

Lawrence Berkeley National Laboratory

Recent Work

Title

CARTE: A THEMATIC MAPPING PROGRAM

Permalink

<https://escholarship.org/uc/item/3v12376v>

Authors

Wood, P.M.

Austin, D.M.

Publication Date

1974-07-01

Presented at the Conference on Computer
Graphics and Interactive Techniques,
Boulder, CO, July 15 - 17, 1974

LBL-3073
c.2

RECEIVED
LAWRENCE
BERKELEY LABORATORY

JUL 1 1975

LIBRARY AND
DOCUMENTS SECTION

CARTE: A THEMATIC MAPPING PROGRAM

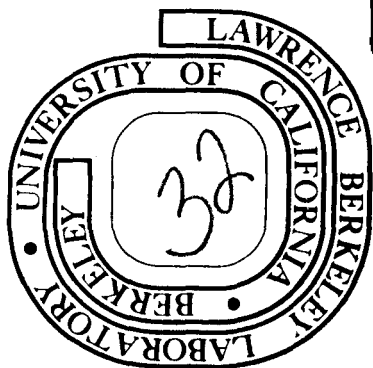
P. M. Wood and D. M. Austin

July 1974

Prepared for the U. S. Energy Research and
Development Administration under Contract W-7405-ENG-48

TWO-WEEK LOAN COPY

*This is a Library Circulating Copy
which may be borrowed for two weeks.
For a personal retention copy, call
Tech. Info. Division, Ext. 5545*



LBL-3073
c.2

DISCLAIMER

This document was prepared as an account of work sponsored by the United States Government. While this document is believed to contain correct information, neither the United States Government nor any agency thereof, nor the Regents of the University of California, nor any of their employees, makes any warranty, express or implied, or assumes any legal responsibility for the accuracy, completeness, or usefulness of any information, apparatus, product, or process disclosed, or represents that its use would not infringe privately owned rights. Reference herein to any specific commercial product, process, or service by its trade name, trademark, manufacturer, or otherwise, does not necessarily constitute or imply its endorsement, recommendation, or favoring by the United States Government or any agency thereof, or the Regents of the University of California. The views and opinions of authors expressed herein do not necessarily state or reflect those of the United States Government or any agency thereof or the Regents of the University of California.

CARTE: A Thematic Mapping Program

by

P. M. Wood and D. M. Austin

Lawrence Berkeley Laboratory

University of California

Berkeley, California 94720

ABSTRACT

CARTE is the graphics display program of the LBL Computer Mapping System, producing thematic maps on microfilm at one hundredth the cost of producing negatives by hand. The program matches a geographical area or a symbol representing such an area with statistical data to produce graphic output on 35mm film in the form of cross-hatched maps for single-color printing (either computer-generated dot screens or total mask frames for photographic screening).

A versatile set of directives allows the user to design the map and a corresponding report, and to specify such features as automatic placement of area names, calculation of a smooth distribution for color coding, and boundary clipping to specified limits for sectioning a map.

TABLE OF CONTENTS

I.	Introduction	1
	A. Components of a Thematic Map	4
	B. Program Structure and Information Flow	7
II.	The Control Module	10
III.	The Setup Module	13
	A. Clipping Limits	15
	B. Label Placement	18
	C. The Shading Algorithm	21
IV.	The Report Generator and the Data Set Specifications	23
	A. Form of the Data Sets.	25
	B. The Report Template.	27
	C. Some Features of the Report.	28
	D. Generating a Table	29
V.	The Distribution Module.	31
VI.	The Graphic Output Module	34
	A. Area Outlines, Labels and Shading	34
	B. Creating Titles, Legends and Other Cosmetics	34
VII.	Data Structure and Memory Management	37
	A. Storing the Map.	37
	B. Dynamic and Fixed Length Structures	40
	C. Memory Management.	42
	D. Random ACcess Mass Storage	44
VIII.	The Symbol Mapping Nodule.	45
	A. Introductory Considerations.	45
	B. Basic Data Structure	47
	1. Symbol Point Storage	47
	2. Data Item Elements	47
	3. Structure for Searching the Map.	48

4. The Index	48
C. Algorithms	49
1. Overview	49
2. Depth	49
3. Symbol Center and Label Coordinate Generation	49
4. Searching the map	49
5. The Clipping Algorithm.	51
6. Producing the Map	54
D. Added Directives	56
E. References	57
F. Map and Table Examples.	58

CARTE: A Thematic Mapping Program

by

P.M. Wood and D.M. Austin

Lawrence Berkeley Laboratory, University of California

Berkeley, California 94720

I. Introduction

The graphic display of statistical data by geographic area is known as thematic (or choropleth) mapping. Thematic maps have traditionally been produced using a combination of manual and photographic techniques. One technique requires cutting a sheet of textured material to fit a geographic boundary and pasting the material onto a paper outline of the map. A more recent manual technique involves scribing of the map outline of a photographically opaque material. A separate sheet is required for each data range (color tone or textured pattern). Each area in a data range is peeled from the material to form a photographic mask which is then reduced to print size.

Computer techniques for producing maps cheaply and quickly were pioneered by Harvard University's Laboratory for Computer Graphics and Spatial Analysis. The best known is the printer plot technique, which provides low-resolution maps with a minimum of special hardware [1].

With the advent of high speed, high precision devices for computer output on microfilm (COM devices), the possibility of producing high quality film images by computer became an attractive alternative to both the manual techniques and the printer plot compositions. Producing COM film images, as compared with manually prepared film images, has proven to be cheaper by almost two orders of magnitude, making thematic

mapping a resource available to a much wider range of applications than ever before.

The primary problem with computerized mapping is in building a machine-readable geographic data base. Statistical data, such as the various census counts, have long been available in machine-readable form, but a data base of geographic areas of interest, such as state, county and census tract boundaries were not generally available and typically involved a laborious hand-digitizing and editing process.

In order to develop automated cartography into a useful tool for data display, the Mathematics and Computing Group of the Lawrence Berkeley Laboratory initiated a computer mapping project in collaboration with the Manpower Administration of the Department of Labor, the Geography Division of the Bureau of the Census and the Department of Housing and Urban Development. The purpose of the project is to create an accurate geographic data base containing the boundaries and identification codes for the United States by state and county and for all the Standard Metropolitan Statistical Areas (SMSAs) by census tract (some 35,000 tracts). The automated map digitizing and editing system is described in a separate report [2].

This data base is to be used in the production of statistical atlases for use by Federal agencies in areas such as manpower resource planning and evaluating population characteristics. This report describes the display portion of LBL's Computing Mapping System, a FORTRAN program for thematic mapping called CARTE.

Given a geographic data base and a set of data keyed to geographic area, CARTE produces film images of maps and tables which are photographically enlarged to produce printer plate negatives.

There are five major functional parts to the program, as shown in Figure 1. This report describes each part in succeeding chapters.

A. Components of a Thematic Map

The objective of a thematic map is to communicate geographical relationships. In order to accomplish this goal, there are several criteria which must be met, as the following list indicates.

- a. Recognizability - The geographic area being mapped should be easily recognizable, usually by providing appropriate labels and by retaining as many geographic features (such as rivers, lakes and coastlines) as possible. Also, the scale should be chosen so that all the data is readable.
- b. Titles - The map should contain enough titling information to completely specify the meaning of the map.
- c. Clear Data Representation - The method for distinguishing data ranges on the map must provide a number of distinct patterns, either by color tones, textures or symbols. Legends should indicate unambiguously the correspondence between representation and data range.

Any program for thematic mapping must provide facilities for designing these features into a map. In addition, general design features which allow arbitrary placement of the map components is necessary because of the wide variation in shapes of geographic areas. CARTE provides a set of directives which allows the user to design both the map and a table listing the actual data that is being represented.

There are three forms of map graphics available in CARTE: cross-hatching, dot screening and total masking.

Cross-hatching is used for single-color maps or slides, and is therefore the least expensive to print. Variation of direction and spacing of the cross-hatching provides eight distinct textures on the COM device in use at LBL.

When multi-color maps are desired, usually three tones of each color are distinguishable by the human eye - 10%, 50% and 100% screens are usually used. CARTE provides a method of generating these dot screens on the COM device, so that all three tones for a single color can be drawn on a single frame. This is done by using three different dot sizes spaced on a regular array of raster points. The main advantage to dot screening is the savings in cost by combining three tones on one frame of film. Thus a 9-tone, three color map, plus black outline requires processing and registration of only four frames of film. The disadvantage in this technique is the loss of resolution caused by the dot sizes required to generate distinct tones. For example, on a device with 4096 addressable points, dot sizes of 2, 4 and 8 raster points in a hexagonal array reduce the resolution to $4096/8 = 512$.

The total masking technique requires the complete masking of each area to be a certain tone of any color. That is, for a 9-tone, three color map, plus black outline, 10 frames of film are generated. The tone screens are produced photographically after enlargement of the 35mm film to print size, thus achieving a much finer screen. The main advantage to this technique is the use of 2-raster point vector to mask an area.

This results in a resolution of $4096/2 = 2048$ raster points, a factor of 4 better than the dot screening. The disadvantage is, of course, the processing and registration of more than twice as many frames of film - not too serious a problem with pin-registered cameras and a good photographic shop.

B. Program Structure and Information Flow

The functional structure and control flow for CARTE are diagrammed in Figure 1. There are five functional modules: a control module where directives are processed, a setup module which generates the graphic data structure from the input map file, a report generation module, a distribution module which assigns representation codes (color or texture) from the data, and a graphic output module which generates the film images, including titles, legends, outlines and masks.

The program runs most efficiently when many maps of the same area are made in one job step. The initial processing of the map file done in the setup module creates a very efficient representation of the geographic area which is used in generating a series of maps from corresponding data sets - i.e., the data, title, legends, etc., change with each map, but the geographic area stays the same.

After the setup function, the map is ready to be combined with nominal, ordinal or interval data to produce a thematic map. The flow of control passes from the directive interpreter to the report module (if specified), where a film image of the input data is produced in tabular form. The distribution module then assigns a graphic code (color or texture) to each of the data items keyed to geographic areas. Control then passes to the graphic output module which creates the film images for the map.

The control loop from directives to table to graphic output is repeated for each data set provided, with the incremental cost of computer

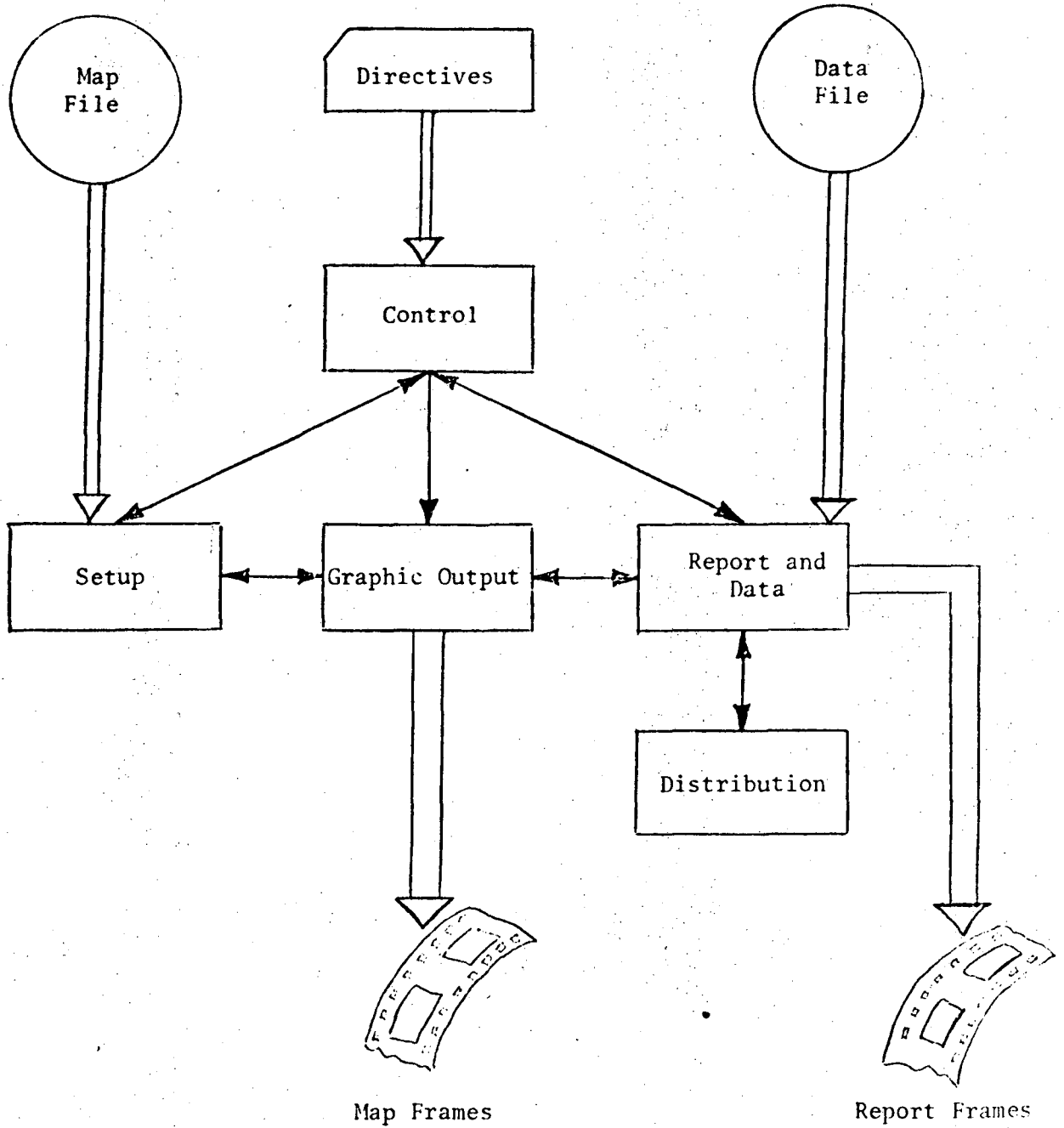


Figure 1. DIAGRAM OF PROGRAM STRUCTURE AND FLOW OF CONTROL

time per map decreasing with the number of data sets. In a typical example, the cost for producing maps of the Phoenix, Arizona SMSA (233 census tracts) was as follows:

setup	\$5	
10 maps @ \$0.50	<u>\$5</u>	
	\$10	or \$1 per map

Comparable costs for producing these film images by hand has been estimated at over \$100 per map.

The major expense in producing thematic maps, especially multi-color maps, is in the printing process, but the fact remains that preparation of film images for the printer has become economically feasible for large scale projects.

II. The Control Module

The control module interprets a set of directives which set the variables controlling program operation and allow the user to design the map layout in detail. The directives consist of a control character (an asterisk in column 1), and a keyword followed by parameters in free format. Keywords may be abbreviated, since only the first character is used.

Each directive has a number of parameters associated with it which, if left unspecified, take on default values. The parameters are numbers (either integer or floating point is acceptable) separated by commas. Following the parameters, user comments can be entered, as the scan is terminated on non-numerical data.

An example of a directive is:

```
*KEYS, 4= 8 KEY 4 on the data set matches KEY 8 on map.
```

Directives specifying titles and legends require a set of cards describing textual information to be placed on the map. Thus a packet of titles may be specified as follows:

```
*TITLES,2 There will be 2 title packets
```

```
1,1,4, (Title 1, on 1 card, size 4)
```

```
TITLE ONE FOR MAP ONE
```

```
2,1,2, (Title 2, 1 card, size 2)
```

```
LAWRENCE BERKELEY LABORATORY
```

The control character serves as an interrupt which causes the control module to terminate the processing of the current directive and examine

the next directive. Thus, to title the second map, the directive would appear as follows:

```
*TITLES,2      Still two titles
1,1,4          (Title 1, 1 card, size 4)
TITLE ONE FOR MAP TWO
*GO
```

In this case, the second title remains the same as originally specified, and the new directive (*GO) causes the reading of titles to be terminated.

