LATTICE STRUCTURES IN GEOGRAPHY

Alan Saalfeld
Statistical Research Division
Bureau of the Census
Washington, DC 20233

ABSTRACT

Many geographic files, such as the Census Bureau's Master Reference File on regional areal subdivisions, are partially-ordered by inclusion, but do not have a purely hierarchical relation. Although the Master Reference File was initially conceived as a hierarchical file, the necessity of examining additional levels of geography forced the structure to lose the hierarchy originally present. The ability to add levels of geography is a necessary feature of a geographic system. A lattice-based file structure permits the addition and deletion of areal references and furthermore adds special elements and algebraic operations to the set of file elements in order to facilitate searches and general file manipulations.

This paper describes the elementary properties of lattices and presents illustrations from an implementation of a computerized lattice file structure.

INTRODUCTION

In 1980 James Corbett and Marvin White described a mathematical model of maps which encoded geometric and topological relations of 0-cells, 1-cells, and 2-cells via incidence matrices and which proposed encoding regional partition relations and subdivision relations via lattices. The theory of lattices that was needed to implement the mathematical model was not delineated at the time for one good reason: the theory was not yet available. While the topological and geometric theory in Corbett and White's mathematical model of maps had been successfully computerized at the Census Bureau and elsewhere, no one had yet developed a computerized structure for representing the elements and relations of a lattice. Indeed, no one knew the nature and properties of the complete geographic lattice of regions described in Corbett and White's model and some suspected that the underlying set would be enormous and the lattice operations inefficient. Dr. Lawrence Cox of the Census Bureau had studied properties of manually constructed lattices of large geographic tabulation zones for his important work on disclosure analysis; however, the creation and manipulation of the necessary lattices was done entirely by hand.

The mystery of Corbett's lattices comes from two principal sources. The first is that lattices themselves constituted a new and generally unfamiliar mathematical construct for the geographer or cartographer. The second source of mystery is more fundamental. A lattice is an augmentation of the usual set of regions dealt with in geography. There are diverse regional entities of geography such as states, counties, urbanized areas, congressional districts, etc.; and each of the individual entities constitutes an element of a partially-ordered set or poset (ordered by inclusion). Additional elements need to be added to the "natural" partially-ordered set in order to create an "unnatural" lattice; and the number and relations of those newly added elements are not *immediately obvious* even to the experienced mathematician. At least in principle, partially-ordered sets are familiar to geographers. The enclosing lattices are not familiar in name or in principle. Moreover, the enclosing lattices are not even uniquely determined by the underlying partially-ordered sets, all of which adds further to the confusion.

One of the first attempts to remedy the confusion was an internal Census technical document, "Lattice Building," (Saalfeld, 1983), which described for any poset a unique minimal enclosing lattice and which further described an original constructive method for building that lattice. The methods described in the paper were implemented for small sets using relational matrix representations. The matrices, of size n-by-n for posets of size n, limited the size of the posets that could be handled by the new methods.

A second internal Census memo, "Structuring a Poset/Lattice File," (Saalfeld, 1983), presented a data representation strategy that exploited the so-called "bottom-heavy" nature of complete geographic lattices. That strategy was further refined and implemented for small test data sets using B-tree storage routines by Brian Smartt at the Santa Clara County Center for Urban Analysis in 1983.

The purpose of this paper is two-fold: the first is to clear up the mystery of lattices in general with an illustrated overview of basic lattice theory as it applies to geography; and the second is to introduce the foundations research that has taken place at the Census Bureau so that others wishing to implement Corbett and White's model can build upon that research.

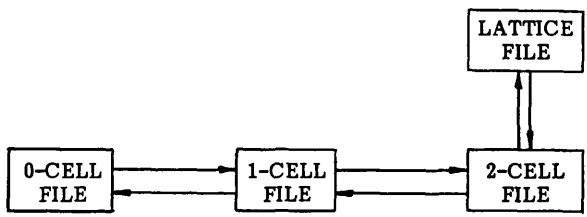

Figure 1. Some Interactions in Corbett and White's Map Model.

The lattice file in Corbett and White's theory permits the geographer to access important combinations of 2-cells in an efficient manner. Internal lattice file operations allow the geographer to add and subtract layers of geography and readily recover the resulting relations of inclusion and intersection of the regions involved.

