all libraries' contributions;
- e. tabulate, from the data compiled in (d) above, several elaborate tables of summary statistics; these statistics are described later in this paper.

The number of libraries the program LNRLST can accommodate is a variable and is entered as an execution-time parameter along with the library names and code designations. The main program occupies approximately 150,000 bytes of core memory.

### THE OUTPUT

A sample of the *Register* entries appears in Figure 2. A simple one-letter designation was used to identify each library rather than the usual National Union Catalog (NUC) designation in order to save space in the printout. These letters appear alphabetically to the right of each LC number. A typical page of the *Register* contains ten columns of up to six-digit LC numbers, with the two-digit series number appearing only once at the beginning of each series. Thus each page contains about 600 LC numbers. The latest cumulation of 1,100,000 volumes (560,000 LC numbers) consists of nearly 1,000 pages. The entire output was produced on five pieces of fiche directly from the cumulated tape. The COM program was written by the commercial firm which contracted to run it.

The computer output microfiche was issued on five 4x6 pieces in 42X. Each piece contains 208 frames and each frame contains an average of 1,126 volumes and 573 titles. The data can be produced on 24X fiche as well as roll film.

#### STATISTICAL SUMMARY

The large samples of holdings (from an initial 5,000 volumes, through successive cumulations to 90,000 and, the most recent, 1,100,000) provide an excellent data base for statistical analysis. We believe the samples may be the largest title by title comparison of monographs ever tabulated in this format. Very little analysis is presented in this paper, but the data base and its format will be explained. Even without analysis, many interesting observations can be made.

973109 973606 PS	76	4449 E 4587 E 4607 E 4690 BCEN 4729 M	9097 D 9157 AE 9236 B 9314 Z	15440 E 15503 C 15972 D 15980 E	75448 AEZ 75456 AZ 75500 EZ 75527 EJ 75535 BD
15168 112600	1	4788 E 4859 C 4891 E 4903 ACED	9611 E 9717 D 9792 BE 9944 Z	16003 E 16109 E 16141 ED 16393 A	75551 E 75578 M 75586 A 75614 E
5	77 A	4938 E 5087 BJLD 5158 AB	10294 U 10349 E 10354 Z 10357 E	16405 E 16472 E 16649 E 16681 E	75630 ELMO 75728 A 75736 Z 75779 AI
100 214 257 360	BP BE	5190 A 5296 D 5564 C 5568 E 5647 E	10361 J 10365 E 10460 E 10468 A	16728 E 16752 E 17260 CE 17567 E	75787 AE 75823 AE 75866 ABIZ 75874 EZ
407431	Â CP C	5655 A 5785 D 5813 AE	10631 Z 10645 E 10661 AE	17733 D 18103 E 18154 E	75937 ABCMN
738 876	ABCEH	5927 E . 6112 E	10716 EU 10723 Z 10774 B	19038 E 19056 E	76051 ACIOP

Fig. 2. Portion of a typical page of the computer printout showing the 2-digit 76 and 77 series, a typical prefix—PS, the serial numbers with the series, and letter codes to the right of each serial number. For example, Library A has the book 77-5; seven libraries—A, B, C, M, N, O, and Z hold the book 77-75937. Each page contains ten columns; only five are shown.

Most of the tabulations are designed to throw light on the various aspects of the overlap problem, since a decisive factor in determining the utility of the *Register* is a knowledge of the number of titles held in common by all the libraries. Over the years there has been continuing interest in overlap. Probably the first and most elaborate of the early studies was by Leroy Merritt (5), and one of the most recent by Leonard, Maier, and Dougherty (6). Continuing interest is expressed in such proclamations as that by Ellsworth Mason where he claims that materials are "being acquired in duplications that are rather staggering across the country." (7).

The following statistics were tabulated from input for current acquisitions, the most recent being a total of 90,302 volumes, rather than the retrospective and current totals in the production runs. The 90,302 volumes were acquired for the most part during the two year period, fall 1969 to fall 1971. The statistics show holdings for sixteen libraries.