The need for multiple cards to specify long titles requires that the user specify the number of cards to be read as a parameter. This overrides the interrupt features for that number of cards, so that text may begin with an asterisk. Also, the comment feature does not apply to actual text cards - only the parameter cards. There are three directives (*MAPTYPE, *INTERVALS, *WATCH) which can have character parameters and to which the comment feature also does not apply.

The control language allows full flexibility in designing the map layout with a minimum of information. Default values are provided for all parameters except textual, and parameters need only be specified when the values change.

The directives fall into four classes:

- a. Setup directives specifying information about the geographic area from the map file.
- b. Graphic directives specifying information about titles, legends and other "cosmetics".

- c. Report directives specifying information about the format of the table.
- d. Distribution directives specifying how the data is to be treated

The setup directives are as follows:

- *MAPTYPE
- *ZOOM
- *PICTURE SPACE UNITS
- *XYMAP PICTURE SPACE
- *CONSTANTS

The graphics directives are:

- *ARROW
- *TITLE
- *SCALE
- *OUTLINE
- *LEGEND
- *BOXES
- *GO

The report directives are:

- *FORMAT
- *KEYS
- *HEADINGS
- *NOTES
- *REPORT

The distribution directives are:

- *DATA
- *INTERVALS
- *EXTRACT

III. The Setup Module

The setup phase of the program processes the map file according to the setup directives to produce a data structure which allows for efficient output of the map graphics. The functions carried out on this "first pass" on the mapfile include:

- a. Clipping the map to prescribed boundaries
- b. Establishing map limits and storage requirements for the mapfile and shading arrays
- c. Addition of label boxes (rectangular areas for label placement which will not be shaded)
- d. Generating a graphic data structure representation of the map in random access mass storage (RAMS)
- e. Generating a test frame for COM setup purposes.

Some functions done in the setup module depend upon user-specified directives - e.g., clipping boundaries - while other functions are universal - e.g., memory requirements. Figure 2 diagrams these functions of the setup module. The major algorithms used in this module are discussed below. The method for storing the graphic data structure and calculating the memory requirements are discussed in Chapter VI on the data structure.

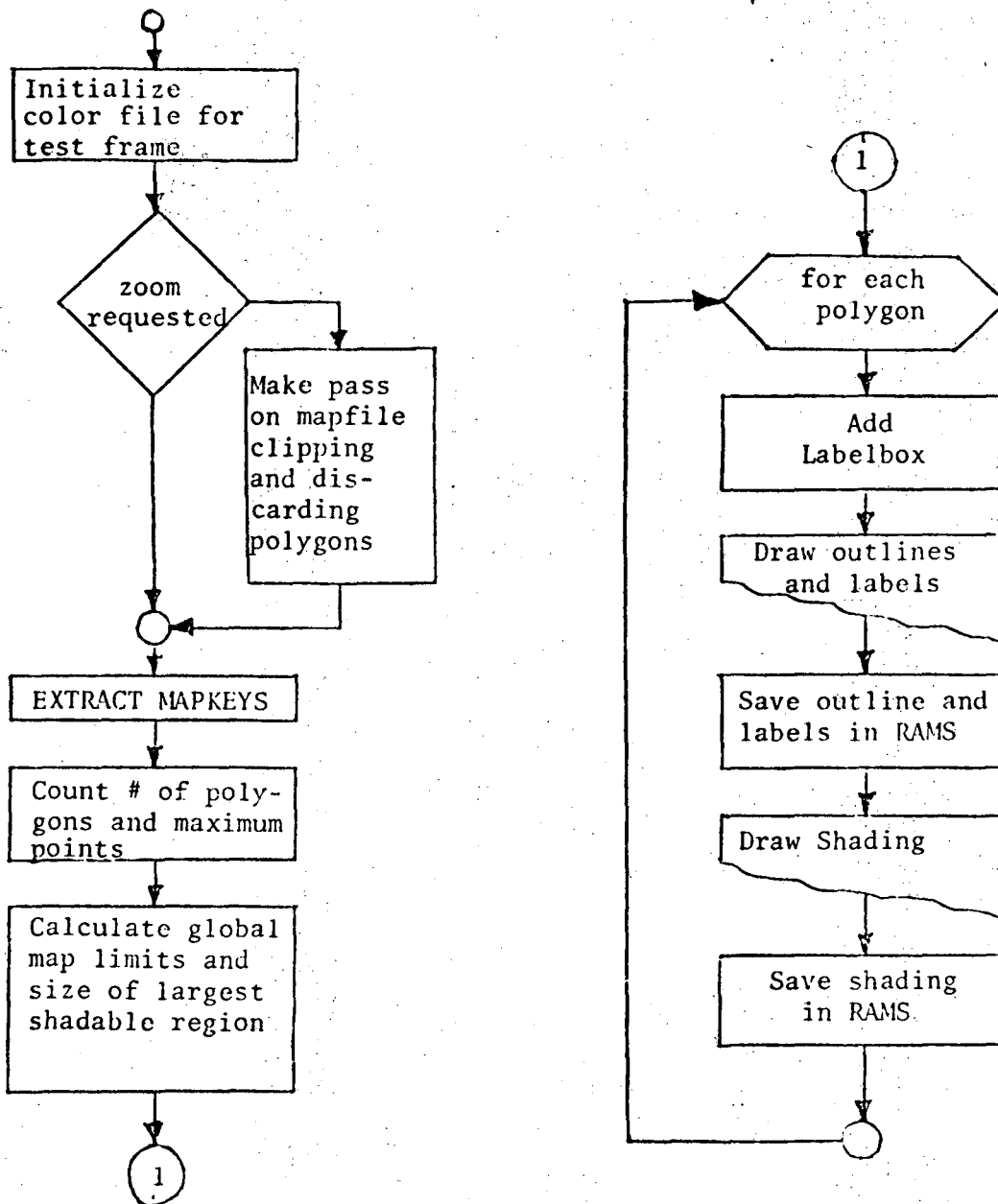


Figure 2. FUNCTIONS OF THE SETUP MODULE

A. Clipping Limits

In the absence of a *ZOOM directive, the map coordinate limits are taken from the map file - that is, the entire map is drawn and no clipping is necessary. However, it is often the case that only a portion of a map is wanted, either for clarity in viewing or for selection of subareas by coordinate limits. In this case, a directive of the form

```
*ZOOM,<long.min>,<long.max>,<lat.min>,<lat max>
```

specifies a rectangular area (in map coordinates) which is to be drawn (see Figure 3). There are three cases to consider:

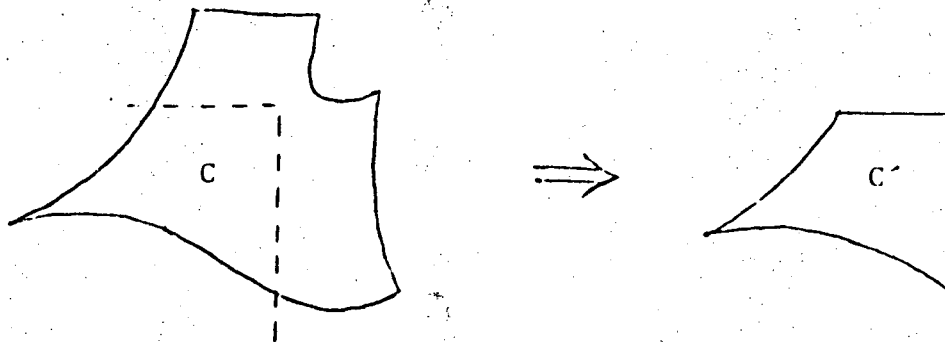
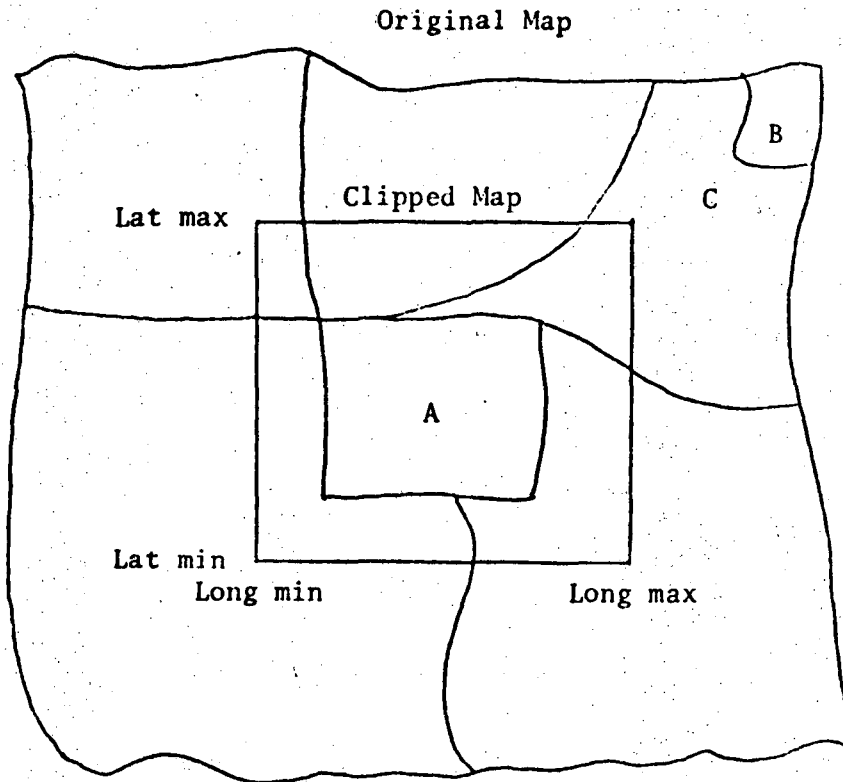
- a. A polygon lies entirely inside the limits
- b. A polygon lies entirely outside the limits
- c. A polygon intersects the limit rectangle.

The polygons are passed to a subroutine CLIPOLY for processing.

For case a, CLIPOLY returns the original points. For case b, CLIPOLY returns no points. For case c, the subroutine discovers which vertices of the limit rectangle must be added to the polygon (eliminating external vertices) in order to have a closed polygon entirely within the limits (polygon c' in Figure 3).

The CLIPOLY routine first determines the direction in which the polygon is drawn, so that the limit vertices can be inserted in the proper order. Then each side is sent thru the clipping routine TVCLIP, which flags lines intersecting the rectangular boundary and calculates the point of intersection (the cut point). If a side crosses the boundary,

Figure 3. CLIPPING TO A RECTANGULAR AREA



a flag is set and the cut point stored. Subsequent vertices are deleted until a side reenters the limit rectangle. A test is made on the exit and entry points to decide whether zero, one, two or three limit vertices must be inserted to preserve the closed polygonal structure.

Once all the polygons of a map file have been clipped to the limit rectangle, the polygons can be shaded with no further clipping. This represents a substantial saving of computer time if more than one map is to be made in a single run.

Another option is called INSET clipping, in which any polygon which intersects with the clipping rectangle is rejected. The routine will return either every polygon entirely inside the clipping rectangle or every polygon partially or entirely outside. This produces either a base map with irregularly shaped holes or the insets which exactly fit the holes in shape. More than one clipping rectangle can be specified for multiple insets.

B. Label Placement

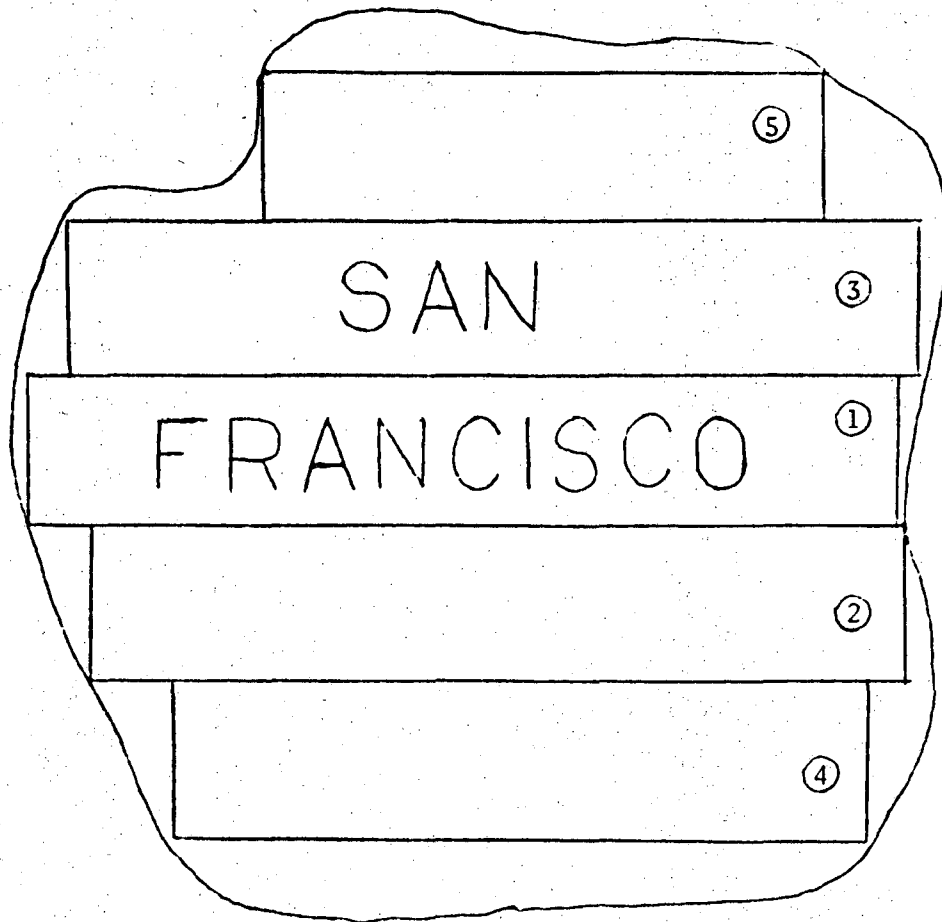
It is sometimes useful to display the name of a geographic area in a central part of that area, especially for single-color cross-hatch maps. Placing names manually is not only tedious, but must be redone when scale changes are made. Therefore, a label placing algorithm was developed which attempts to place a label of a user specified size within the polygonal boundaries of an area. The algorithm works as follows:

1. Calculate the limits of a rectangle into which the label of a specified character size will fit.
2. Starting at the center of the polygon and working outward symmetrically, search for a place where the rectangle will fit entirely within the polygon. If a place is found go to step 5.
3. Repeat step 2, only try for a verticle rectangle.
4. Check the label for an embedded blank to see if the label can be broken into more than one word. If so, repeat steps 2 and 3 for each word, retaining the proper order.
5. Place labels in largest area available, truncating the label to fit.
6. If the map is to be single-color, add the vertices of the rectangle to the polygon in a manner than retains a simple closed curve. This ensures that the labels will not be shaded over.

Feasible regions are found using the shading program to compute a set of parallel lines (horizontal or vertical) spaced a character height apart. These lines approximate a set of rectangles, and the number of characters which will fit into the rectangle is easily computed. In the first two attempts, the search for feasible regions begins in the center and works outward, building a set of rectangles sorted in descending order of length. These are then further sorted by spatial order, top to bottom. A list of suitable regions corresponding to each word of the label (step 4) is prepared, indicating the upper and lower spatial boundaries allowed for that word (i.e., the first word of a label must go above the second word for the horizontal case, and to the left for the vertical case). Finally, a loop is begun in which, at each iteration, the longest region in range for the longest word is chosen, the word is centered within the region (truncated if necessary) and the remaining range lists are updated.