BASIC CONCEPTS IN LATTICE THEORY

A partially ordered set, or poset, $\{ P, \leq \}$, is a set P with a binary relation $\leq$ between some pairs of elements of P such that:

(1)	For all p in P, $p \leq p$.	(Reflexivity)
(2)	If $p \leq q$ and $q \leq r$, then $p \leq r$.	(Transitivity)
(3)	If $p \leq q$ and $q \leq p$, then $p = q$.	(Antisymmetry)

A lattice is a poset in which every pair of elements posesses a unique least upper bound and a unique greatest lower bound.

A least upper bound is also called a meet and notation used is: lub(x,y) or $x V y$. A greatest lower bound is also called a join, and notation used is: glb(x,y) or $x \Lambda y$.

A least upper bound is comparable to every upper bound:

(4) $x \leq x V y$ and $y \leq x V y$.
(5) If $x \leq p$ and $y \leq p$, then $x V y \leq p$.

Similarly a greatest lower bound is comparable to every lower bound.

(6) An upper ideal in a poset P is a subset X of P such that if x is in X and $x \leq y$, then y is in X also.

In other words, an upper ideal contains every element that is bigger than any one of its elements. Similarly one may define a lower ideal:

(7) A lower ideal in a poset P is a subset Y of P such that if y is in Y and $z \leq y$, then z is in Y also.

Given any set X in a poset P, the collection of all elements in P simutaneously greater than or equal to all elements of X forms an upper ideal, denoted X^*, where

$$X^* = \{ p \in P \mid \forall x \in X,\ x \leq p \}.$$

Similarly X_* represents the lower ideal:

$$X_* = \{ p \in P \mid \forall x \in X,\ p \leq x \}$$

It is easy to show that X is contained in X^* (or X_*) if and only if X is a singleton, {x}. It is also easy to show that X is always contained in $(X^*)_*$ and in $(X_*)^*$.

Furthermore, the operators upper star, *, and lower star, $_*$, are order reversing. That is, if $Y \subseteq X$ as subsets of a poset, then $X^* \subseteq Y^*$ and $X_* \subseteq Y_*$. The operators * and $_*$ behave as follows with intersections and unions of subsets:

(8) $(A \cup B)^* = A^* \cap B^*$,

(9) $(A \cup B)_* = A_* \cap B_*$,

(10) $(A \cap B)^* \supseteq A^* \cup B^*$,

(11) $(A \cap B)_* \supseteq A_* \cup B_*$.

In general, equality does not hold in (10) and (11).
The following properties of the upper star and lower star operators are easily verified.

(12) For all $X \subseteq P$, $X^* = ((X^*)_*)^*$

(13) For all $Y \subseteq P$, $Y_* = ((Y_*)^*)_*$

(14) For all $X \subseteq P$, $(X^*)_* = (((X)^*)_*)^*)_*$

Property (14) follows from either property (12) or (13). However, it is property (14) that makes the combined operator, $(\ ^*)_*$, a closure-type operator in the sense that the operator is idempotent, i.e. applying the operator twice gives the same result as applying it once.

Property (14) is also useful for limiting the number of generating sets X one must examine in order to study all sets of the form $(X^*)_*$. One may regard property (14) as saying that one need only look at X's which are already lower ideals (and, more specifically, X's which are already lower ideals of the form $(Y^*)_*$ for some Y).

An upper ideal, I, is called principal if there is an element x in P such that

$$I = \{ x \}^* = \{ p \in p \mid x \leq p \}$$

Similarly a lower ideal is principal if it has the form $\{ y \}_*$ for some element y.

In a lattice, L, the following results hold:

$$X^* = \{ \text{lub}(X) \}^* \quad \text{for all } X \subseteq L$$
$$X_* = \{ \text{glb}(X) \}_* \quad \text{for all } X \subseteq L$$

Therefore all ideals arising from upper star, *, and lower star, $_*$, operators are principal in a lattice.

THE NORMAL COMPLETION: A MINIMAL CONTAINING LATTICE

The groundwork has been laid for the following main result:

Given any poset, $\{ P, \leq \}$, consider the family $F = \{ (X^*)_* \mid X \subseteq P \}$ of lower ideals. For each $p \in P$, associate the element $(\{p\}^*)_* = \{ p \}_*$ of F. Let the usual set inclusion, $\subseteq$, give F a partial order. Then $\{ F, \subseteq \}$ is a lattice which contains a copy of $\{ P, \leq \}$ under the injective mapping $i : p \to (\{p\}^*)_*$.

The mapping i preserves order: $p_1 \leq p_2$ implies $i(p_1) \subseteq i(p_2)$. Moreover if P is already a lattice, then every element of F is principal and i(P) equals all of F. The lattice $\{ F, \subseteq \}$ is called the normal completion of the poset $\{ P, \leq \}$.

