"The Investigation of a Computer-Aided Technique for the Solution to the Space Allocation Problem"

by

Peter G. Rowe

A THESIS SUBMITTED IN PARTIAL FULFILLMENT OF THE REQUIREMENTS FOR THE DEGREE OF

Master of Architecture in Urban Design

Thesis Director's signature:

Houston, Texas

May 1971
ABSTRACT

"The Investigation of a Computer-Aided Technique for the Solution of the Space Allocation Problem"

by

Peter G. Rowe

A general class of problems which frequently arises in the course of architectural and planning design activity is the development of an appropriate configuration of spaces in proper functional relationship to each other. This is referred to as the space allocation problem.

The initial statement of such a problem generally occurs during the programming phase of the design activity and is most often able to be expressed in the form of a set of simple verbal statements. These are descriptions of the proximal nature of the relationship between a set of spatial units and can be said to fully describe certain major locational requirements of a specific design problem. The solution to such a problem must maximize the satisfaction of these locational requirements and be capable of being expressed by means of an appropriate graphic representation to allow further information processing.
This paper proposes an alternative approach to the solution of this general class of problems than has been traditional, utilizing a computer-aided technique. The problem is first formulated in a precise form suitable for systematic information processing. A general strategy by which this problem can be optimally resolved is outlined and the computer program involved, CAPLAN, presented and described in detail. This solution strategy basically involves the use of a specially developed cluster-seeking algorithm with a graphic output capability.

Several representative and easily interpretable case studies are developed to further describe the application of this specific technique in the solution of the space allocation problem. Some experimental validation of the technique is presented in support of the efficacy of the approach.

The investigation shows that this particular technique performs all the necessary supportive analysis for solution of the space allocation problem, as stated, in a manner which is deterministic and readily interpreted.
# TABLE OF CONTENTS

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>I.</strong> INTRODUCTION</td>
<td>1</td>
</tr>
<tr>
<td><strong>II. ASPECTS OF THE SPACE ALLOCATION PROBLEM</strong></td>
<td>3</td>
</tr>
<tr>
<td>1. State of the Information</td>
<td></td>
</tr>
<tr>
<td>2. Complexity of the Task</td>
<td></td>
</tr>
<tr>
<td>3. Traditional Solution to the Problem</td>
<td></td>
</tr>
<tr>
<td><strong>III. PROBLEM FORMULATION</strong></td>
<td>9</td>
</tr>
<tr>
<td>1. Statement of the Problem</td>
<td></td>
</tr>
<tr>
<td>2. Spatial Elements</td>
<td></td>
</tr>
<tr>
<td>3. Interrelationships</td>
<td></td>
</tr>
<tr>
<td><strong>IV. TECHNIQUE OF ANALYSIS</strong></td>
<td>14</td>
</tr>
<tr>
<td>1. Solution Strategy</td>
<td></td>
</tr>
<tr>
<td>2. The Distance Coefficient</td>
<td></td>
</tr>
<tr>
<td><strong>V. DESCRIPTION OF COMPUTER PROGRAMS CAPLAN AND PLANDS</strong></td>
<td>24</td>
</tr>
<tr>
<td>1. General Characteristics</td>
<td></td>
</tr>
<tr>
<td>2. The Cluster-finding Algorithm</td>
<td></td>
</tr>
<tr>
<td>3. Structural Characteristics of the Clusters</td>
<td></td>
</tr>
<tr>
<td>4. Input Modes</td>
<td></td>
</tr>
<tr>
<td>5. Control Parameters</td>
<td></td>
</tr>
<tr>
<td>6. Command Language</td>
<td></td>
</tr>
<tr>
<td>7. Output Mode and Representation of the Analysis</td>
<td></td>
</tr>
<tr>
<td><strong>VI. RESULTS</strong></td>
<td>42</td>
</tr>
<tr>
<td>1. A Partial Proof</td>
<td></td>
</tr>
<tr>
<td>2. Case Study No. 1 - Domestic Suburban Dwelling</td>
<td></td>
</tr>
<tr>
<td>3. Case Study No. 2 - Hospital Complex Master Plan</td>
<td></td>
</tr>
<tr>
<td>4. Discussion</td>
<td></td>
</tr>
<tr>
<td><strong>VII. SUMMARY</strong></td>
<td>75</td>
</tr>
<tr>
<td>1. Conclusion</td>
<td></td>
</tr>
<tr>
<td>2. Future Work</td>
<td></td>
</tr>
<tr>
<td><strong>VIII. BIBLIOGRAPHY</strong></td>
<td>78</td>
</tr>
</tbody>
</table>
I. INTRODUCTION

It is difficult to imagine any argument with a general description of design activity as a goal oriented or purposive form of human behavior. Design activity arises out of the recognition of a need for which it attempts to provide the best solution, either in the form of an idea (solution image) or as a piece of hardware - a designed artifact. The solution is invariably expressed by the way in which physical form is organized to meet that need.

One of the essential aspects of a design problem is the development of an appropriate arrangement of spaces in proper functional relationship to each other, as expressed by that need. This is the space allocation problem, and it represents one of the more complex tasks in design problem solving.