# THE BASIC TABULATION—TITLES HELD IN COMMON BY UNIQUE COMBINATIONS OF LIBRARIES

The basic tabulation sections which are shown in Table 2 actually fill seven pages of computer printout. The tabulation is designed so that each unique and actual combination of libraries is separately listed, and the books held by each combination are counted. Thus, in the table, although the total number of books held in common by Libraries A and B is 127, the

8 AB ABC ABCE ABCEGHJLZ	52 10688 1 17041 2 46227 1 69111	<del>:18</del> :22	AHZ 32 AI 17 AIJLNOP 1 AIJLNOP 1	16942 7473 11626 29997	07 KU 10 KP 01 KZ 83 LM	5 65 4 37 1 83 73 86
ABCEIZ ABCEL ABCEM ABCEM2	1 56754 1 51118 1 50030 1 57792	.01		12364 17244 11764 15235	.06 LMNP LMNZ .04 LMD LMDP	8 129 1 162 2 207 6 146 1 178
ABCHIMPZ ABCHJLMP ABCIJKMPZ ABCIJLZ	1 39064 1 37581 1 39308	:01 .02	AJ 41 AJL 2 AJLND 1 AJM 3	8861 13752 24011 12664	02 LMP 02 LMZ 90 LN 01 LND	3 119 4 164 83 91 3 151
ABCL ABCM ABCD ABE	1 21932 1 21932 1 20844 1 23009 2 39874	.01	AJNP AJP 2	13152 16373 14829 12082	04 LNP 01 LNZ 70 LD 02 LDP	5 124 3 169 77 108 4 140
ABEH ABEHIJM ABEHIMNPZ ABEHJ	2 44346 1 55067 1 66188 1 48499	:83	AJPZ 1 AJZ 13 AK 2 AKZ 1	19844 16623 5271 13033	.80 LP LPZ .37 LZ	65 81 1 158 48 126 53 80
ABEHJMZ ABEHJN ABEHMU ABEHZ ABEHZ	1 60064 1 52790 1 54117 1 52108 1 52108	of oi	ALMP 22 ALMZ 2	13402 16623 21164 13890	02 MND 01 MNP MNZ	4 140 2 113 1 158 61 97
ABEL ABEL ABEO ABEZ	1 44765 1 51439 1 51842 4 47636	.02 .34 .01	ALNZ 1 ALZ 4 AM 29 AMN 29	21652 17361 8511 12802	01 MOZ 49 MP	2 129 3 175 35 70
АЗ <u>БН</u> МР АЗБС 1 АВН 1 АВН 1 АВНZ	1 28681 1 17185 3 15160 3 22922	:01	AMD 2 AMP 1 AMZ 8 AN 33	14479 11732 16273 8999	43 NO 01 NOP 01 NOPZ	45 102 2 134 1 212 2 180
ABI ABJ ABJKEP ABKNŬ	1 13453 1 14841 1 23516 1 21510 1 21510	.04 .16 .02	ANZ 7 AU 18 AUZ 4 AP 6	1490/ 16761 10676 18438 7929	69 NP NPZ 31 NZ 40 DP	52 75 1 152 38 120 37 91
ABLP ABM 1 ABN	1 18800 1 14491 2 14979	5.43 1.18 .03	AZ 678 BC 146 BCE 14	12470 12333 41519	29 02 :18 PZ	40 137

Table 2. Titles Held in Common by Each Unique Combination of Libraries

number of books held in common by them and no other library is only 52. The number of books held by Libraries A, B and Z, and no other library is 18. None of these 18 is included in the count of 52, and none of the 52 in the 18. They are mutually exclusive. But the 18, plus the 52, plus the small counts in each of the other combinations in which A and B share holdings is 127.

The percentage of common holdings for each combination is also given except when the percentage is less than .01. Thus libraries A and B have .48 percent in common of their total combined holdings of 10,688 volumes.

It is interesting to note that of the 65,535 possible combinations, in only 444 combinations did the percentage of common holdings exceed .01 percent, and in only 8 did the percentage exceed 1 percent. Of these, the highest is 5.43 percent (A and Z). This 5.43 percent means that 678 of A and Z's common holdings were held by no other library. The total of A and Z's common holdings that were also held by other libraries is 1,315, or about 10.5 percent of 12,470. Again this is the highest percentage of any combination.

### Summary of Titles Held in Common

The basic tabulation of titles held in common is summarized in Table 3. Column 1 is the number of libraries from 1 to 16 in each combination. Column 2 is the total number of titles counted in all combinations. For example, 59,907 titles exist in unique copy, thus there were only 59,907 copies (column 3), but there were only 8 titles which as many as 9 libraries held, for a total of 72 copies (column 3).

Column 4 shows that all 16 libraries contributed unique titles and that there were 117 different combinations of two libraries, out of a possible 120 (column 5). Thus there were 3 combinations of 2 libraries which had no titles in common. It is also most interesting that there were only 7 combinations of 9 libraries out of a possible 11,440, and no combinations of 10 or larger.

According to the binomial distribution, there are 65,535 theoretical ways that 16 libraries can combine (total, column 5), whereas, in this sample, only 1,198 combinations occurred (total, column 4).