The following are properties of the normal completion operation are easily proved. They form a theoretical basis for formal iterative construction procedures used to build a minimal containing lattice.

If the operator, $\bar{\ }$, represents the normal completion operator, one has (up to isomorphism of posets or lattices):

A. For every poset P, $P \subseteq \bar{P}$.

B. For every poset P, $\bar{P} = \bar{\bar{P}}$.

C. For posets $\{ P, \leq_p \}$ and $\{ Q, \leq_q \}$ with $P \subseteq Q$ and $\leq_q$ equal to $\leq_p$ on elements of P, one has $\bar{P} \subseteq \bar{Q}$.

D. For a poset $\{ P, \leq_p \}$ contained in a lattice $\{ L, \leq_l \}$ (where $\leq_l$ restricted to P is equal to $\leq_p$), one has $\bar{P} \subseteq L = \bar{L}$

E. $P = \bar{P}$ if and only if P is a lattice.

EXAMPLE OF A GEOGRAPHIC POSET

The simple figures below will be used to illustrate definition and methods. Consider the following subdivisions of the same region:

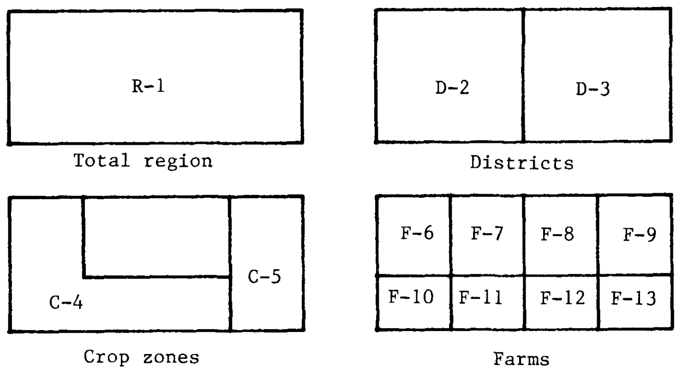

Figure 2. Illustration of Geographic Subdivisions Forming a Poset.

The above geographic relations may be described as a **poset** with the following **Haase diagram:**

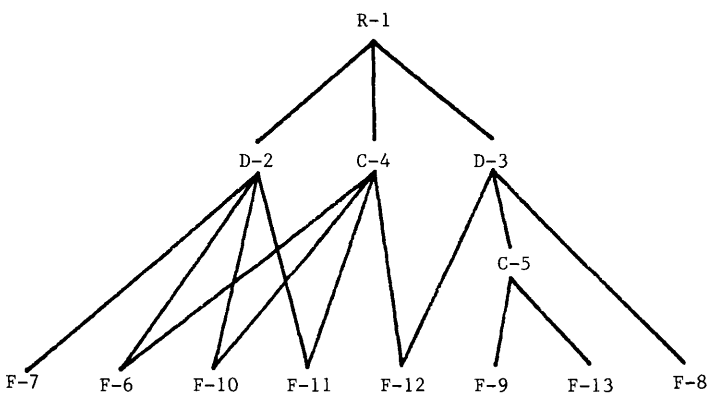

Figure 3. Haase Diagram for Containment Relations of Illustrated Geographic Poset.

The links of a Haase diagram show immediate containment, where there are no intermediate sets. For that reason, no link is drawn between F-9 and D-3 since C-5 lies between Farm 9 and District 3. Notice that the farms are numbered out of order to facilitate the drawing of containment lines with fewer intersections of those lines. At times it is impossible to avoid intersections.

The collection of farms, crop zones, districts and region do not form a lattice, only a poset. Adding the null set as a least element is not sufficient to bring the collection to lattice status. The farms, F-6, F-10, and F-11, have no least upper bound, but rather have two incomparable upper bounds, C-4 and D-2. The addition of the intersection of C-4 and D-2 as a new element will make the collection a lattice.

MATRICES, VECTORS, AND BOOLEAN ARITHMETIC

The above example also serves to illustrate the matrix/vector representation of posets which lend themselves to straightforward arithmetic operations. Consider the above thirteen regions and the inclusions among them; and define the 13-by-13 matrix $(\mathbf{a_{ij}})$ as follows:

$a_{ij} = 1,$ if region **i** is contained in region **j**, and

$a_{ij} = 0,$ otherwise.