This paper is intended to propose an alternative approach to the solution of this problem utilizing a computer-aided technique. This treatment begins by recognizing aspects of the problem as it initially arises in problem solving activity and briefly discusses the traditional approach to its solution. On the basis of this discussion, the problem is formulated in a
more precise form suitable for systematic information processing.

A general strategy by which this problem can be optimally resolved is outlined and the computer program involved described in detail. Several representative and easily interpretable case studies are developed and illustrated to further describe the application of the technique and as a demonstration of its effectiveness. An experimental validation of the technique is presented to a classical form of the space allocation problem, as something of a proof of the efficacy of the approach.

The intent of this paper is to present and investigate an alternative technique by which the necessary analysis for the solution of the space allocation problem can be performed and an arrangement in the form of a diagramatic spatial representation directly generated.
II. ASPECTS OF THE SPACE ALLOCATION PROBLEM

1. State of the Information

A class of problems which frequently arises in the course of architectural and planning design activity is the development of an appropriate configuration of spaces in proper functional relationship to each other. This is referred to as the space allocation problem.¹

The initial statement of such a problem generally occurs during the programming phase of the design activity and is most often able to be expressed in the form of a set of simple verbal statements.² For instance, in the case of a domestic dwelling: "The kitchen must be next to the dining facilities" or "The living room must have access to the front door."

It can be readily recognized that the essential features of such verbal, functional statements are that they consist of a spatial unit description, e.g., "kitchen, dining room," etc., and some specification of the degree of the desired physical relationship between them, e.g., "must have"... etc. In other words, they are clear
single specifications of an important set of functional requirements of a building or designed artifact.\(^3\)

The spatial units involved are physically identifiable spaces which are acceptable fundamental units of the physical environment, in the context of a specific problem. At an elemental level we can regard them as physical space or facility descriptions having associated with them the idea of a single activity, and we shall refer to them as spatial elements.\(^4\) Perhaps an immediate conclusion that might be drawn is that these spatial elements are what we refer to with common room names. However, if we think of a space in this way, e.g., "living room," there would seem to be no simple relationship between the space and the activities that can commonly occur in this space. For instance, it can usefully be divided into such things as: a) a place where visiting groups aggregate and talk, b) a place of inter-family entertainment, c) a place of indoor child recreation and so on.

2. Complexity of the Task

The relationship between the spatial elements is the
critical aspect of the problem. These must be resolved for a particular arrangement of spaces to be said to function in a desired manner. In a building of any complexity this turns out to be a complex task. The initial statement of the problem, in the form just described, often contains functional requirements that are in some way mutually contradictory and cannot be solved by the use of a simple (syllogistic) logical procedure. Such as, if A goes next to B and C next to B then C is in the desired relationship to A. Trade-offs have to be made, as different spatial elements compete with one another for proximity to some other set of spatial elements. (The real built solution only provides for the most part 2, and to a lesser extent, 3 dimensions in which to manouver this problem). An acceptable solution to a problem of this type would seem to lie in finding a solution which most satisfactorily answers all the prescribed functional requirements in the best possible manner. An optimum solution would seem desirable.

Another aspect of the problem which further adds to its complexity is the extent of the information it is required
to process at one time. For instance, if we say
the kitchen must be next to the dining room and be away
from the front entrance, whereas the dining room must
be part of the formal entertainment area having direct
access to the front door, in locating any one of these
spatial elements, the required interrelationships with
the other areas must also be resolved. This is a task
of complex information processing.

Recent investigations in the psychology of problem-
solving and human attention and performance have led
to the general conclusion that the human information
processor has a rather severely bounded capability in
dealing with situations of this type. The number of
chunks of information that can be simultaneously pro-
cessed in a designer's operational memory in a prob-
lem situation of this type places severe restrictions on
his ability to directly resolve even a small set of func-
tional requirements necessary in producing an adequate
solution to a given space allocation problem.

3. Traditional Problem Solution

This aspect of a designer's information processing
capability has strongly influenced the manner in which he resolves the problem. The traditional approach to the problem has been to draw on some past body of experience and to simplify the problem into a number of information units which can be successfully handled.

In other words, aggregations of spaces, or their general logic, which have been shown to be in some way contiguous or strongly related and functionally satisfactory for a given problem type are organized with respect to one another. Following this, a further breakdown of the aggregations is made until all the spatial elements are resolved into an appropriate configuration.

Needless to say, the problem can be solved. However, one would suspect on somewhat a priori grounds that solution approaches of this type, comparative to a direct approach, are subject to over-simplification and are liable to the use of grosser forms of false premises.

The transformation of formal types outlined by Ambasz, in which the prototype solutions of one era are an outgrowth of similar solutions in a previous period, is in many ways demonstrative of the more traditional approach. Clearly the solution to the expression of
physical functional relationships of a new solution are contingent in their organization upon some preceding, presumably satisfactory solution. Here, design, by the incremental modification of prototypes, simplifies a designer's task considerably. 10

The current use of an alternative form of information processing in the form of a digital computer has given rise to a specific area of design investigation referred to as Computer-Aided Design. The purpose of these investigations is to provide the means by which the capability of a designer as an effective problem-solver can be reinforced.