In the example depicted in Figure 4, the label SAN FRANCISCO was broken into two words and placed in the optimum regions.

Figure 4. PLACEMENT OF LABELS IN A POLYGON



Regions

Order by Length	Vertical Order
1	5
3	3
2	1
4	2
5	4

Range	List
SAN	FRANCISCO
5	3
3	1
1	2
2	4

C. The Shading Algorithm

For multi-color maps, the shading algorithm is used during the setup phase to produce a masked region which is saved as part of the data structure. For single-color or cross-hatched maps, a different texture will be used for each characteristic being mapped, so the shading must be redone each time.

The shading algorithm in CARTE was adapted from a paper from the Polytechnic Institute of Brooklyn by William Dwyer, who credits Dr. Frank Sinden with its invention. The subroutine SHADE accepts a polygon of N vertices, a vector \vec{u} normal to a family of shade lines, and a distance d between shade lines. It returns an array of cut points for each shade line sorted in order of algebraically increasing distance from the origin of the coordinate system.

The sides of the polygon are traversed in order from the vertex closest to the lowest shade line to the vertex farthest, and then back down. The cut points are stored in arrays $X(I,J)$, $Y(I,J)$, for the I th cut point along the J th shade line. The cut points are then sorted as follows:

Let \vec{W} be a unit vector normal to \vec{u} (i.e., parallel to the shade lines). Then the points $\vec{P}_{i,j} = (X(I,J), Y(I,J))$ for shade line J are sorted in the order

$$(\vec{P}_{i,j} \cdot \vec{W}) < (\vec{P}_{i+1,j} \cdot \vec{W}).$$

In addition, if $|\vec{P}_{i,j} - \vec{P}_{i+1,j}| < \epsilon$, where ϵ is some resolution parameter (typically, $\epsilon \sim 0.01d$), this segment is eliminated. The

cut points are thus paired, with the first point of a pair being the entrance to the polygon along a shade line and the second being the exit.

To produce a cross-hatched map, the shade routine is called twice - once with a vector \vec{u} and once with a vector normal to \vec{u} . By varying the distance d between shade lines, and the direction of \vec{u} , at least eight distinct textures are available.

For dot screen or total masking (for multi-color maps) the shade routine is called with a vertical vector (horizontal shade lines) and a spacing appropriate to the dot size or line width. The best resolution is obtained by total masking using two raster point spacing (out of 4096 addressable points), providing a 2048 raster point resolution on the film. For dot screens, dot sizes range from 2 to 8 raster points in order to produce 3 distinct tones for each color, so the resolution is reduced to 512. This can give a stair-step effect along diagonal borders.

IV. The Report Generator and the Data Set Specifications

A table of the data being represented on a thematic map is often a useful addition to an atlas. The report generation module in CARTE allows user design of such tables thru directives which specify the format of the table (see Figure 5). The table is built from a template which contains key items specified by the directives, such as titles, column headings, and footnotes. Items in the table are taken from the data set, formatted according to the directives and output on a separate frame.

Figure 5. TEMPLATE FOR THE TABULAR REPORT

T I T L E 1 Title 2			Title 3		
Heading 1	Heading 2	Heading 3	Heading 1	Heading 2	Heading 3
Name 1	Datum 1	Datum 2	Name N+1	Datum 1	Datum 2
← First Repetition of Body Template →			← Second Repetition of Body Template →		
Name N	Datum 1	Datum 2	Name Last	Datum 1	Datum 2
*FOOTNOTE 1					

A. Form of the Data Sets

Data sets are composed of data items, each of which normally consists of three elements: one or more keys for matching, zero or more words (ten characters constitute a word) of area name, and one or more data values. The user may specify the format in which to read each data item by the *FORMAT directive. A data set is completed by a header card indicating the number of data items and the number of keys, words of area name, and the data values per data item.

The KEYS are the same geographic area codes that appear on the map file. The match between keys on the data and keys on the map file is specified by the KEY directive:

*KEY, key i on data = key j on map , (key k) = (key l,) ...

For example,

*KEY, 1 = 8

indicates that the first key on the data card must match the eighth key on the map file.

The NAMES are usually the names of the geographic areas on the map. Currently, the names are used only in the report, and can consist of up to 30 characters.

At least one datum is required on each data card - this is the statistic being represented on the thematic map. This item can be any of the following:

1. A number, to be used by the distribution module in assigning color or shade codes to the areas.
2. The color or shade code itself (i.e., a number between zero and the maximum number of colors or shades desired).

3. A nominal assignment (such as FORESTS, WETLANDS, etc.).

Any additional data items will appear in the table, but only one is used for generating the map. This is specified by the *EXTRACT directive.

B. The Report Template

Since the report generator must handle data for sets ranging from less than ten to several thousand areas, it was written to be flexible and algorithms were included to do most of the format specification automatically. Certain elements are assumed to be present for any table: the data to be displayed, one or more titles, column headings, and row headings (the area names from the data set). Figure 4 shows this basic report template. The report body template may be repeated several times across one frame, in order to fit more data on one page.

The text and placement of the titles may be specified thru the *REPORT directives. The text of column headings and footnotes is entered by the *HEADINGS and *NOTES directives. The placement of all other elements is done automatically within the general template above.

C. Some Features of the Report

In many applications the areas being mapped occur as elements of groups in the data (e.g. counties of states, states of regions). As it is important for these to be clearly delineated in the table, an array of keys, one for each list (the first specified on the *KEYS directive) is required, and, if an area's key value is zero, then it is set off from the rest of the body of the table by one line being skipped before it, its name being entered in capitals, and the text backspaced one character.

There are several other useful features which are included as options that the user can specify for table generation. Any column of the table, including the row names, can be input in either numeric or character form, column headings can be spread over more than one column, any footnote can be bound to a particular title, column or row heading, and as implied above, more than one column of data can be input in order to display with the characteristic mapped other values of interest. The program will also place commas in numbers as needed, indicate data suppressions by an asterisk, and make multi-page tables as necessary.

D. Generating a Table

With the possible exception of the titles, the user specifies only qualitative information to the table generator. The program handles the quantitative details.

Several operations must be completed before the format of the report body is decided.

First, the program counts the number of lines that will be skipped by examining the array of keys, obtaining an accurate count of the actual number of lines to be used by the table. Then the field width of each column is estimated by examining the row headings and data values, counting characters in the case of alphanumerics and computing a floating point format with space allowed for commas as needed, in the case of numerical values.

These estimates are then compared with the individual column headings and adjusted as necessary. The space required for titles and footnotes is computed. Finally, the values of the parameters defining the body of the report can be computed.

The key variable for the report is the number of times the report body template is to be repeated across the page. In CARTE, this is calculated by the following formula:

$$R = \min((NL-1)/LP+1, w/(C*S(K)))$$

where

R = the number of times the body template is to be repeated

NL = the number of lines to be drawn in table body

LP = the number of lines possible per template repetition

W = frame width

C = total characters per data line

S(K) = horizontal width of characters at size K.

This formula is tested for varying values of K and LP until it is apparent that all the information will fit on one frame or not. Once this value is known, the tab stops for the body, centering information for headings, and actual number of data lines to be displayed per body template repetition are computed.

The table is then produced. While looping through the data items, at each new frame, the report sends out the titles and footnotes and a few major lines; at each body template repetition the module sends out the column headings and the tab stops for the body of the report. Each data item constitutes one line in a body template repetition which is enclosed in a temporary buffer until it is complete.

V. The Distribution Module

The distribution module transforms a set of data into a set of color codes to shade the map.

There are three methods of generating colors from a data set in CARTE. Two of these merely divide the data into groups at given division points. This occurs when internal ranges are given or nominal data is used. The third, automatic distribution with rounding, should be selected when a user does not know where to divide the data. The algorithm is set to attempt an equal number of area items per division. This was chosen because it is a relatively simple yet informative method and in the case of nearly equal areas, results in a nearly equal division of the map into shaded regions. It was decided that the usefulness of the mean or equal area approach did not outweigh the cost, especially because the rounding process, needed for a pleasing display, disturbs most distributions anyway. Dividing data into equal groups is a straightforward task except in cases where the distribution is extremely skewed, in which case some arbitrary range must be imposed.

In CARTE's algorithm, the data is first sorted into 100 bins by a simple transformation. Thus the smallest unit of the data space is effectively 1/100th. Then these bins are divided into an arbitrary number of subgroups, usually eight or less, the number being set by the user, thus generating an expected number of items/group. While cycling through the bins, a count is kept of the number of items accumulated since beginning and since the last division was made. When the expected number of items is past a decision is made as to where to place the dividing line. After each such decision the expected number of items per subgroup is

recalculated from the number of items actually allocated, and if needed, fewer subgroups are used.

After transforming the calculated division points back into the data space, the program proceeds to round these into acceptable numbers for presentation on a map. The user controls the number of significant digits to remain in each division (ns) and the range these digits are allowed to take over the set of division points (nmag). This second parameter insures that the division points will all end in the same column. As an example, assume the division points calculated are:

.712, 6.711, 24.381, 78.172.

With ns=2, nmag=3 these will be rounded to:

.7, 6.7, 24.0, 78.0

A maximum of two significant digits are allowed, but the first point is allowed only one because the divisions must fit in a three column range. If nmag = 2, then the result of rounding would be:

1., 7., 24., 78.

The process the program follows is to normalize all division points to less than one, multiply by the calculated ns, add .5, fix the number and multiply again by the adjusted normalization factor. In some cases the division boundaries will now be equal. These are then adjusted by the minimal allowable amount (given ns, nmag) and divisions are eliminated if the adjustments make them extend too far beyond the range of the data.

The colors are then assigned to the data items, followed by the encoding of the division ranges into the display form (the color shade code). The minimal field width for the set of divisions is calculated and used, including space for commas in large numbers as needed. Also,

a histogram of the data as divided is automatically compiled and displayed.

The shades or textures, stored in data item order, are then assigned to the polygons on the map file. There are two steps in this process. First each polygon is assigned its appropriate color. Then the polygon list is reordered so that all areas with the same color are grouped together. This enables the use of one film file, thus allowing an essentially unlimited number of shades, and requires random access to the map in the graphic output module.

VI. The Graphic Output Module

The next function step in CARTE is to put out the map image as specified on the appropriate files. There are two parts to this process: putting out area outlines, labels and shading, and creating the titles, legends and other cosmetics.

A. Area Outlines, Labels, and Shading

There are two cases to be considered here. The first is that of cross-hatch mapping. In this case all of each map is on one frame, but more importantly, the areas are in NICKEL format [2]. Thus for each area, for each map, the display file is generated from the stored map. The process is straightforward conceptually, but the generation of shade lines relatively costly. It is the fact that the shading for an area will probably vary that requires us to store the map in its source form. The second case is for multi-color mask or dot maps. Here, in the setup phase, the display file for the entire map was generated and saved. Thus to generate an area's outline, labels, and shading, we have only to transfer its pre-generated display code to the color buffer with dot sizes being changed for dot maps. As can be expected, the cost of this method is much cheaper than that of cross-hatch.

The effect of the color, generated by the distribution module, depends upon the type of map being produced. In the case of cross-hatch, the orientation and spacing of the shade lines is affected. In the case of mask, the choice of output frame is affected. In the case of dot, the choice of output frame and dot size is affected.

B. Creating Titles, Legends, and Other Cosmetics

The titles and legends may vary from map to map. Thus the display file for them must be generated each time.

Every title is described internally by several characteristics: a centering point, character size and orientation, number of lines allowed below the centering point, width of a line in characters, number of characters, and the text of the title. The title drawing routine takes this information and centers it below the given point, creating as many centered lines of text (broken at blanks) as necessary.

Drawing the legend is a bit more complicated. The program is given an area in which to place the legend, the legend text, the text describing the color code, the allowable height and width range for the box for each color, and the allowable range of character sizes. Since the given legend area and/or the number of colors used, and/or text lengths may vary from map to map, the allocation of space within the legend box is recalculated for each map.

The process is as follows: first, for maximum character size, enough space is allocated to draw the legend. Then with the remaining area, the program decides whether to place the color code and boxes in a row horizontally or vertically. Once this is decided, an estimate is made of the unit width or height respectively (the space needed per interval). This estimate is then tested to see if the code and box for a color can fit in. If not, the estimate is revised by changing the number of text lines allowed, the code and legend character size, and the box size until a solution is reached or infeasibility is shown. The algorithm is set to maximize the size of the individual color boxes. The area required

by the legends and color codes is derived by an inverse operation of the title drawing routine, i.e., given text, its display space is computed.

Once the space for each element within the legend box has been allocated, the elements are centered within the given area and drawn. If so directed, the program will give the type of fit tried, the text displayed, and the spacing information it computed.

VII. Data Structure and Memory Management

For large maps, memory becomes a scarce resource. Thus CARTE minimizes its memory requirements by storing the map on auxiliary storage in a compact form (for screen type maps); by having the size of many of its working arrays determined at execution time; by swapping these working arrays in and out of memory as needed; and by using only as much small core memory and auxiliary storage as needed.

A. Storing the Map

A map in CARTE consists of a series of closed polygons, each describing an area that can be shaded. Accompanying the points for each polygon is a set of descriptors primarily for matching with the data and a set of labels or area names associated with that polygon.

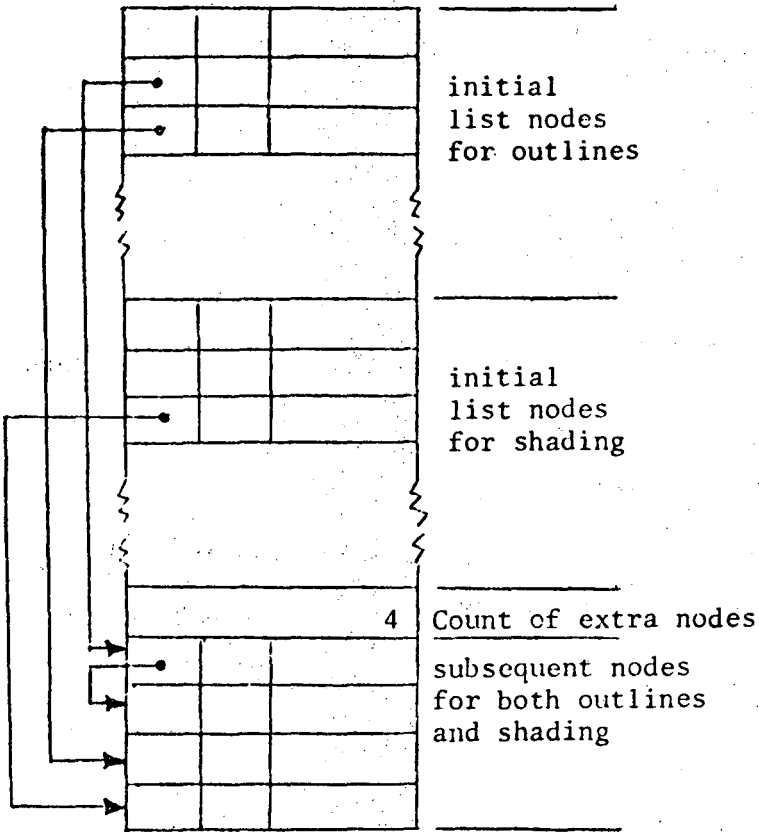
The set of polygons comprising a map is stored differently depending upon the type of map to be produced. In cross-hatch mapping, an image of the map input file is kept on disk thus allowing for later modification of the map (e.g., further windowing, changing the viewport). This approach requires the calculation of shade lines for each map. In the case of screen type maps, the shade lines need be generated only once, so a different, less expensive, approach may be used. The approach is to create a graphical data structure as described by Newman and Sproull [3]. In this case, the raw map is no longer the data base, the display file created in the setup phase of the program is used as the data base for the map. This has two major advantages: it requires less space to store and the cost of generating a new display file for each map is eliminated. Its only disadvantage is that the map cannot be easily redesigned later in

a run.