It is easy to see that $(\mathbf{a_{ij}})$ looks like:

$$
\begin{matrix}
1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
1 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
1 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
1 & 0 & 1 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
1 & 1 & 0 & 1 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
1 & 1 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\
1 & 0 & 1 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\
1 & 0 & 1 & 0 & 1 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\
1 & 1 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\
1 & 1 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\
1 & 0 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\
1 & 0 & 1 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1
\end{matrix}
$$

Consider the following Boolean arithmetic tables of 0's and 1's:

Addition:

+	0	1
0	0	1
1	1	1

Multiplication

*	0	1
0	0	0
1	0	1

By defining a Boolean arithmetic (logical **and/or** operations) on the entries of the square matrix, one may perform matrix multiplications such as $(\mathbf{a_{ij}})^2$ or $(\mathbf{a_{ij}})^n$. If $(\mathbf{a_{ij}})$ is the matrix of a poset, then $(\mathbf{a_{ij}})^n$ will equal $(\mathbf{a_{ij}})$ for all positive integers **n** because of transitivity and reflexivity. Furthermore, one may check antisymmetry by using the same Boolean arithmetic to verify that $\mathbf{a_{ij}*a_{ji} = 1}$ only when **i=j**.

From the above remarks, one sees that a poset's matrix representation can be quite useful. Moreover, a given set of relations may be expanded to include all reflexive relations by adding 1's to the diagonal and all transitively-implied relations by matrix exponentiation as described in the preceding paragraph. The resulting set of relations may then be tested for antisymmetry to see if it is a poset.

THE HAASE DIAGRAM AND THE HAASE MATRIX

The Haase matrix is simply $(\mathbf{h_{ij}})$ where:

$h_{ij} = 1,$ if element **i** is directly below element **j**, and

$h_{ij} = 0,$ otherwise.

The Haase matrix for the poset used in the illustration above is:

$$
\begin{matrix}
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 1 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 1 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 1 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0
\end{matrix}
$$

If (a_{ij}) is the complete relation matrix of a poset, I is the identity matrix, and (h_{ij}) is the Haase matrix; and if (a_{ij})-I is the matrix (a_{ij}) with the diagonal changed to zeros, then:

$$(h_{ij}) = [\,(a_{ij})-I\,] - [\,(a_{ij})-I\,]^2.$$

The subtractions in the above expression make sense because a **1** entry is never "subtracted" from a **0** entry.

Conversely, to go from (h_{ij}) to (a_{ij}), one has: $(a_{ij}) = [\,(h_{ij})+I\,]^m$ for some integer **m**.

The Haase diagram is a directed graph; and the Haase matrix is the usual relational matrix for a directed graph. Many interesting results from graph theory are applicable. In particular, decomposition results on chains and antichains have proven useful in generating ideals efficiently.

VECTOR ARITHMETIC FOR LATTICE ELEMENT CONSTRUCTION

Other meaningful Boolean operations may be performed on entries in vectors and matrices. Some of these are used to construct the normal completion lattice of a poset. A few of those operations are described briefly here which correspond to the "upper-star" and "lower-star" opeators introduced in the section on the normal completion lattice.

The rows and columns of the poset matrix are vectors which represent all principal upper and lower ideals of the poset, respectively. In each case the **n-vectors** are characteristic vectors for the set of **n** elements: each possible combination of **1**'s and **0**'s corresponds to a possible subset of the **n** elements, with **1** signifying that an element is present, **0** that it is not.

For example, the third column has **1**'s in cells 3, 5, 8, 9, 12, and 13, indicating that element D-3 contains the fifth, eighth, ninth, twelfth, and thirteenth elements in addition to itself. Similarly the sixth row has **1**'s in positions 1, 2, 4, and 6, indicating that the element F-6 is contained in the first, second, and fourth elements in addition to being contained in itself. The third column is the characteristic vector for $(D\text{-}3)_*$; and the sixth row is the characteristic vector for $(F\text{-}6)^*$.

Boolean arithmetic defined termwise on corresponding cells of characteristic vectors amounts to the usual intersection and union of the represented subsets:

Intersection: $(a_i)*(b_j) = (a_k*b_k) = (a_1*b_1,\ a_2*b_2, \ldots, a_n*b_n)$.

Union: $(a_i) + (b_j) = (a_k + b_k) = (a_1 + b_1,\ a_2 + b_2, \ldots, a_n + b_n)$.