In the following sections it is proposed to illustrate an alternative method, using a computer-aided technique by which the space allocation problem can be directly and efficiently dealt with.
FOOTNOTES

1. The problem class has been referred to by this name, e.g., DMG publications; Proceedings of EDRA-1 (1969) and EDRA-2 (1970), etc.


3. The need for one spatial unit to be placed in close proximity with another is an important functional requirement of a designed artifact. However it is readily acknowledged that there are also other important functional considerations not necessarily pertaining to proximity. The problem here is confined to this one aspect of a design problem.

4. To Alexander and Poyner (1966), p. 6: "The atoms of environment structures are relations. Relations are geometric patterns. They are the simple geometric patterns in a building which can be functionally right or wrong."
   To Havilland (1967), p. 4: "...... activity/ space...... activity which takes up space and has a generally common set of facility implications."
   For further discussion see: Sommer (1969); Hall (1967); Proshansky (1970).


6. Miller (1956), ..... experiments on digit span, numerosity judgement and the span of immediate memory impose severe restrictions on the amount of information that we are able to receive, process and remember. By organizing the stimulus input simultaneously into several dimensions and successively into a series of chunks, we manage to break this informational
bottleneck to some extent. (7± chunks...)

Waugh and Norman (1965), ... experiments showing that in a situation where interception occurs in compatible short term memory input/output tasks only the first 2 of a sequence of items are retained with reliability.

7. Past experience. The intended meaning here is knowledge of a situation held intimately by a designer or a corporate body of knowledge to which he has or could have access.


"...for the cultural process' progress is not linear, and prototypes are made with parts of other prototypes, types and stereotypes. Indeed it is in the form of this combination that the prototype truly occurs."
...(A Jungian view.)
See also, Baird and Jencks (1969).

10. For further discussion see:
Van der Ryn (1966); Gombrich (1959); Alexander (1968).
III. PROBLEM FORMULATION

1. Statement of the Problem
   To find an appropriate configuration of spatial elements which maximizes the possible satisfaction of a set of functional requirements, regarding the spatial inter-relationships of a set of spatial elements, deduced from a set of verbal statements describing specific aspects of the problem. For the solution to be capable of generating further solution representation in the form of an "affinity diagram" and consequently a schematic plan of a building, or building complex.

2. Spatial Elements
   Further to the foregoing discussion the spatial elements may be said to have the following properties:
   a. Representative of an activity or activities that has any physical design implication.
   b. That is sufficiently important in the problem context to warrant a designer's attention.
   c. That is elemental in the sense that there is no useful information to be gained by further subdividing the spatial element into smaller units in the context
of a given problem. (We have suggested that perhaps the most elemental level could be described as a spatial element description associated with a single activity.)

3. Interrelationships

Characteristically, the strength of a relationship between various spatial elements is often expressed by terms such as "adjacency," "next to," and the like. In some problems they have been expressed in terms of the desired real distance between the units as an expression of proximal relationship. A variety of scoring systems has been devised for expressing this relationship.

In general terms we would like to be in a position to fairly immediately translate the meaning expressed in a set of verbal statements describing the problem, as well as the rarer situations of an expression of actual desired distance between the spatial elements. It is also clear that, in addition to being able to consider the positive aspects of proximal relationships such as "must be next to" or "should be next to," the negative
aspect such as "must be isolated from" etc., should be incorporated.

As a general scale for evaluating the relative strength of the relationship, it is proposed to express it in the following manner: (Note should be made that an individual designer in using this technique is not necessarily confined to this notation system for a particular problem)

(+++) must have a close proximal relationship (next to, etc.)

(+) desirable to have a proximal relationship

(0) indifference to proximal relationship

(-) desirable to have some measure of isolation

(-->) must be isolated from a given spatial element.

It is felt that this formal expression of proximal relationship is consistent with the level of detail and meaning inherent in the basic data base (statements) and is hence appropriate for formally describing the spatial relationships between a set of spatial elements.

In order to incorporate this scale in the technique, we can express it in terms of an interval scale of positive integers, for instance:
++ - 4
+ - 3
0 - 2
- - 1
-- - 0  (......a five-point scale).

(The difference in the description of relationship
strength is ultimately expressed in the same manner as
a notation containing negative integers, e.g., ++ = +2;
-- = -2. For example, from -- to ++ is four units on
both scales.)

The most convenient formal expression of this aspect
of the problem is in the form of a symmetric inter-
action matrix (see case studies). This representation
expresses all the desired pair-wise relationships be-
tween the spatial elements and forms a specific and
precise representation of the problem as stated, without
any undue loss of information in the translation process.

The use of an interval scale scoring technique is a
divergence on the part of this method from other simi-
lar problem formulations. In the past the scoring has
been expressed in the form of "match-mismatch,"
(1 or 0) decision-making. This was felt to conform to
the theoretical idea of "satisficing" decision-making,
characteristic of much of design problem-solving.