In this approach, two blocks of memory are created for each polygon: one for labels and outline, one for shade lines. These display file blocks are stored in random access mass storage in sequence as generated. Since it is unpredictable whether or not the shading for a polygon will be used on any given map, pointers and block lengths are kept for each block.

The method of creating the graphical data structure is to save the contents of the color buffer as the test frame for the COM is generated. This has the advantage of minimizing memory requirements during the creation and later generation of maps. This implies that when the display file block for a polygon's outline or shading is longer than the buffer size it is broken up into buffer loads. Thus the array of map block pointers is an array of singly-linked lists of the buffer loads comprising a block. Figure 6 summarizes the structure of this array. It is organized so that the first buffer load (usually the only one) is accessed by the polygon number plus a constant.

Display code blocks for each polygon are described by two singly linked lists - one for outlines and labels one for shading



Hypothetical Picture of Map Pointer Array

List pointer for polygon i is i for outlines and labels $i+n_{poly}$ for shading

List node structure

Link	Word Count	RAMS Address
------	------------	--------------

Link = 0 indicates terminal node of list
 $\neq 0$ gives index of next node

Figure 6. Structure of Map Pointer Array

B. The Dynamic and Fixed Length Structures

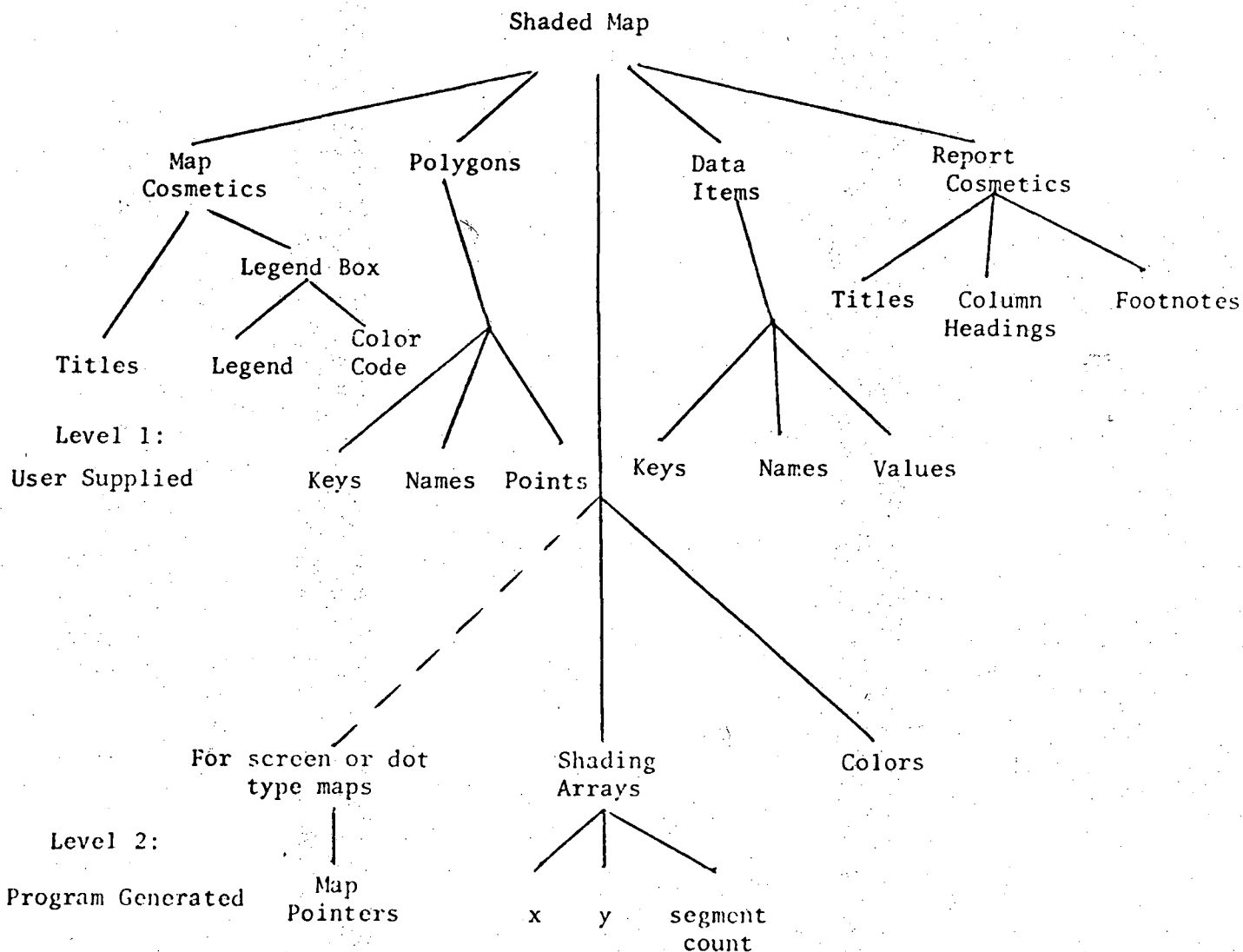
The maps drawn by CARTE so far have ranged from 9 to 3100 polygons. Thus the size of many arrays varies widely, and more importantly, would overflow memory if kept in core at one time. For these a variable sized dynamic allocation was adopted. Other arrays, such as those for map and table titles, have small variation in size. These are kept at a fixed size.

Figure 7 depicts the elements of CARTE's data structure. The terminal nodes of the tree comprise the major arrays of the program.

The fixed length arrays are those under map and table cosmetics since these vary within small ranges. Also of fixed length is the array of polygon names, because the names and points for only one polygon at a time need be in core. The points array is variable because the maximum number of points needed can range from ten to thousands. Thus including the points array there are ten arrays of variable sizes. The size of the variable data arrays is computed from the data set header card, while the size of the map arrays is computed from a preliminary pass through the mapfile in the setup module. The next section describes the management of the variable arrays.

Figure 7. DATA STRUCTURE OF CARTE

Terminal nodes are data structure elements



C. Memory Management

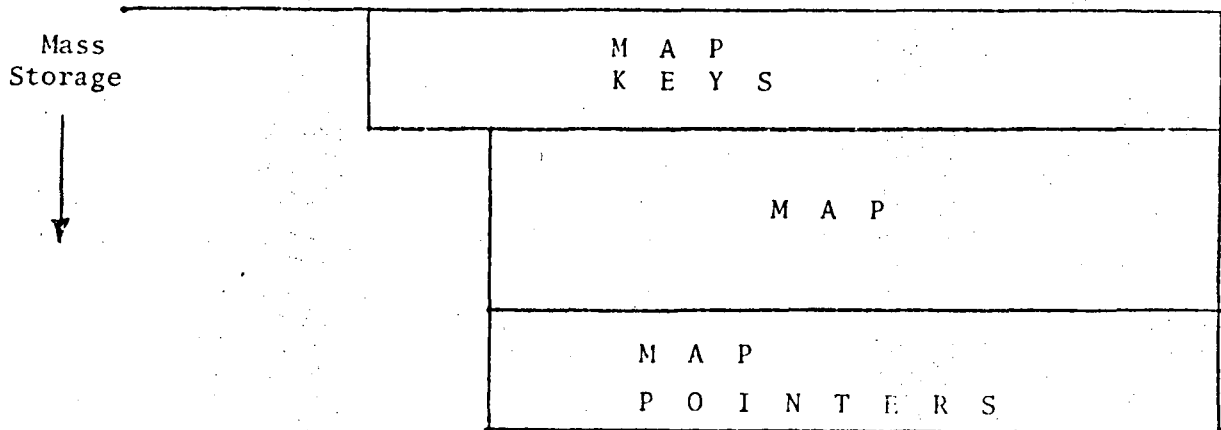
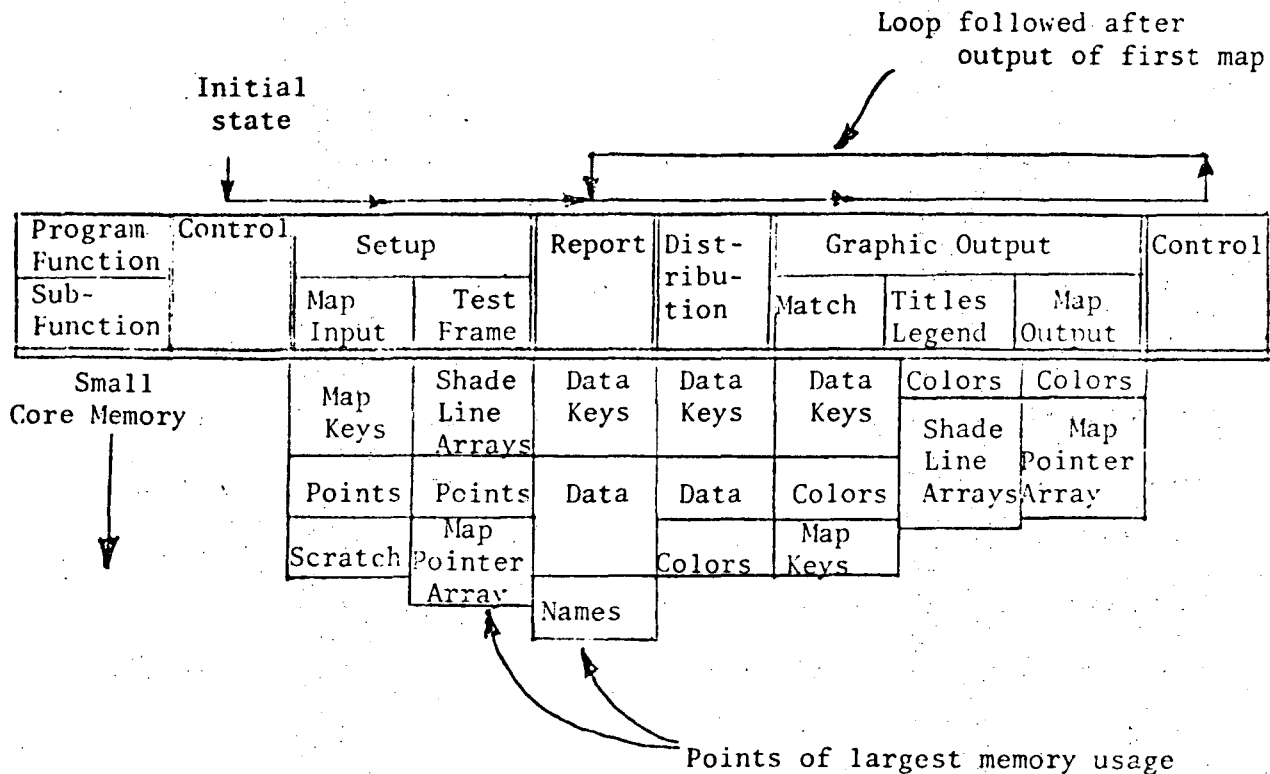
Because the operations of the program are known, the management of memory is also known. Small core memory, where the current arrays are kept, is treated like a modified stack. Thus the two basic stack operations are used: putting a block on top of the stack, and freeing the block on top of the stack.

Any block in the stack is accessible and the operation of freeing a block low on the stack frees all memory above it.

When a block of memory is created, a status flag indicates whether this is a place for new data or a place to put data being saved in auxiliary storage, in which case the block from auxiliary is moved into its new small core location. When a block of memory is freed, the status flag indicates one of three options: free the memory; move the block down to stack location x , freeing all above; or free the memory, saving the block at the next available location in auxiliary storage.

Thus, for each of the ten arrays, a block length, a small core address, and a large core address are maintained. The memory map in Figure 8 depicts the contents of SCM and auxiliary storage over time.

Figure 8. MEMORY MAP FOR CARTE* OVER TIME



*Chart is for dot or mask type maps

D. Random Access Mass Storage

The CARTE program is designed to make efficient use of random access mass storage. On the CDC 7600 in use at LBL, this takes the form of 512,000 words of large core memory (LCM) and two 75 million word disk files. All processing takes place in 65,000 words of fast memory (called small core memory or SCM), so CARTE transfers blocks of data between fast memory and auxiliary storage as required. Large core memory is always treated as a random access file, with dynamic allocation of blocks of memory. In the case the data structure overflows LCM, the program automatically switches to random access disk files. This combination of fast memory, large core memory and disk files allows CARTE to handle a wide range of map files in an efficient and cost-effective manner. Experience with the program has shown that memory management is a necessary feature for a mapping program, since maps produced have ranged from nine areas with a few thousand points to over 3,000 areas with more than 100,000 points (over 200,000 words).

VIII. The Symbol Mapping Nodule

A. Introductory Considerations

CARTE produces shaded maps. It was first designed for choropleth mapping, where geographic entities are accurately described by polygons. Geographic entities may also be points and lines. To shade point data it is necessary to represent each point by a symbol. (Lines can be shaded by expanding them in width.) Shading point data is called symbol mapping. User-defined symbols, each defining a locality type, are placed at specified coordinates on a base map, and shaded according to the range of a common attribute. Examples are allocation of funding by type of sponsor and power output by type of generation facility.