The vector/matrix/Boolean manipulations required to create new lattice elements (least upper bounds of a collection of poset elements) are sketched below as an illustration of the required arithmetic procedures:

Given poset elements $\{p_1, p_2, \ldots, p_k\}$, find the characteristic vector of $(\{p_1, p_2, \ldots, p_k\}^*)_*$, the upper-lower-star operator of the set.

First, let (v_{ij}) be the row vector of the poset matrix corresponding to p_i^* for $i=1,2,\ldots,k$. Then $(\{p_1, p_2, \ldots, p_k\}^*)$ has characteristic vector:

$$(v_{1j}*v_{2j}*\cdots*v_{kj}) = (v_{11}*v_{21}*\cdots*v_{k1},\ v_{12}*v_{22}*\cdots*v_{k2}, \ldots, v_{1n}*v_{2n}*\cdots*v_{kn})$$

since the product * corresponds to the intersection $\cap$; and

$\{ p_1, p_2, \ldots, p_k \}^* = p_1^* \cap p_2^* \cap \ldots \cap p_k^*$.

Let $(w_j) = (v_{1j} * v_{2j} * \ldots * v_{kj})$.

Then (w_j) is the characteristic vector for the subset:

$$W = \{ p_1, p_2, \ldots, p_k \}^*,$$

and it remains to find the characteristic vector for W_*.

If (a_{ij}) is the square complete relational matrix of the poset described earlier, then (z_r), the characteristic vector for W_*, is given by:

$$z_r = [w_1 * a_{r1} + (1-w_1)] * [w_2 * a_{r2} + (1-w_2)] * \ldots * [w_n * a_{rn} + (1-w_n)],$$

$$\text{for } r = 1,2,\ldots,n.$$

The expression $(1-w_i)$ above is just the Boolean complement of w_i; that is, $(1-w_i)$ is **0** when w_i is **1**, and **1** when w_i is **0**. By adding (in the Boolean sense) $(1-w_i)$ to each factor, one is saying simply to ignore the factor unless w_i is **1**.

CONCLUSION

This short paper reviews the theoretical foundations and some of the applications of lattices being explored at the Bureau of the Census. The arithmetic examples given above are presented simply to illustrate the computational possibilities in representing posets and lattices as matrices. Ongoing research in geographic lattice applications at the Bureau of the Census is focusing on efficient storage and retrieval of matrices and sparse matrices, and on additional representation and computational strategies for poset and lattice elements and operations.

BIBLIOGRAPHY

Birkhoff, G., 1967, Lattice Theory, Providence, RI, The American Mathematical Society.

Cohn, P.M., 1965, Universal Algebra, New York, NY, Harper and Row.

Cox, L., 1975, "Applications of Lattice Theory to Automated Coding and Decoding," **AutoCarto II**, Reston, Va., ACSM/Census.

Cox, L., 1980, "Suppression Methodology and Statistical Disclosure Analysis," **Journal of the American Statistical Association**, 75(370), pp. 377-385.

Crawley, P., and R.P. Dilworth, 1973, Algebraic Theory of Lattices, Englewood Cliffs, NJ, Prentice-Hall, Inc.

Preparata, F.P., and R.T. Yeh, 1973, Introduction to Discrete Structures, Reading, MA, Addison-Wesley.

Saalfeld, A., 1983, "Lattice Building," internal Census Bureau technical document.

Saalfeld, A., 1983, "Structuring a Poset/Lattice File," internal Census Bureau document.

White, M.S., 1984, "Technical Requirements and Standards for a Multi-purpose Geographic Data System," **The American Cartographer**, Vol. 11, No. 1, pp. 15-26.

TOPICS IN ADVANCED TOPOLOGY FOR CARTOGRAPHY

Alan Saalfeld
Statistical Research Division
Bureau of the Census
Washington, D.C 20233
(301) 763-1496

ABSTRACT:

Cartography has embraced the fundamental principles of combinatorial topology with great benefit to the development of the mathematical theory of maps Many other useful tools available from other areas of topology have gone unrecognized and unused. Several of those areas are illustrated here Usually the cartographer considers only the most basic topological properties of static, simply-connected, two-dimensional manifolds or surfaces. Here we present some of the advanced topological theory that may be utilized to address the more difficult, higher-dimensional, dynamic problems of cartography Some of the problems viewed from a new topological perspective include generalization, deformation over time, 3-D representation and 2-D representation of 3-D features, algebraic operations on map features, unified theory of map projections, and orientation Some topological tools used to examine these problems include homotopy theory, homology and cohomology theory, topological groups and vector spaces, and global analysis These sophisticated tools are simplified for use by the mathematically adept cartographer.