(An expression of whether a potential solution or concept does or does not satisfy a set of requirements.)

However, the work of Bruner and others\(^3\) has developed the area of relative human judgement in decision-making situations; whereby the decision-making can often express (accurately) how "far off the mark" a particular solution is, in fact. Clearly in the problem illustrated here the designer is in a position to say a little more about a relationship between two spatial elements, apart from the fact that they do or do not interact, and an interval scale form of scoring seems applicable.\(^4\)
FOOTNOTES

1. Techniques of a similar generic type (though not necessarily in their specific application) use "match-mismatch" scoring systems or an approximate, e.g., CLUSTR Milne (1968); VTCON 2 Owens (1967); HIDESS series Manheim and Alexander (1963); CLUMP1, 2 & 3 Mitchell (1969-70); MATRAN Miller (1969).

2. "Satisficing Decision-Making" - A "satisficing" solution is what is referred to in Operations Research as one of the feasible solutions which meets all design requirements and goals. For further discussion see: Simon (1957); Milne (1967), p. 5.


4. Blalock (1960), chapter 4 on interval scales.
IV. TECHNIQUE OF ANALYSIS

1. Solution Strategy

The central issue confronting a technique for the solution of this class of problems is the formation of aggregations of the spatial elements in a manner which maximizes the formal expression of their interrelationships.

Each spatial element has associated with it a profile of positive integers expressing the strength of its proximal relationship to the other elements.

<table>
<thead>
<tr>
<th></th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Kitchen</td>
<td></td>
<td></td>
<td></td>
<td>4</td>
</tr>
<tr>
<td>2. Formal Eating</td>
<td>4</td>
<td>4</td>
<td></td>
<td></td>
</tr>
<tr>
<td>3. Child Bedroom</td>
<td>1</td>
<td>1</td>
<td>4</td>
<td></td>
</tr>
<tr>
<td>4. Formal Entertaining</td>
<td>2</td>
<td>3</td>
<td>0</td>
<td>4</td>
</tr>
</tbody>
</table>

Clearly those spatial elements whose relative proximal interrelationship is high (3 or 4) should be located in close proximity to each other. If in the above example we take the independent variable "kitchen" as a single dimension of our problem, then those dependent variables with a high (4) and similar relationship to "kitchen"
should be located near to one another, e.g., "formal living."

Also, if we take a somewhat abstract example,

<table>
<thead>
<tr>
<th></th>
<th>A</th>
<th>B</th>
<th>C</th>
<th>D</th>
<th>E</th>
<th>F</th>
<th>..</th>
<th>N</th>
</tr>
</thead>
<tbody>
<tr>
<td>I</td>
<td>4</td>
<td>2</td>
<td>1</td>
<td>4</td>
<td>4</td>
<td>2</td>
<td></td>
<td></td>
</tr>
<tr>
<td>J</td>
<td>4</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>3</td>
<td></td>
<td></td>
</tr>
<tr>
<td>K</td>
<td>4</td>
<td>1</td>
<td>2</td>
<td>4</td>
<td>4</td>
<td>2</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

it is clear that on the basis of the independent variable dimensions expressed here, that I should be placed nearer to K than to J although in the A-th dimension they are all strongly interrelated.

The reason for this is that, given all the independent dimensions of the problem A, B - - - - F, the desired relationship of I to A, B - - - - F is almost exactly the same as that of K to A, B - - - - F, and considerably different from J. This is not at all unreasonable because if two spatial elements functionally depend on the same set of single spatial units in the same way, then (all other things being equal), they shall share similar locations in order to fulfill this functional dependency.

For example, suppose we arrange four spatial units (A, B, C and D) in the following manner satisfying a
given set of proximal contraints:

Suppose a further element is added to the aggregation with the following profile:

Then the aggregation would necessarily appear in the following manner:

Again, suppose that another element is added with a similar profile to that of E:

Then its placement within the group shall necessarily be close to that of E in order to continue to satisfy the proximal requirements specified for the problem.
The point of these illustrations is to demonstrate the complexity and logical nature of the desired connection between the spatial element within the context of the whole problem. It can clearly not be a case where spatial elements reside together on the basis of the necessary and sufficient condition of sharing another spatial element in common. Rather, it is more a situation that two particular spatial elements can be quite unalike in their profiles of inter-element relationships and yet, in terms of all the other profiles and the problem as a whole, be in fact more similar in their proximal dependencies than any other given element in the study.

The converse of this is also equally true. Suppose the profile scores for two elements were identical, e.g.,

<table>
<thead>
<tr>
<th>X</th>
<th>4</th>
<th>4</th>
<th>2</th>
<th>1</th>
</tr>
</thead>
<tbody>
<tr>
<td>Y</td>
<td>4</td>
<td>4</td>
<td>2</td>
<td>1</td>
</tr>
</tbody>
</table>
Then given the theoretical possibility of a comprehensive data set, X and Y would be, for the study, synonymous. The only way they could satisfy the set of proximal relationships would be to reside in the same location. (Obviously, at another level and in an operational sense, this does not mean that X is not different from Y; rather, for the purposes of the functional arrangements as specified by the problem, they are effectively synonymous). This would describe a further constraint for the element string in so far as the spatial elements should differ from one another in at least one dimension of the interaction matrix.