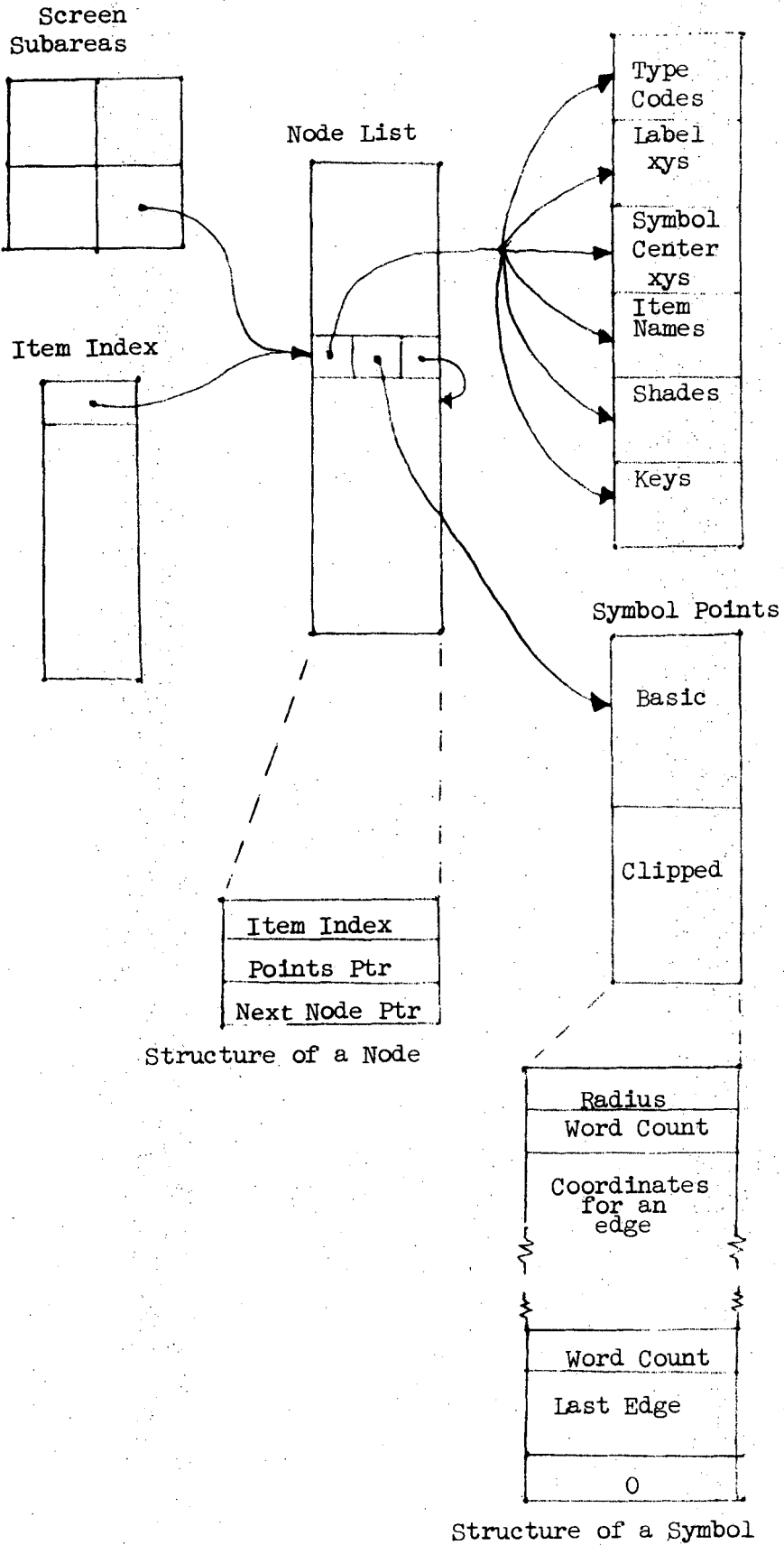
The placement and shading of a polygon presents no problem, but complications arise if symbols are allowed to overlap. (In choropleth mapping non-overlapping symbols are drawn.) If two symbols overlap, it must be decided which goes on top. Thus the conception of depth is added. Also, algorithms for general polygon-polygon clipping are needed, so that just the visible portion of a symbol will be drawn and shaded. And an efficient means of storing and searching the already compiled portion of a map is needed, in order to determine if symbols do indeed overlap.

A symbol may have many instances; it may be placed at many different places on the map. Each instance of a symbol may remain whole, become partitioned into fragments, or have holes (islands) added to it. This requires a flexible data structure for storing symbols. Also, a user may not know the map coordinate for a symbol. He may wish it to be placed anywhere within a geosarea. Thus, the program must be able to place the symbol in a suitable location.

Experience has shown that it is useful to annotate each symbol. Thus some facility for generating (placing) symbol annotation must be provided.

In the output phase, as in choropleth mapping, the data items are output

Figure 9. Basic Data Structure for Symbol Maps



in shade, not data item, order. An index pointing to all necessary information to draw and shade each symbol must be built.

These considerations are built into the symbol mapping module of CARTE and will be discussed in more detail below. The basic data structure will be discussed first, as it is the foundation upon which the algorithms are built.

B. The Basic Data Structure

The following sections describe the basic elements of the data structure used in symbol mapping. They are basic in the sense that they are global to the compilation and output phase of producing a symbol map.

1. Symbol Point Storage. A symbol may be defined by a collection of closed polygons a code defining its type, and a text definition. The points of each user-defined symbol are stored in a block which includes the radius of the symbol about the origin, and the actual points as a series of closed polygons, also relative to the origin. This makes it easy to place a symbol anywhere on the screen and allows many data items to use the same symbol definition. The radius simplifies checking for overlap.

The outer edges of a symbol go in one direction while islands, or inner edges are indicated by a flag and go the other way, e.t., counter-clockwise on the outside, clockwise on the inside. This is essential for correct traversing of the symbols when clipping.

The collection of basic symbol definitions is saved in RAMS until it is time to compile the symbol map. At this time they are loaded into dynamic memory allowing them to be treated in the same way as generated symbols by other parts of the data structure. Variations on the basic symbols generated by overlap are stored in the same manner as the basic symbols. They are stored at the next available location in dynamic memory.

2. Data Item Elements. Associated with each data item to be represented on the map are, as in choropleth mapping, a set of keys, a name, and a shade. Also

stored in data item order are symbol type code, and space for a symbol and label xy coordinate (to be generated if not already supplied).

3. Structure for Searching the Map. The screen is partitioned into subareas, each larger than any permissible symbol radius. Each subarea of the screen is represented by a linked list of nodes indicating the data items placed in that subarea. As each data item is placed a new node is added to the appropriate subarea list.

A node consists of three parts -- the data item number, a pointer to the points defining that instance of the appropriate symbol type, and a pointer to the next node on the list. The data item number acts as a pointer to the elements described in section 2.

There may be data items which are not drawn. Thus the nodes are stored in dynamic memory interspersed with points for new symbols.

4. The Index. The index consists simply of pointers to the appropriate node locations in dynamic memory. These pointers are stored in data item order.

C. Algorithms

1. Overview. The process of producing a symbol map is divided into two phases - that of compilation and that of output. This enables the production of more than one map, i.e. different shading, from one compilation. Figure 10 shows a flow chart of the compilation phase, which is described in sections 2-5. This condensation allows the compilation phase to be relatively costly, since it is performed only once. The clipping algorithm also can be relatively costly, since it is executed for just a small subset of cases.

2. Depth. Selection by depth forms the outermost loop of the compilation module. Depth is defined by symbol type in order of definition. Thus if symbol type 1 to n are stacked on top of each other, symbol 1 will be completely visible, while symbol n will be least visible.

This means all data items with symbol type 1 are compiled first, all those with symbol type n last. If two symbols of the same depth (i.e. type) overlap, the one compiled first will be most visible.

3. Symbol Center and Label Coordinate Generation. In some cases, the symbol is to be placed anywhere within a polygon from the base map. Thus the symbol center must be generated. After finding the appropriate base map area by matching keys, the program shades the area finding feasible regions in the same manner as automatic label placement. These regions are tried until one with no overlap is found. If no such region exists, then the symbol center is the center of the rectangle circumscribing the area.

Label coordinates, if not supplied, are created so that the item name will appear one character width to the right or left of its symbol.

4. Searching the Map. After being transformed from data to screen space, the symbol center is combined with the points of its basic symbol type. These absolute points are used by the clipping algorithm. Using the radius of the symbol, part of the already compiled map is searched for overlap.

Figure 10. Overview of Symbol Compilation Phase

For depth $d = 1$ to n do

 Find data item with depth = d

 Load symbol center

 If center not supplied, then generate

 Transform center to screen space

 Load symbol points and radius

 For each of nine subareas do

 For each node on subarea list do

 If old symbol and new symbol overlap then

 Load points of old symbol

 Clip new symbol to old symbol

 Store new symbol node

 Store index entry

 Store points if clipping occurred

 Store center and label coordinates

The node lists for at most 9 subareas are traversed (the one the symbol center falls into and the eight surrounding). Since the screen is partitioned into 100 subareas, this search includes at most 1/10 of the screen. The surrounding subareas must be checked because, even though the new symbol center falls in only one subarea, its points may overlap into others, and symbols centered in other subareas may overlap into the subject's subarea.

For each node, the circle described by its center and radius is compared with the circle described by the new symbol's center and radius. If the circles overlap, it is probable that there is symbol overlap and the clipping routines are activated. Only then are the points of the old symbol loaded.

This method provides a quick and efficient means of searching the map. When the search is complete, a new node is added to the appropriate subarea list and the symbol points stored, if clipping occurred. An entry into the index for that data item is also made.

5. The Clipping Algorithm. When the probability of overlap has been established, a four step algorithm is followed to clip the new symbol so that it conforms to the boundaries of the already compiled symbol. A symbol is here defined as a set of one or more fragments (closed polygons representing outer edges), and zero or more islands (closed polygons representing inner edges). The four steps are -- build a relation table, compute intersection points, traverse the symbols, and compose the clipped symbol.

First a relation table is built, indicating precisely the type or relation that holds between each component of a new (subject) and old (object) symbol. There are four possibilities -- S inside O; S outside, enclosing O; S outside, disjoint from O; and S intersecting O. These relations are determined by testing each point of each component of S against the points of each component of O, using the Shrimrat Straddle algorithm described in the paper, Point in Polygon Algorithms

by HRP Ferguson. If S is outside O, the reverse test of each point of the O component with respect to the S component is required to determine enclosure or disjointness.

Next, all intersections for the two symbols are computed. This step uses the point-slope method of computing intersection points between line segments. To facilitate traversing, each intersection point is represented as: x, y, its S line, its O line, and a status flag indicating whether or not it is active.

With this information and the different orientations of inner and outer edges, the symbols are easily traversed, generating fragments and/or composite symbols. The most difficult part of the traversing algorithm is finding the correct point to start the traverse. An intersection point must be found which can act as the origin of an S line segment. This is done by testing the midpoint of the line segment formed by a prospective intersection point and the next S vertex as to whether it is outside the O component the intersection point was generated from.

Once a correct starting point has been found, the algorithm is as follows:

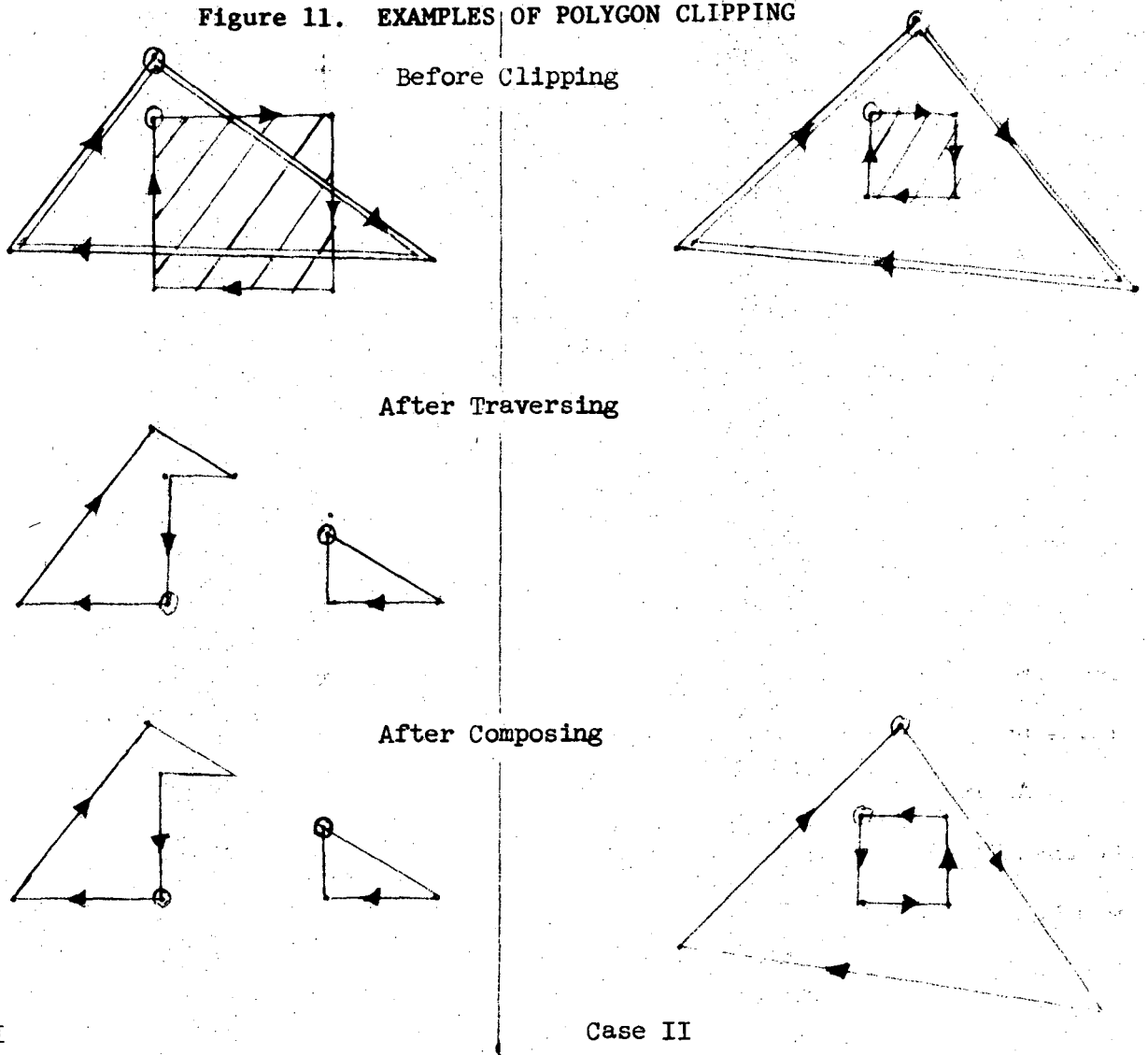
1. Follow S (in line direction) until next intersection point.
2. Follow O (opposite line direction) until next intersection point.
3. If not starting point, go to 1.
4. Compute next starting point, if any, and go to 1.

Finally comes the step of composing the clipped symbol from the collection of original S and O edges and the polygons generated by traversing. This task is made easy by the relation table. The algorithm is as follows:

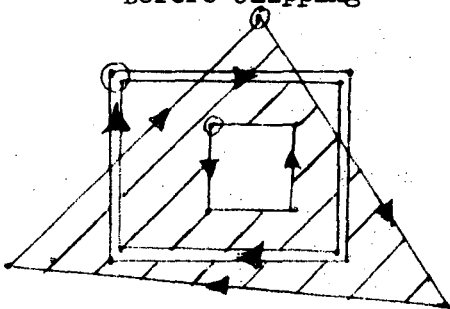
1. ENTER all components of S not inside the outer edge(s) of O.
2. ENTER all components of S inside or intersecting an inner edge of O.
3. ENTER all components of O inside an outer edge of S and outside all inner edges of S.

For this procedure to work, two things are needed. First, the identity of each

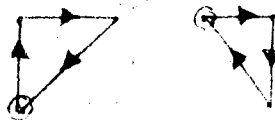
Figure 11. EXAMPLES OF POLYGON CLIPPING



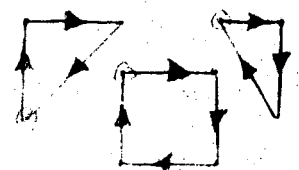
Case III
Before Clipping



After Traversing



After Composing



Examples of Clipping

- New Symbol
- ▨ Old Symbol
- Line Orientation
- Edge Starting Point

Note changes in line orientation and edge starting point produced by algorithm

S and O component and polygon generated by traversing must be known. This can be done simply by keeping a list of pointers to the current use(s) of each component, updated during the traverse. In complex cases, several components can be used in one generated polygon, or one component can be split into several polygons.

Second, the ENTERing procedure must perform several additional functions besides simply entering points. It must make sure all uses of a component are entered, deactivate all components used, and preserve line orientation. Any outer edge of O will become an inner edge and any inner edge of O will become an outer edge of the composed symbol, if entered.

The process of composing a clipped symbol is illustrated in figure 11. The composed symbol replaces the original symbol being processed and the program moves on to the next node on the sector list.