Column 6 is the result of column 2 divided by column 4. Thus 3774.19 is the average number of unique titles contributed by each library. 74.92 is the average number held by any combination of 2 libraries, and 6.89 is the average held by any combination of 3.

# SUMMARY OF EACH LIBRARY'S MULTIPLICATED TITLES

The administrators of each library are especially interested to know how many of their own titles are also held by other libraries. This information for total input (i.e., for titles with LC prefixes from 1900 to the present) is given in Table 4. (Tables were also produced giving the same kind of

Column 1 No. of Libraries in Each Combination	Column 2 Total No. of Titles in all Combinations	Column 3 Total No. of Copies in all Combinations	Column 4 No. of Times a Combination Occurred	Column 5 Theoretical No. of Times a Combination can Occur	Column 6 Average Title Overlap Per Combination	
1	59 907	59 907	16	(Binomial Distribution)	3 774 19	
2	8766	17 532	117	120	74.92	
3	2,453	7 359	356	560	6.89	
4	782	3,128	360	1.820	2.17	
5	279	1.395	214	4.368	1.30	
6	84	504	75	8,008	1.12	
7	43	301	41	11.440	1.04	
8	13	104	12	12,870	1.08	
9	8	72	7	11,440	1.14	
10	0	0	0	8,008	0.00	
11	0	0	0	4,368	0.00	
12	0	0	0	1,820	0.00	
13	0	0	0	560	0.00	
14	0	0	0	120	0.00	
15	0	0	0	16	0.00	
16	0	0	0	1	0.00	
Totals	72,335	90,302	1,198	65,535	60.38	

# Table 3. Summary of Titles Held in Common by Unique Combinations of Libraries (Spring 1971 tabulation)

Regional Numerical Union Catalog/McGRATH and SIMON 225

# Table 4. Summary of Each Library's Multiplicated Titles (1900-1971 imprints)

Column 1 Library	Column 2 Library Code	Column 3 Number of Volumes Contributed by Each Library	Column 4 Each Library's Volume as a X of Total Volumes	Column 5 No. of Titles for Which Copies are also Held by Other Libraries	Column 6 Each Library's Multiplicated Titles as a % of Own Titles (Col. 5÷Col. 3)	Column 7 Each Library's Multiplicated Titles as a % of Grand Total (Col. 5÷Total, Col. 3)
Louisiana State						
Library	A	4,708	5.21	2,497	53.03	2.76
Louisiana Tech						
University	в	5,980	6.62	2,378	39.76	2.63
University of South-						
western Louisiana	С	6,353	7.03	1,932	30.41	2.13
Louisiana State Uni-						
versity-Baton Rouge	E	29,186	32.32	6,190	21.20	6.85
Louisiana State Univer-						
sity Medical Center	F	580	.64	168	28.96	.18
Grambling	G	1,606	1.77	471	29.32	.52
Centenary	H	4,472	4.95	2,061	46.08	2.28
Louisiana State Uni-						
versity-Alexandria	I	2,765	3.06	1,087	39.31	1.20
Southeastern Louisiana	I	4.153	4.59	1,849	44.52	2.04
Northwestern Louisiana	K	563	.62	230	40.85	.25
Northeastern Louisiana	L	4.891	5.41	1,980	40.48	2.19
Lovola-New Orleans	M	3.803	4.21	1.744	45.85	1.93
Louisiana State Uni-						
versity-Shreveport	N	4.291	4.75	1.749	40.75	1.93
Louisiana State Uni-		-,=• -		-,		
versity-New Orleans	0	5,968	6.60	1.783	29.87	1.97
Nicholls	P	3,221	3.56	1.048	32.53	1.16
New Orleans Public	7.	7 762	8.59	3 228	41.58	3.57
	Totals	90.302	100.00	30.395		0.01
	Average	5,644	6.25	1,900	37.78	2.09

226 Journal of Library Automation Vol. 5/4 December, 1972 information by decade and for the last two years, but are not reproduced here.)

The column labels are self-explanatory, but it may be observed that the total in column 5, 30,395, equals the difference between the total copies, 90,302 (column 3, table 3) and the number of titles held by one library only, 59,907 (columns 2 and 3, table 3).

# DISTRIBUTION OF TITLES PUBLISHED AND MULTIPLICATED BY DECADE

Table 5 shows that the very largest overlap, in current acquisitions, occurs among books with recent imprints. This is to be expected since these figures do not make any comparison to older books recently acquired by one library to those already in another library, and since the acquisition of older books is from a much larger universe than that for current books.