Having modelled the problem in a manner in which what is effectively an empirical taxonomy is desired, and with the aim in mind of producing something akin to an affinity diagram (schematic layout), some desirable characteristics of a technique can now be specified.

It is apparent, given the inherent complexity of the task before us in satisfactorily resolving this set of proximal affinities, that a systematic data handling technique with an extended information processing capability is
required. A particular form of cluster-seeking technique, developed by the author, was found to be appropriate for the solution of this class of problems and in a manner which would appear to have few of the disadvantages of some other approaches to the problem's solution. 3

To facilitate the task, it is necessary to control the aggregation of spatial elements in such a way that the critical relationships between them can be described. In other words, the way in which they converge into groups is important to the problem solution. One possible strategy is a cluster-seeking technique which combines two elements, or an element and a cluster, or two clusters, at a time. The resultant representation in the form of a binary tree, as we shall see, is an important step in the problem's solution.

In discussing clumping techniques Ball makes the following comment.

One application for which such techniques seem particularly valuable... is that of developing taxonomies... The problem is tracing them (éléments) back along an evolutionary tree, combining branches as
clumps become close together.\textsuperscript{5} This is essentially the type of procedure required for analysis of the data related to this problem in the manner formulated.

2. The Distance Coefficient

Before describing the computer program in any detail, some general discussion related to the appropriate procedure for expressing the proximal affinity between the spatial elements must be made. It has been argued within the bounds of this problem that if two spatial elements are alike with respect to their proximal affinities along all the dimensions, then the proximal relationship between them shall be strong. Texts dealing with the subject of numerical taxonomy and cluster-seeking techniques abound with many such measures.\textsuperscript{6} Many are terms of geometric and distance considerations; others are in terms of graph and network linkages.

The basic geometric interpretation of the interaction matrix cell values specifies the location of these points in a multi-dimensional space. The position of these
points (cells) is fixed in this hyperspace by the cell value. In other words, a spatial element is being expressed in terms of a determinable distribution of points in a hyperspace. A three-dimensional case of this clearly shows the efficacy of this model.

\[
\begin{array}{ccc}
A & B & 0 & 3 \\
A & 0 & 3 \\
B & 3 & 0 \\
\end{array}
\quad
\begin{array}{ccc}
A & B & C & 0 & 5 & 3 \\
A & 0 & 5 & 3 \\
B & 5 & 0 & 2 \\
C & 3 & 2 & 0 \\
\end{array}
\]

\[d = \sqrt{(B_A - A_a)^2 + (A_B - B_b)^2} \]

\[AB = \sqrt{(A_A - B_A)^2 + (A_B - B_B)^2 + (A_C - B_C)^2} \]

The development of the "generalized distance statistic" allows the extension of this formulation into multidimensional space. In so doing a measure of proximal affinity (connectedness) between a pair of spatial elements can be determined.
General expression of the distance measure is

\[ D_{\alpha\beta} = \sqrt{\sum_{i=1}^{n} (X_i - X_i')^2} \]

Where \( D \) = the distance measure between elements \( \alpha \) and \( \beta \) and \( X_i\alpha \) and \( X_i\beta \) - the position of \( \alpha \) and \( \beta \) profile points in the \( i \)th dimension. This expression is of particular interest, as it clearly allows distance (location) to be specified figuratively (in terms of an interval scale value system), or quite literally for special types of problems. The general analysis is the same, as both expressions are not mutually dissimilar with respect to the technique.

It is interesting to note that by providing a dimension with a possible geometric (proximal) interpretation for each spatial element that the necessary analysis for the space allocation problem can be performed by a method of direct determination rather than the more trial-and-error oriented optimization procedures which have been previously developed for the solution of this class of problems. Herein lies a significant difference between this approach and others and the point of departure of this
particular formulation of this class of problems. Fixation on a local optimums, which is one of the inherent dangers of the "hill-climbing" versions of many strict optimization approaches to the problem, is consequently avoided.
FOOTNOTES

1. The notion of Aristotelean sets.

2. The notion of Adanson sets... as used in numerical taxonomy.

3. See: Whitehead (1961); Souder and Clark (1964), development of CRAFT and CORELAP, etc... Also, Delon (1969); HIDECS series Manheim and Alexander (1963); PLAN9, McAdams (1969); ALDEP and MATRAN, Miller (1969)... etc.

4. For further discussion and definition see: Sokal and Sneath (1963); Smirnov (1960); Cole (1969).


6. For further discussion see: Sokal and Sneath (1963); Smirnov (1960); Cole (1969); Wirth, Eastbrook and Rogers (1966); Ball (1966).


8. (Refer to footnote number 3.)
V. DESCRIPTION OF COMPUTER PROGRAMS CAPLAN AND PLANDS

CAPLAN and an interactive time-sharing version PLANDS were the computer programs developed in this investigation, on the basis of the foregoing theoretical discussion. They serve as the primary information component by which the problem data structure is analyzed, clusters found, and interrelationships described. In this section it is proposed to specify various aspects of the program(s) in order to better describe their role as aids in the design decision-making process.