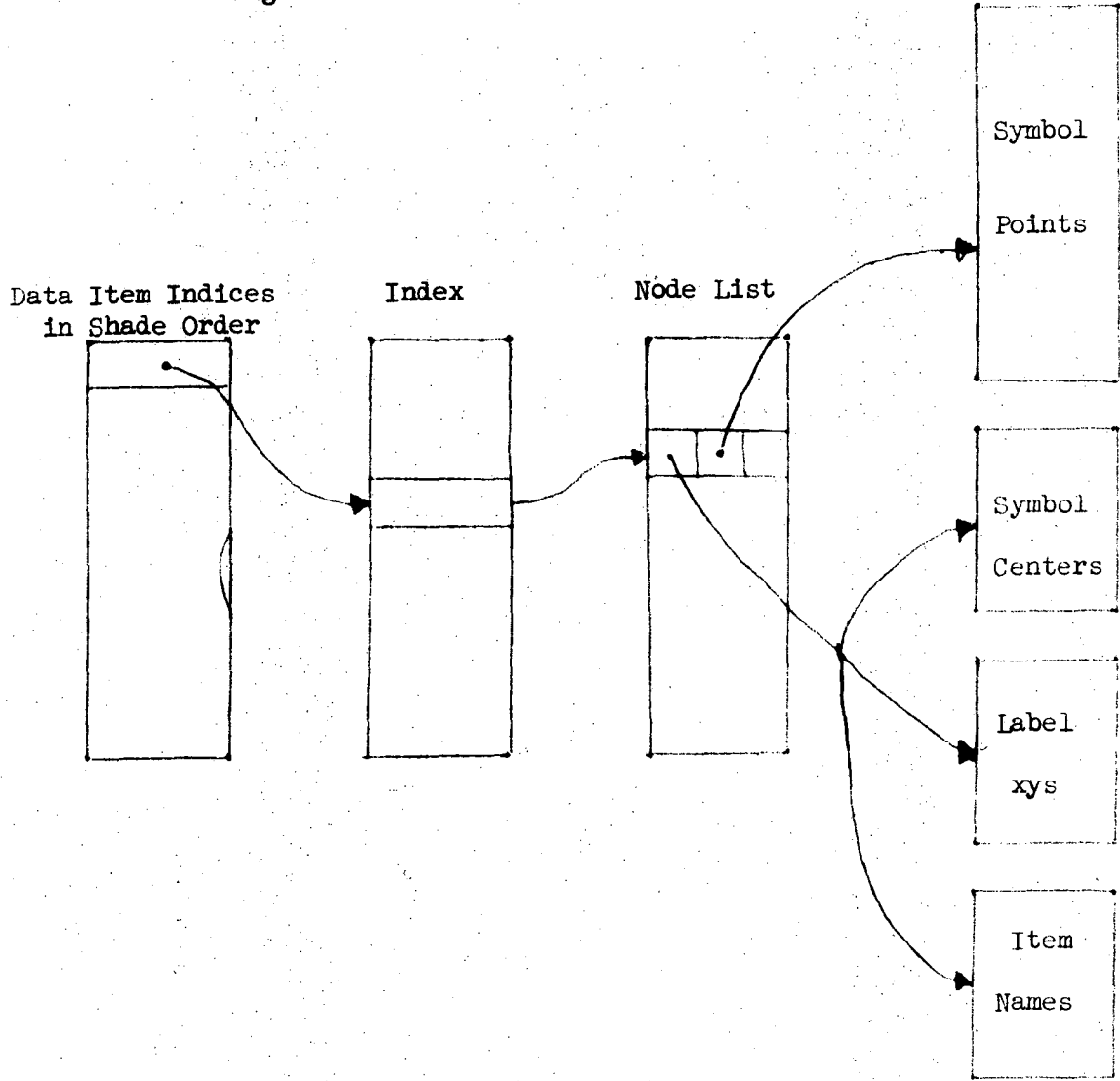
An interesting feature of the algorithm is that inner edges do not have to be associated with their enclosing outer edges. The program must simply know whether each component is an outer or inner edge. This is because symbols are considered as a whole both in the clipping algorithm and in the output phase of the program.

6. Producing the Map. Once the compilation phase is complete, the symbol map can be produced for many different sets of data values.. Because output is on only one film file, two passes must be made through the data structure.

The first pass draws the symbol outlines. The program simply passes through the dictionary, drawing each component of each symbol separately. Names are also drawn during this pass.

The second pass produces the shading. First the shades are processed to produce a list indicating which data items are to be drawn for each color frame. Then, as each data item index is encountered, the appropriate dictionary entry is accessed, and, through the node, all points and the center for that symbol are combined and passed to the shading algorithm which generates the correct shading.

Figure 12. OUTPUT PHASE DATA STRUCTURE FOR SYMBOL MAPS



D. Added Directives

For the functions described above to be made available to the user, two (only) directives were required: One for defining symbol types, and one for plotting a symbol legend. They are:

*USER-DEFINED SYMBOL, code, text def, N1, points, N2, points....
outer edge, 1st inner edge, etc.

*VERTICAL SYMBOL LEGEND, N, xmin, xmax, ymin, ymax

REFERENCES

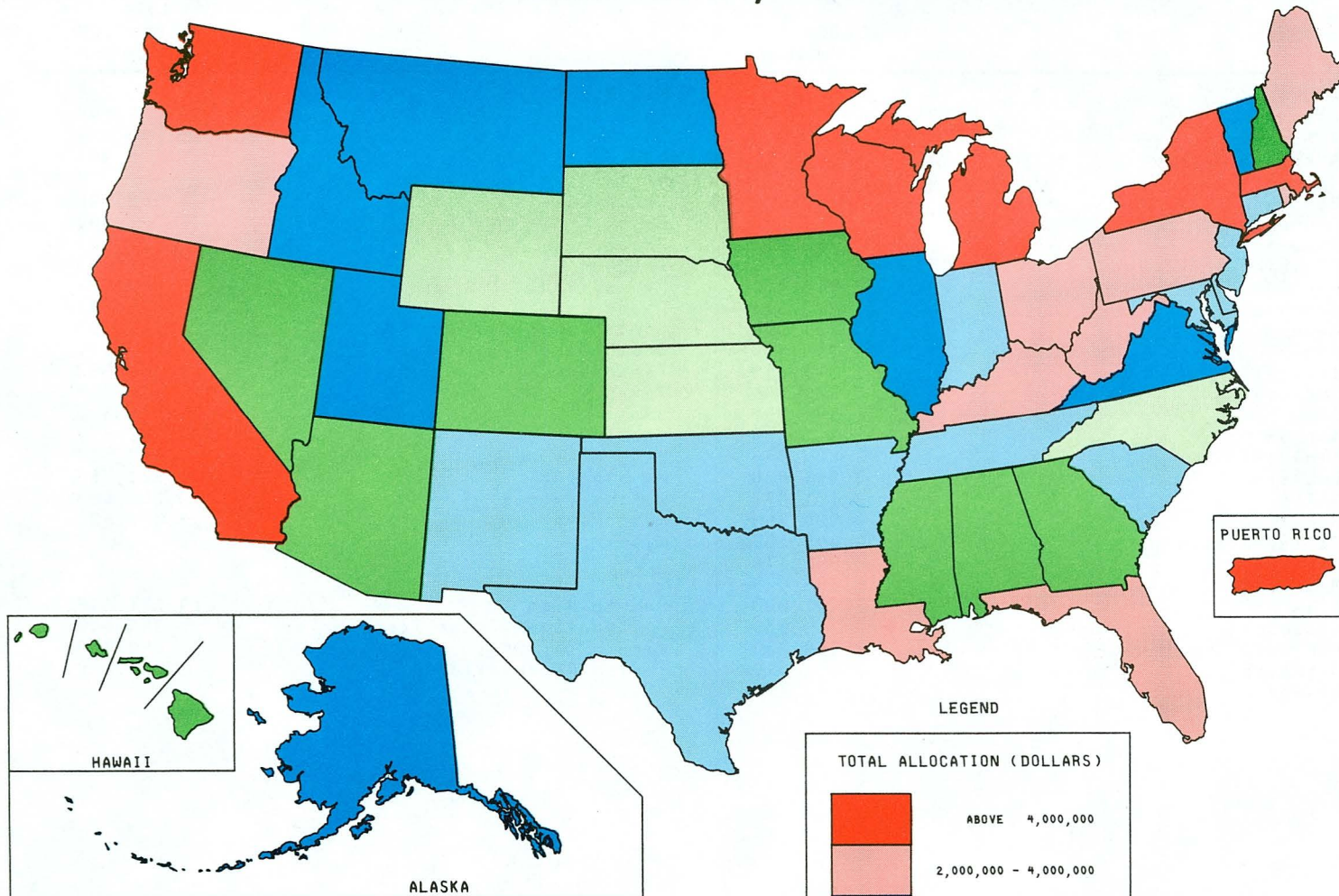
1. SYMAP Manual, Version 5.15, Laboratory for Computer Graphics and Spatial Analysis, Harvard University, Cambridge, Mass. 02138
2. Holmes, H.H., Austin, D.M., and Benson, W.A., "The MAPEDIT System for Automatic Map Digitization", Computers and Graphics, July, 1974.
3. Newman, William M., and Sproull, Robert F., Principles of Interactive Computer Graphics, McGraw-Hill, 1973.
4. Robinson, A.H., and Sale, R.D., Elements of Cartography, John Wiley and Sons, 1969.

Map and Table Examples

1. U.S. by State (color)
2. U.S. by County (color)
3. U.S. by SMSA (color)
4. Denver base and inserts (color)
5. Upper Mississippi (cross hatch)
6. Prime Sponsor Symbol Map (color)

MAP 2 - Allocation of State Government Portion of CETA Title II Funds For Public Service Employment - FY 1974 United States by State

PAGE 4
RUN DATE 03/01/75



MEDIAN STATE ALLOCATION
\$ 1,368,000

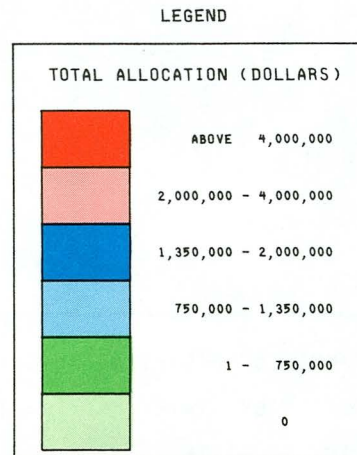


TABLE 2 - ALLOCATION OF STATE GOVERNMENT PORTION¹
OF CETA TITLE II FUNDS FOR PUBLIC SERVICE EMPLOYMENT - FY 1974

PAGE 3
RUN DATE 03/01/75

U.S. DEPARTMENT OF LABOR
MANPOWER ADMINISTRATION

UNITED STATES BY STATE

LAWRENCE BERKELEY LABORATORY
UNIVERSITY OF CALIFORNIA

STATE	TOTAL ALLOCATION (DOLLARS)	STATE	TOTAL ALLOCATION (DOLLARS)
Alabama	479,465	New Hampshire	311,234
Alaska	1,842,989	New Jersey	1,019,144
Arizona	140,500	New Mexico	1,330,479
Arkansas	852,774	New York	5,576,351
California	5,212,978	North Carolina	0
Colorado	205,078	North Dakota	1,368,400
Connecticut	769,000	Ohio	2,741,000
Delaware	1,266,312	Oklahoma	947,054
District Of Columbia	2,258,500	Oregon	2,544,293
Florida	2,914,400	Pennsylvania	2,049,722
Georgia	273,353	Rhode Island	2,038,887
Hawaii	738,350	South Carolina	1,200,383
Idaho	1,996,800	South Dakota	0
Illinois	1,379,359	Tennessee	1,334,332
Indiana	1,221,564	Texas	914,039
Iowa	494,300	Utah	1,854,100
Kansas	0	Vermont	1,552,279
Kentucky	2,236,856	Virginia	1,779,814
Louisiana	3,915,830	Washington	5,321,565
Maine	2,771,077	West Virginia	3,261,649
Maryland	909,364	Wisconsin	4,496,520
Massachusetts	10,849,185	Wyoming	0
Michigan	4,845,776		
Minnesota	4,793,023	Puerto Rico	10,677,698
Mississippi	702,000	A.Samoa-Guam-Trust Territories	345,300
Missouri	314,600	Virgin Islands	246,700
Montana	1,860,200	Indian Reservations	1,855,000
Nebraska	0		
Nevada	473,759		

1. FUNDS TO BE ADMINISTERED BY STATE GOVERNMENT FOR BALANCE OF STATE AREA

MEDIAN ALLOCATION BY STATE \$ 1,368,000

MEAN ALLOCATION BY STATE \$ 1,963,000

MAP 3 - Allocation of Total CETA Title II Funds For Public Service Employment - FY 1974 United States by County