# Table 5. Distribution of Contributed Titles Published and Multiplicated by Decade (Titles acquired from 1969 to 1971)

Imprint Period	Number of Titles Contributed	% of Titles Contributed	Number of Volumes Multiplicated	% of Total Volumes Multiplicated
1900-1909	1,483	2.05	23	.13
1910-1919	1,049	1.45	29	.16
1920-1929	1,180	1.63	22	.12
1930-1939	1,816	2.51	74	.41
1940-1949	2,539	3.51	102	.57
1950-1959	5,353	7.40	361	2.01
1960–1971	58,915	81.45	17,356	96.59
Totals	72,335	100.00	17,967	100.00

# OTHER SUMMARY STATISTICS

The foregoing tables illustrate the kind of tabulations that can be made with this type of data. More detailed tables can be compiled, and indeed were—e.g., tables giving the percentage of books acquired for each year and each decade for each library, with ten year totals and averages. Other possibilities would be frequency distributions and summaries for clusters of similar libraries.

This material awaits analysis. We believe it contains many heretofore unsuspected insights.

# FUTURE PLANS

Since the data can be updated so readily, plans are being made to provide funds for the extraction and keypunching of LC numbers in the remaining retrospective collections of the participating libraries. These libraries contain an estimated total of two million volumes. Succeeding cumulations will be readily produced on COM. Most of the cost has been for extracting retrospective numbers from card catalogs. Once the remaining retrospective collections are cumulated, costs for cumulating current input will be negligible.

Any final catalog of course can never list complete holdings since each library has many titles without LC numbers. Those titles could be listed in more conventional form. Since they are in a minority, the expense would be far more reasonable than it would be to reproduce entire holdings in conventional form.

We have said nothing about other aspects of the project. In committee discussions, however, much has been said about the feasibility of using the LC card number to access the information in other major projects such as MARC, and possibly even the data bank in the Ohio College Library Center. Technically, it is feasible to print a conventional bibliographic catalog by matching up our LC numbers with titles listed in the current MARC tapes; pragmatically and economically, of course, it is another matter.

Other possibilities are the printing of a list of specialized holdings by accessing the subject headings on the MARC tapes, assignment of specialized acquisitions, and the gathering of information which might affect development of a joint processing center.

#### ACKNOWLEDGMENTS

This project was supported in part by the Library Services and Construction Act Title III funds administered by the Louisiana State Library.

The authors wish to give special thanks to Miss Sallie Farrell, Louisiana State Librarian, for her enthusiastic support and fine advice. We wish also to thank the other members of the L.L.A. Committee on the Union Catalog: Mr. Sam Dyson, Louisiana Tech University; Mrs. Jane Kleiner, Louisiana State University, Baton Rouge; Mrs. Elizabeth Roundtree, Louisiana State Library; Dr. Gerald Eberle, Louisiana State University, New Orleans; Mrs. Hester Slocum, New Orleans Public Library; Mr. Charles Miller, Tulane University, New Orleans; Mr. Ronald Tumey, Rapides Parish Library, Alexandria; and finally, Mr. John Richard, past president of the Louisiana Library Association, who saw the importance of the project, and who appointed the original committee.

Complete documentation for this project, including computer programs, has been deposited with the ERIC Clearinghouse on Library and Information Science (8).

### REFERENCES

1. Harry Dewey, "Numerical Union Catalogs," *Library Quarterly* 18:33-34 (Jan. 1948).

Regional Numerical Union Catalog/McGRATH and SIMON 229

- 2. Louis Schreiber, "A New England Regional Catalog of Books," Bay State Librarian 55:13-15 (Jan. 1965).
- Samuel Goldstein, et al., Development of a Machine Form Union Catalog for the New England Library Information Network (NELINET). (Wellesley, Mass.: New England Board of Higher Education, 1970) (U.S. Office of Education final report, Project No. 9-0404.) ED 043 367.
- 4. William E. McGrath, "LNR: Numerical Register of Books in Louisiana Libraries," Louisiana Library Association Bulletin 34:79-86 (Fall 1971).
- 5. Leroy C. Merritt, "The Administrative, Fiscal, and Quantitative Aspects of the Regional Union Catalog," in *Union Catalogs in the United States* (Chicago, Ill.: American Library Association, 1942).
- 6. Lawrence E. Leonard, Joan M. Maier, and Richard M. Dougherty, Centralized Processing: A Feasibility Study Based on Colorado Academic Libraries. (Metuchen, N.J.: Scarecrow Press, 1969).
- Ellsworth Mason, "Along the Academic Way," Library Journal 96:1671-76 (15 May 1971).
- 8. William E. McGrath and Donald J. Simon, LNR: Numerical Register of Books in Louisiana Libraries; Basic Documents (Lafayette, La.: Louisiana Library Association, Dec. 1972) (U.S. Office of Education) ED 070 470, ED 070 471.