1. General Characteristics

The present operational version of CAPLAN was written for the Burroughs B5500 system and FORTRAN G compiler of the Rice Computer Center. Because of the necessary number of large arrays which must be stored, the program requires in the order of 200K bytes. Due to the flexible characteristics of the B5500's core storage and the design of the program, 32K word segments are more than adequate for the necessary computational task. The program uses almost no system
specific functions and hence could be adapted to most
FORTRAN compatible systems. CAPLAN handles up
to 100 spatial elements.

2. The Cluster-finding Algorithm

The central or "driving" algorithm of the computer
program makes use of a procedure which can be best
described as "clumping" or "clustering by single link-
age."² It commences by computing all possible pair-
wise relationships or similarity coefficients between
the spatial elements using a version of the general
Euclidean distance formula.³ This computation pro-
duces the initial symmetric relationship matrix. This
matrix (or string of coefficients) is then searched (sorted),
the lowest coefficient value is isolated, and the two
spatial elements concerned are combined to form the
first cluster.³ Their coefficient is then moved to its
diagonal element in the matrix. The relationship co-
efficients of the other spatial elements, as they repre-
sent potential combination with this cluster,⁴ must be
modified. This is performed for each cell after the
method of Ward⁵ by the following formula:
Typical computation for the $i^k$th cell or the Matrix $M$.

\[
Dik = (Dik (Ni + Nk) + Djk(Nj + Nk) + \ldots \\
Din (Ni + Nm) - Dii(Ni) + \ldots + Dnm(Nm))/ \\
(Ni + \ldots + Nm)
\]

$Dik$ represents the potential coefficient for combining the $i$th and $k$th spatial elements; and $N$ represents the number of elements per cluster, and let us say that $Djk$ represents the relational coefficient forming the first cluster.

This computation is performed for each cell in the matrix. After this reduction the matrix is now in the form to allow the selection of the second pairing, again on the basis of the lowest relational coefficient. In order to maintain the presence of those elements that have already formed clusters, this decision is based on determining that cell which has the smallest value when its corresponding diagonal cells have been subtracted, i.e.,

\[
d = Dij - Dii - Djj
\]

Again, the matrix is modified and the procedure repeated until all clusters have been formed revealing the universal set which contains all the spatial elements. Given that there are $N$ spatial elements in the study, this
will amount to N-1 iterations of the algorithm and the formation of N-1 "Threshold" levels in the study. (Thresholds numbered from 101-100+(n-1)l) in output). This procedure with the use of raw data produces satisfactory "natural clusters," except when the element profiles are extremely weakly defined (lacking any clear modality) and the dimensions scaled with very small variances.

The foregoing discussion has emphasized the fact that the dimensions (spatial elements), should be constructed in such a manner as to differentiate as much as possible between spatial elements.

In program CAPLAN the majority of this computation is found in the main calling routine MAIN.

3. Structural Characteristics of the Clusters

In addition to the specification of the data content of a cluster, consisting of a list of spatial elements, other important characteristics are investigated by program CAPLAN. They may be enumerated in the following manner:
a. Specification of the element describing or strongly influencing the specific cluster formation:

All the elements held in common, by value, are computed for the spatial elements forming a cluster. Those with high values (0 = no interaction) can be seen to strongly influence the cluster formation. CAPLAN allows the user to fix the desired value strength on the basis of the original input interval scale by use of a control parameter, in order to identify these attributes. (All profiles are displayed with clusters to facilitate identification of possible decision changes for future input data).

b. Percentage Cluster Connectivity:

This measure is useful in interpreting the overall strength of the relationships between the members of a given cluster. It also gives some measure of the data retention within a cluster. Its computation in CAPLAN is carried out independently of the selection of clusters and represents the within cluster connectedness. The
computation is performed in the following manner. A coefficient of dissimilarity $Cr$ for a study can be computed from the following empirically derived formula,

$$Cr = \frac{(N-(N/2)) (N/2) ((MAX-MIN)^2)}{\text{where } N = \text{number of elements in the cluster,}}
\text{MAX = maximum value of the scoring interval scale on the attribute dimensions, and MIN = minimum value.}}$$

Hence, percentage connectivity for a cluster is given by,

$$P = \frac{Cr}{\sum_{i=1}^{n} \sum_{j=1}^{n} (X_{ij} - X_{i+1,j})^2 / C_k} \times 100 \text{%}$$

where $Cr = \text{coefficient of dissimilarity for the study and } X_{ij} \ldots \text{etc.} = \text{cell entry for the ith element and the jth element dimension.}$

It should be noted that this will never be 100% as the very inclusion of a spatial element in the study implies some degree of connectedness.

c. Specification of the state of the study at a given point in the analysis in terms of the number of
clusters so far formed and the number of elements which remain unclustered.

d.. Data Retention:
A useful piece of information concerns the relative difference between the percentage cluster connectivity at one level and that at another. This allows some monitoring capability of the amount and specific nature of the data retention in a particular branch structure or sub-graph of the total study. It also acts as a measure of cluster isolation and an aid in identifying "natural clusters" within the study. CAPLAN specifies this comparatively for both links forming a new cluster.