PAGE 8
RUN DATE 03/01/75

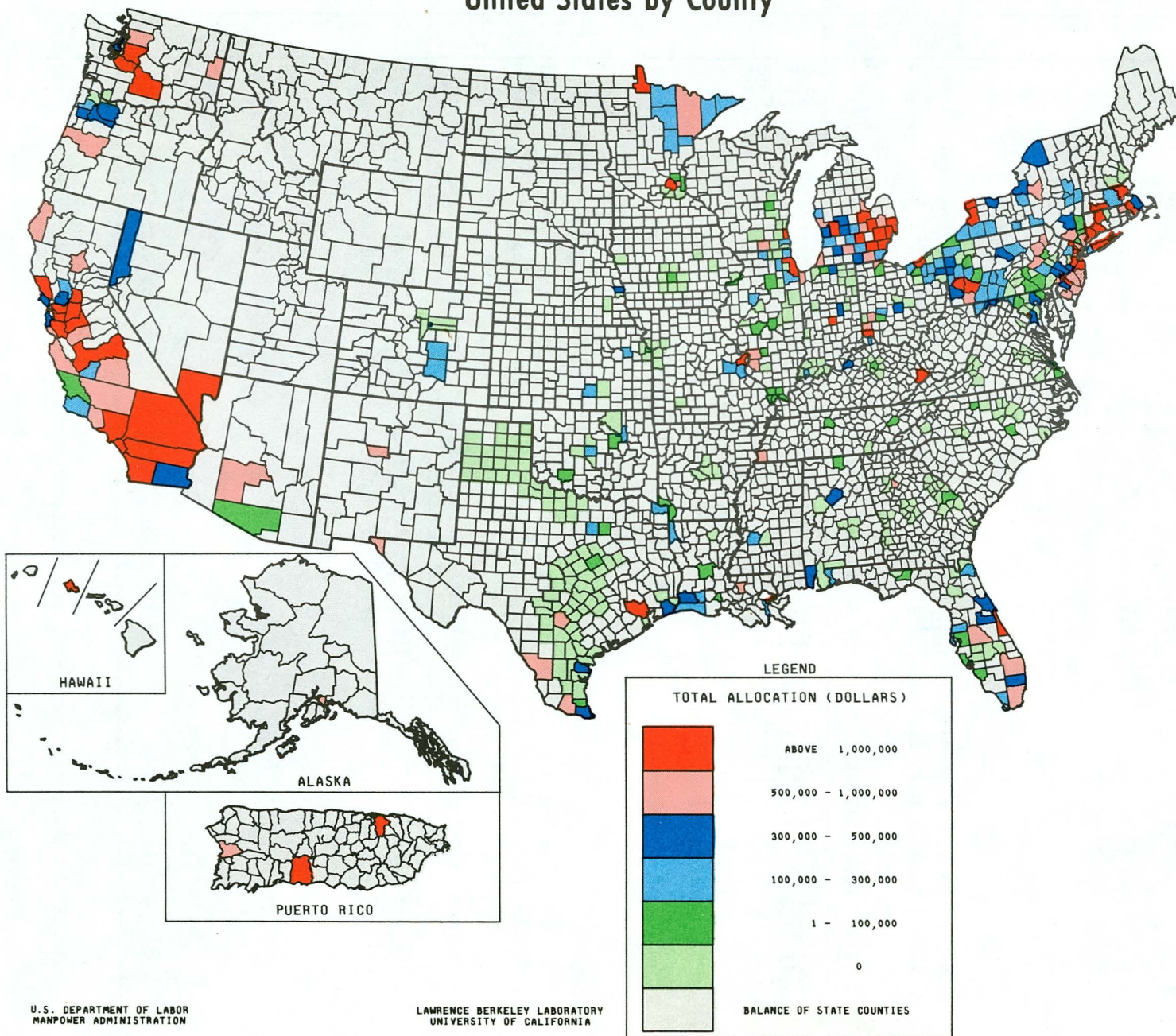


TABLE 4 - ALLOCATION OF TOTAL CETA TITLE II FUNDS
FOR PUBLIC SERVICE EMPLOYMENT - FY 1974

PAGE 11
RUN DATE 03/01/75

U.S. DEPARTMENT OF LABOR
MANPOWER ADMINISTRATION

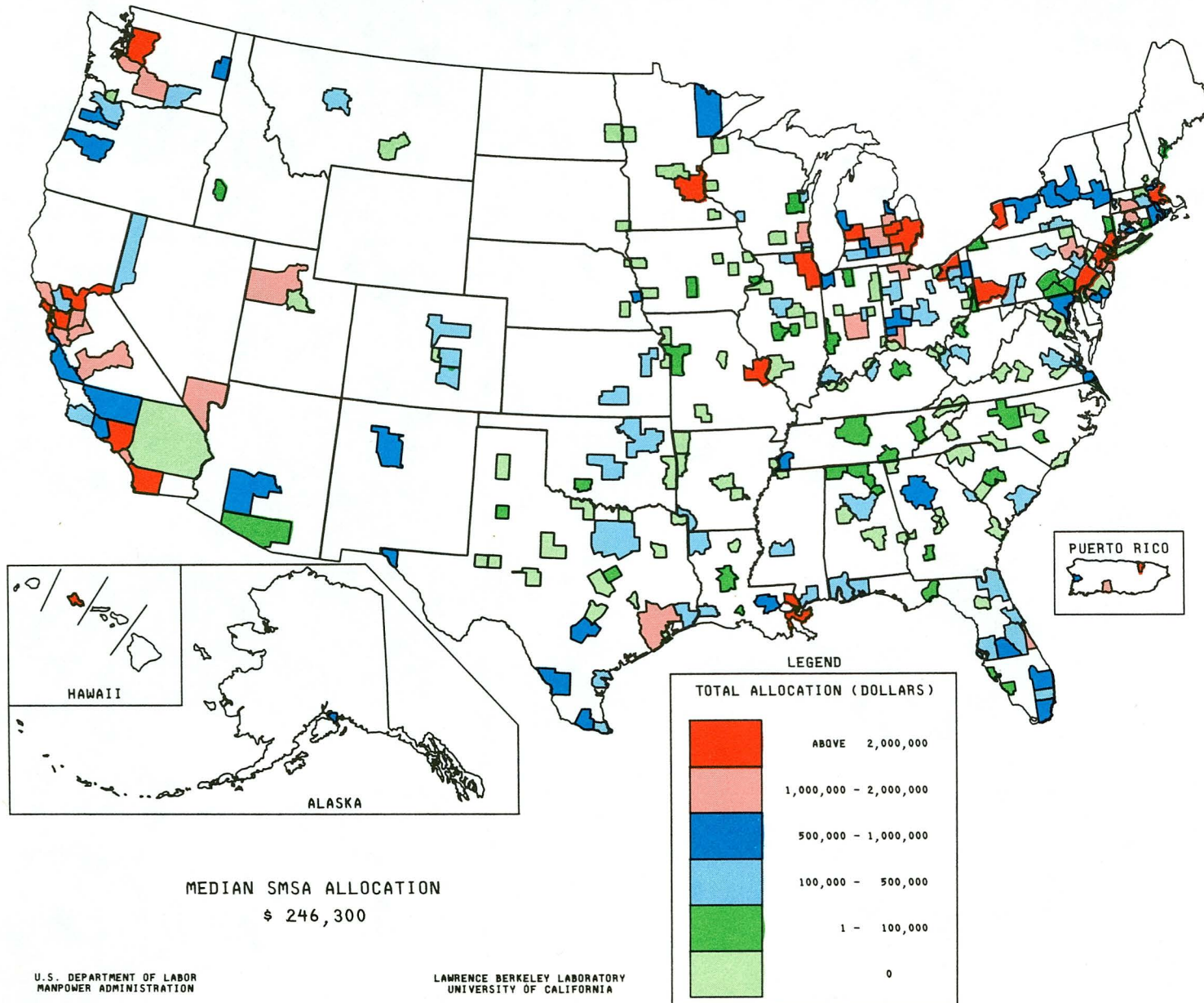
UNITED STATES BY STANDARD METROPOLITAN STATISTICAL AREA

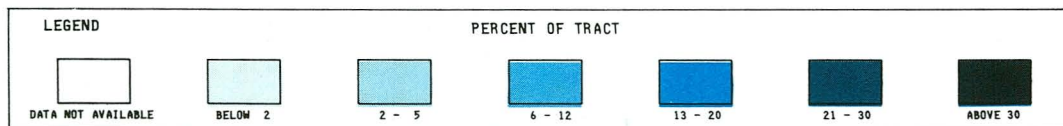
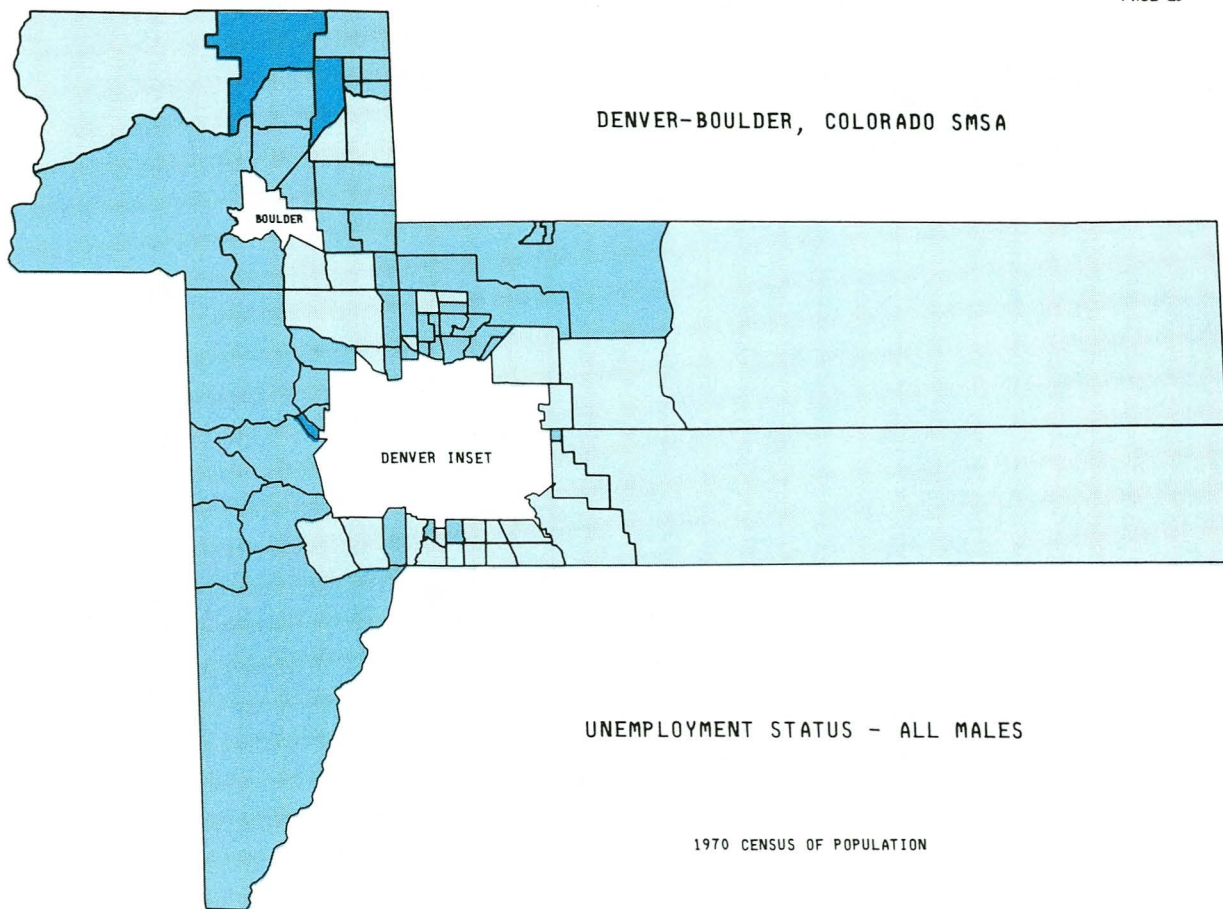
LAWRENCE BERKELEY LABORATORY
UNIVERSITY OF CALIFORNIA

SMSA	TOTAL ALLOCATION (DOLLARS)	PERCENT OF NATION
Florence, Ala	40,400	.02
Bay City, Mich.	624,883	.25
Sioux City, Iowa-Nebr.	0	.00
Tuscaloosa, Ala	0	.00
Danbury, Conn	149,600	.06
Monroe, La	0	.00
Williamsport, Pa	111,277	.04
Texarkana, Tex-Texarkana, Ark	321,358	.13
Boise City, Idaho	48,100	.02
Lafayette, La	76,300	.03
Lafayette-West Lafayette, In	0	.00
Tallahassee, Fla.	35,300	.01
Lawton, Okla.	83,000	.03
Wilmington, Nc	0	.00
Fort Myers, Fla.	56,900	.02
Gainesville, Fla.	0	.00
Bloomington-Normal, Ill.	0	.00
Anniston, Ala	0	.00
Elmira, Ny	223,300	.09
St. Joseph, Mo.	0	.00
Fitchburg-Leominster, Mass.	0	.00
Tyler, Tex.	0	.00
Pittsfield, Mass.	188,700	.08
Albany, Ga	90,000	.04
Burlington, Nc.	0	.00
Sioux Falls, S. Dak.	0	.00
Gadsden, Ala.	75,216	.03
Richland-Kennesaw, Wash.	275,900	.11
Odessa, Tex.	0	.00
Dubuque, Iowa	0	.00
Billings, Mont.	0	.00
Nashua, Nh.	0	.00
Pine Bluff, Ark	0	.00
Rochester, Minn.	0	.00
Sherman-Denison Tex	0	.00
Great Falls, Mont.	181,200	.07
Columbia, Mo.	0	.00
La Crosse, Wis.	150,100	.06
Owensboro, Ky.	0	.00
Laredo, Tex	594,858	.24
Lewiston-Auburn, Maine	0	.00
San Angelo, Tex.	0	.00
Bristol, Conn.	198,500	.08
Midland, Tex	0	.00
Bryan-College Station, Tex.	0	.00
Meriden, Conn.	313,400	.12
PUERTO RICO		
San Juan	3,601,929	1.43
Ponce	1,475,106	.59
Caguas	856,384	.34
Mayaguez	622,196	.25

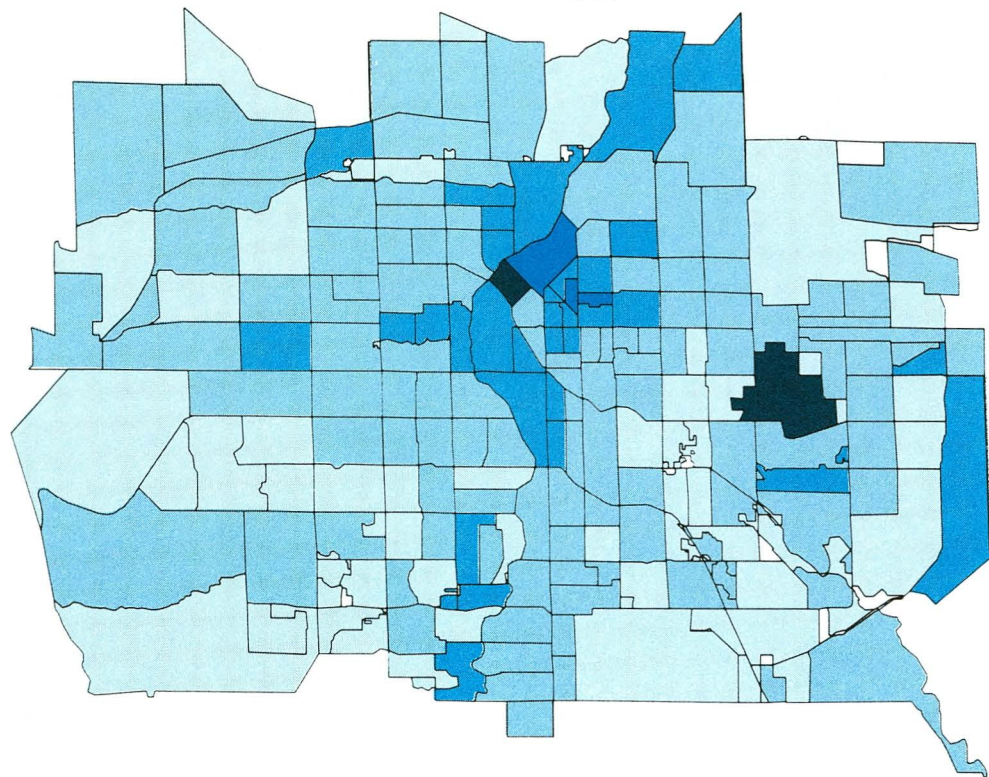
MEDIAN ALLOCATION BY SMSA \$ 246,300
MEAN ALLOCATION BY SMSA \$ 930,100

MAP 4 - Allocation of Total CETA Title II Funds For Public Service Employment - FY 1974 United States by Standard Metropolitan Statistical Area



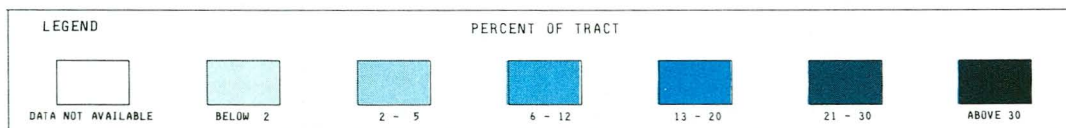


DENVER AREA INSET

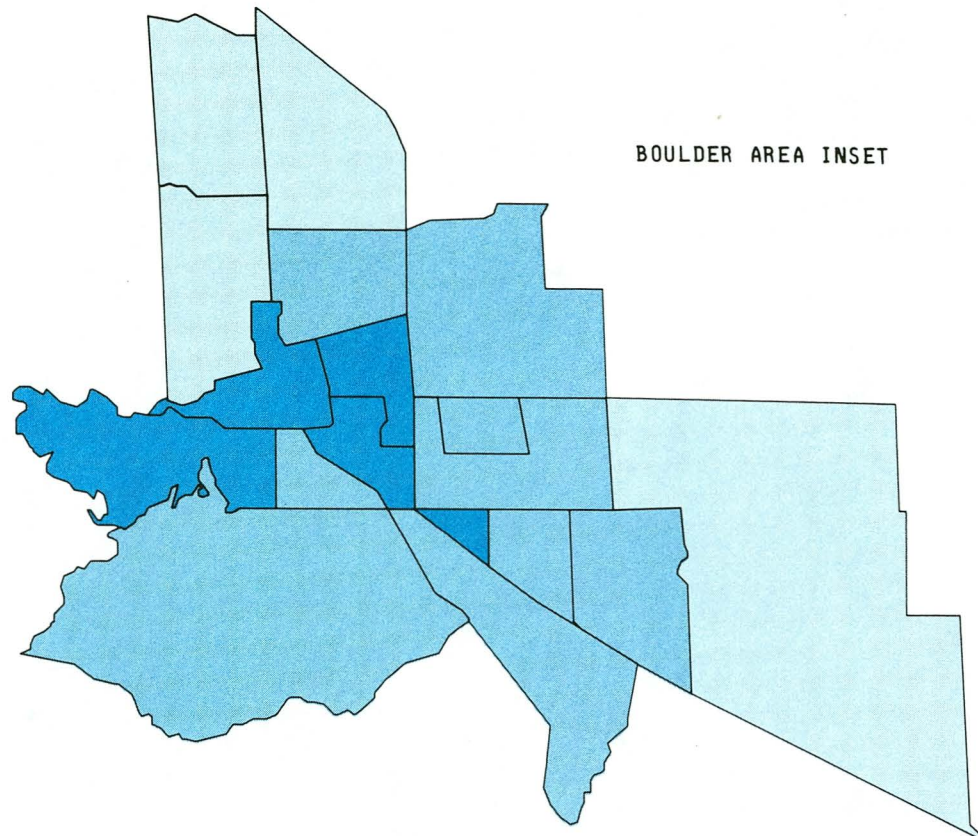


UNEMPLOYMENT STATUS - ALL MALES

1970 CENSUS OF POPULATION

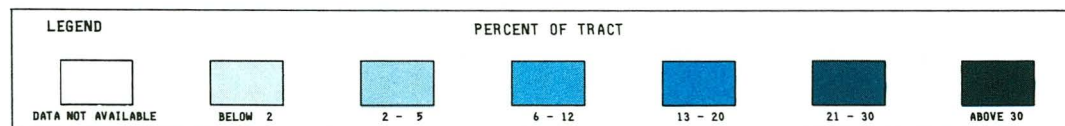


BOULDER AREA INSET



UNEMPLOYMENT STATUS - ALL MALES

1970 CENSUS OF POPULATION



SOCIO-ECONOMIC STUDY
LOCK AND DAM 26
UPPER MISSISSIPPI RIVER
AND
ILLINOIS RIVER

1970 CENSUS OF POPULATION

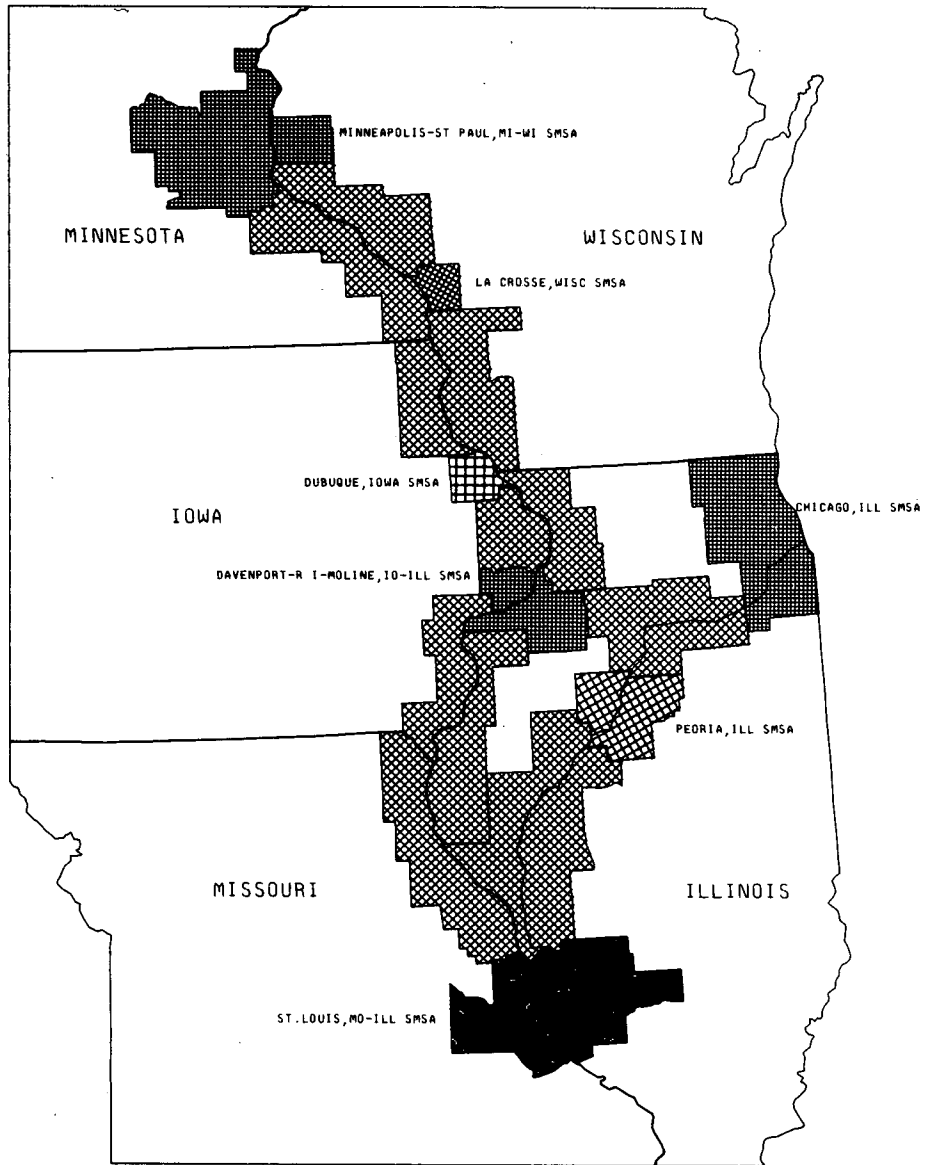
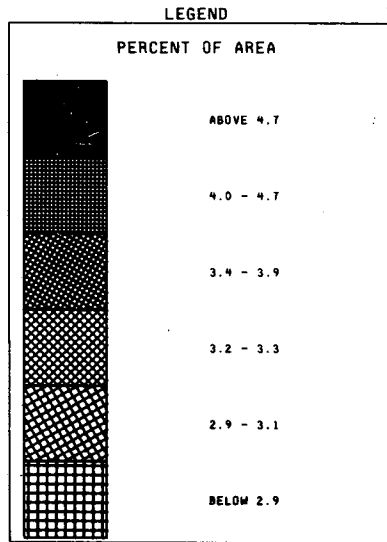


TABLE 83 -- EMPLOYMENT IN THE PUBLIC ADMINISTRATION INDUSTRY

RUN DATE 01/10/75

SOCIO-ECONOMIC STUDY, LOCK AND DAM 26

ST. LOUIS, MISSOURI DISTRICT
U.S. ARMY CORPS OF ENGINEERS

UPPER MISSISSIPPI - ILLINOIS RIVERS

LAWRENCE BERKELEY LABORATORY
UNIVERSITY OF CALIFORNIA

SMSA AND NON-SMSA COUNTY AGGREGATION AREAS	TOTAL EMPLOYED ALL INDUS.	PUBLIC ADMIN. EMPLOYMENT	PERCENT OF AREA	PERCENT OF REGION
REGION TOTAL	5,382,425	238,396	4.4	100.0
Illinois River	138,278	4,628	3.3	1.9
Chicago, Ill Smsa	2,852,017	126,867	4.4	53.2
Peoria, Ill Smsa	134,501	3,954	2.9	1.7
St. Louis, Mo-II Smsa	914,474	51,355	5.6	21.5
Davenport-R. Island, Io-II Smsa	139,926	6,570	4.7	2.8
Dubuque, Io Smsa	33,408	805	2.4	.3
La Crosse, Wi Smsa	30,005	1,028	3.4	.4
Minneapolis-St. Paul, Mn-Wi Smsa	815,273	32,685	4.0	13.7
Upper Mississippi River	324,543	10,504	3.2	4.4

MAP 13 - Allocation of Total CETA Title II Funds For Public Service Employment - FY 1974 Region IX by Prime Sponsor

ARIZONA - CALIFORNIA - NEVADA -
HAWAII - AMERICAN SAMOA - GUAM -
TRUST TERRITORIES OF THE PACIFIC ISLANDS
FEDERAL REGION IX

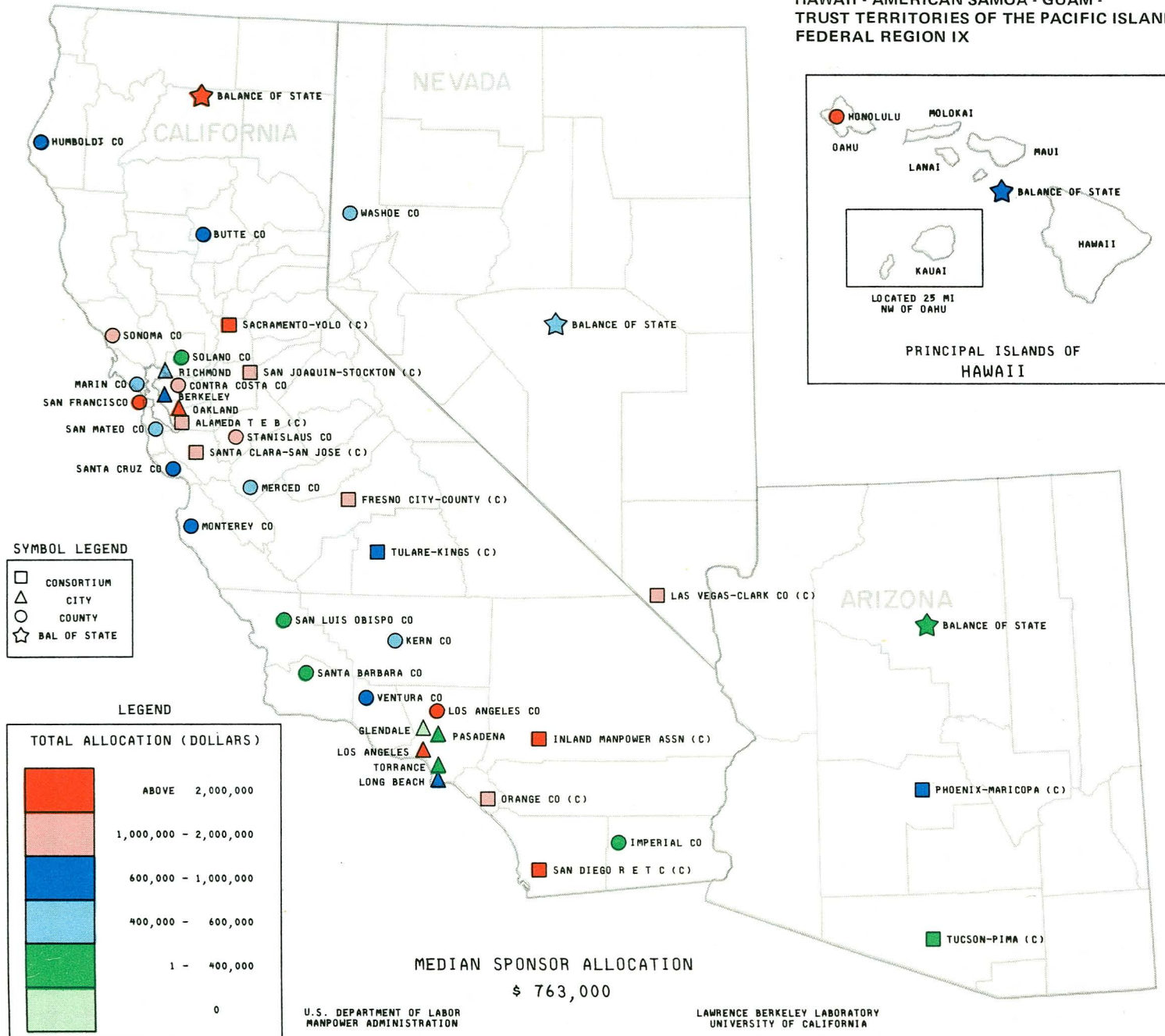


TABLE 13 - ALLOCATION OF TOTAL CETA TITLE II FUNDS
FOR PUBLIC SERVICE EMPLOYMENT - FY 1974

RUN DATE 12/13/74

U.S. DEPARTMENT OF LABOR
MANPOWER ADMINISTRATION

REGION IX BY PRIME SPONSOR

LAWRENCE BERKELEY LABORATORY
UNIVERSITY OF CALIFORNIA

PRIME SPONSOR	TOTAL ALLOCATION (DOLLARS)	PERCENT OF REGION	PRIME SPONSOR	TOTAL ALLOCATION (DOLLARS)	PERCENT OF REGION
REGION IX	70,984,153	100.00	Merced Co	512,005	.72
ARIZONA	1,002,288	1.41	Monterey Co	724,434	1.02
Phoenix-Maricopa [C]	763,288	1.08	Oakland	2,233,139	3.15
Tucson-Pima [C]	98,500	.14	Pasadena	180,824	.25
Balance Of State	140,500	.20	Richmond	429,230	.60
CALIFORNIA	64,769,428	91.24	San Francisco	3,587,946	5.05
Alameda T E B [C]	1,752,888	2.47	San Luis Obispo Co	29,200	.04
Fresno City-County [C]	1,660,842	2.34	San Mateo Co	485,200	.68
Inland Manpower Assn [C]	3,730,289	5.26	Santa Barbara Co	277,600	.39
Orange Co [C]	1,769,800	2.49	Santa Cruz Co	741,589	1.04
Sacramento-Yolo [C]	2,224,452	3.13	Solano Co	338,600	.48
San Diego R E T C [C]	6,829,161	9.62	Sonoma Co	1,044,513	1.47
San Joaquin-Stockton [C]	1,346,328	1.90	Stanislaus Co	1,704,636	2.40
Santa Clara-San Jose [C]	1,506,293	2.12	Torrance	58,000	.08
Tulare-Kings [C]	792,683	1.12	Ventura Co	644,600	.91
Berkeley	891,643	1.26	Balance Of State	5,212,978	7.34
Butte Co	680,703	.96	HAWAII	2,817,732	3.97
Contra Costa Co	1,332,931	1.88	Honolulu	2,079,382	2.93
Glendale	0	.00	Balance Of State	738,350	1.04
Humboldt Co	730,005	1.03	NEVADA	2,049,406	2.89
Imperial Co	388,857	.55	Las Vegas-Clark Co [C]	1,135,347	1.60
Kern Co	594,905	.84	Washoe Co	440,300	.62
Long Beach	974,187	1.37	Balance Of State	473,759	.67
Los Angeles	10,324,021	14.54	A.SAMOA-GUAM-TRUST TERRITORIES	345,300	.49
Los Angeles Co	8,549,332	12.04			
Marin Co	485,600	.68			

MEDIAN ALLOCATION BY PRIME SPONSOR

\$ 763,000

MEAN ALLOCATION BY PRIME SPONSOR

\$ 1,571,000

—LEGAL NOTICE—

This report was prepared as an account of work sponsored by the United States Government. Neither the United States nor the United States Energy Research and Development Administration, nor any of their employees, nor any of their contractors, subcontractors, or their employees, makes any warranty, express or implied, or assumes any legal liability or responsibility for the accuracy, completeness or usefulness of any information, apparatus, product or process disclosed, or represents that its use would not infringe privately owned rights.

TECHNICAL INFORMATION DIVISION
LAWRENCE BERKELEY LABORATORY
UNIVERSITY OF CALIFORNIA
BERKELEY, CALIFORNIA 94720