The measurement is given by:

data loss = Pi - Pj
link 1

data loss = Pi - Pk
link 2

where Pi = % connectivity of new cluster formed and Pj and Pk = % connectivity of previous clusters (elements) combining to form the ith cluster.

S = Pi = min represents the % "data loss" by
clustering at the given level.

e. Percentage Simplicity Gained:

As a measure of the simplification of the study and the reduction of possible information units, this measure is given by:

\[ r = \frac{N - (N_c + N_e)}{(N - 1)} \times 100 \%
\]

where \( N \) = number of spatial elements in study,
\( N_c \) = number of clusters so far formed,
\( N_e \) = number of spatial elements remaining (unclustered).

f. Cost Benefit Ratio:

This allows statistical monitoring of the important "trade-off" previously mentioned, between the simplicity gained in the study and information (element data) lost by the clustering of two units.

Cost Benefit Ratio = \( \frac{r}{s} \)

(This measure also can be interpreted to isolate that section of the analysis which potentially has the greatest amount of useful information in the context of the design problem).
The above measures describing the structure of clusters formed and the state of study analysis generally are illustrated in the worked examples. They form one of the more powerful interpretive elements of the analysis output and the structure of the problem space generally. In addition they represent some of the departures from other like generic types of applications.

4. Input Modes

In the course of the discussion some of the possible forms of the initial problem information were suggested and elaborated upon. Emphasis should be placed upon the desirability of interfacing in a compatible and natural way (languages), the initial data set and the input mode of the computer program. This has been accomplished in CAPLAN by subroutines SYMM and STATE.

a. SYMM:

This subroutine allows the input of information of a form directly descriptive of the spatial element vs. spatial element interaction. To minimise the effort involved in defining the input
matrix, and in view of the symmetric nature of the matrix, only one half of the matrix plus the principal diagonal need be entered.

b. STATE:

A further aid in the easy (for the designer) establishment of the components of the interaction matrix involves the direct coding of an initial list of simple verbal problem statements. Consider for a moment the structure of one of these statements: e.g., "The kitchen must be next to the dining room". (SNP+VP) This is descriptive of two spatial elements involved in a study and the nature of (strength) of the interaction between them. Hence, they can be coded in the following simple manner:

Let, element "kitchen" spatial element No. 3.

element "dining room" spatial element No. 6.

and, interaction type "must be next to" scale score 4.

Hence, the sentence simply becomes (3, 4, 6).
This subroutine (STATE) allows such a coded list to be made in any order and simply sorts (SORT) the component strings in an appropriate manner in order to generate the symmetric interaction matrix previously described. This input mode clearly overcomes some of the difficulties involved in pedantically, and to some, irrelevantly filling in each cell of the matrix. The designer is in a position to freely specify only those relationships he thinks are most critical to the study and the logic of the design situation. This aspect of matrix input is, to the author's knowledge, an innovation on the part of this technique comparative to other generically similar programs involving input matrices of a similar type.

Also included in the data input mode characteristics is the option of specifying the square foot area of the spatial elements, this is aggregated to give the total area per cluster.

Subroutine STAND allows the initial or raw interaction matrix to be prestandardized before clustering begins. This option has been incorporated in the event that con-
siderable distortion is suspected in the matrix due to an unbalanced set of small and large variances along the element dimensions. This could occur in a large matrix in which the previously described systematic data generation technique was not adhered to. In other words the judgemental scale initiating the cell scoring varies across the matrix in an erratic and non-systematic manner. Needless to say, this practice and the use of the standardization option should be avoided where possible.

Program CAPLAN with subroutine LIST provides a means whereby particular element dimensions can be differentially weighted by use of a specified vector string of constants. The primary purpose is to allow the designer to study the effects of a particular group of attributes on the study without necessarily precluding the influence of others. This is perhaps an artificial imposition, but it can be seen to be potentially advantageous in the use of the technique for purposes of simulating an existing aggregation of spaces and manipulating the outcome of this approach. LIST also incorpo-
rates a raw summation and rank ordering procedure which allows spatial element profiles to be grossly compared as to their relative density or strength of their influence on the study.

5. Control Parameters

The subroutine KCDS allows the input of the relevant information concerning the title and date, etc., of the study and the selecting and description of the various options available to the designer in using CAPLAN. The program architecture is such that subroutine MAIN incorporates the branch points to the available options by initially calling the control parameter vector string from KCDS and assigning these values (usually 1 or 0) to the appropriate Boolean switch statements. It also provides the means by which two or more studies can be analyzed in the same run. Apart from the options already mentioned, KCDS provides a means by which the detailed output can be suppressed in a flexible manner, in order to avoid unnecessary voluminous output for large studies. An upper and lower limit of threshold levels can be specified in order to isolate the desired
section of detailed output. (Generally the upper middle section yields the most pertinent results.) Also are incorporated a set of error checks which facilitate the user in determining gross errors in this input data set, e.g., number of elements.

6. Command Language

The interactive time-sharing program PLANDS is essentially the same as CAPLAN with the addition of a command and edit language allowing the designer to freely manipulate the input, analysis and output of the program. At the time of writing this application is still undergoing development but this section is incorporated in the discussion as it is the logical extension of the batch-mode application and allows for a higher-level communication between the designer and the computer. The basic features of the system is the application of a controlling FORTRAN command and edit language and the necessary subroutines. This incorporates some twenty-five commands in English by which the designer can add or delete elements from the study and make single decisions and evoke the desired input
and output modes. The basic constraint in the design of this aspect of the system was to provide the designer with flexible control of the analysis at the level of detail required without rigidly constraining him to a single mode of operation or a pedantic time-consuming process for inputting the necessary information.

7. Output Mode and Representation of the Analysis

The output from CAPLAN is in both verbal and graphical form. It has been designed to be comprehensible and useful to a non-specialist with just a basic understanding of the nature of the aggregation strategy.

The verbal section of the output mode begins by specifying the threshold number, allowing for easy cross reference to the graphic display for the purpose of obtaining more information about a particular cluster represented in this graph. Each cluster formed is then listed as to its spatial element membership, their profile of decisions, the total area (square feet) of the cluster formed, and an indication of the "reason" for this cluster formation by way of a verbal statement of these elements most strongly effecting the cluster and
a statement of the number of similar elements held in common by the cluster members. A detailed description of the structural characteristics of the cluster formed at the particular level includes the percentage connectivity, the percentage simplicity gained by clustering at that level, and the amount of information (relationship information concerning commonly held elements) lost by clustering at that level. In this way the internal connectedness of the clusters, their degree of isolation within the study, and their meaningfulness with respect to the problem can be accurately monitored. The point is that the more tightly connected or homogeneous the clusters are while occurring at a intermediary or higher level in the hierarchy (i.e., not primitives,) the greater the legitimacy for treating them as a single unit. The advantage is that with a considerable reduction in the number of "discrete" information units and a minimum loss of information the task of effecting an appropriate configuration of spatial units is greatly simplified.

The graphical output mode takes the form of a hier-
archical binary tree, beginning with the spatial elements as system primitives and ending with the universal set of all elements. The vertical axis of the tree is measured in degree of percentage cluster connectivity, and the clusters are placed in this dimension in accordance with this rule. The hierarchical tree form is a standard representation for this type of technique. The binary nature of the tree allows the simplest representation of the interrelationships involved, and allows for a more conventional spatial representation to be easily generated, (affinity diagram). The tree is produced by a "tree walking" algorithm incorporated in the main framework of the program.

Upon inspection of the tree and its branch structure, relatively independent spatial aggregations can be quickly identified. Large gaps in the hierarchy indicate that more information has been lost in forming the higher cluster. The verbal output can be easily consulted, with reference to that cluster, to determine something about the nature of this loss. An important interpretive characteristic can be attached to a spatial
element that remains unclustered through most of the study (for example, Student Housing shown in the Master Plan case study). Usually the number of nodes between its cluster formation and the rest of the study elements is at a minimum. This suggests that it is singularly independent or that it pervades most of the study. Taking the latter inclusive sets interpretation, this would seem to suggest that the spatial element becomes a locus for the complex, or if this did not meet with the general formal aims of the design, that it be a decentralized function. Hence, although the technique deals explicitly with the combination of spatial elements into exclusive sets, the output can, on occasions, be interpreted inclusively. The technique then clearly becomes an aid to design concept development, one of the functions of the programming phase.
1. CAPLAN an acronym standing for "Cluster Finding Algorithm for use in Planning."


3. Equation of the form \( D = \sum_{i=1}^{n} \frac{(X_{i\alpha} - X_{i\beta})^2}{2} \).

where \( X_{i\alpha} \) and \( X_{i\beta} \) = cells \( \alpha \) and \( \beta \) of the \( i \)th attribute dimension.

(Division by 2 - as there are two elements in the potential cluster).

4. In the event that there are two pairs of the same threshold value only one is chosen to form the first cluster. They shall be sufficiently independent pairs not to distort the analysis by this seemingly arbitrary action. The independent calculation of cluster statistics assumes no loss of accuracy in taking this action.

5. The Adansonian concept of acts.

6. In this respect the algorithm draws on the work of Ward (1963), p. 236. Ward in this paper proposed a method for numerical classification which was designed to optimize the error sum of squares objective function in a hierarchical fusion process.

7. For further discussion see Carmichael, George and Julius (1967), pp. 144-145.


VI. RESULTS

1. A Partial Proof

Before describing in detail a set of case study examples showing the application of this technique, it is proposed to offer some significant experimental validation of the technique and program CAPLAN.

Suppose we consider the following problem examples.

The expected statement of the result can be expressed in the following manner.

\[
\begin{array}{c}
\text{TWO} \\
\text{THREE} \\
\text{ONE} \\
\text{FOUR} \\
\text{FIVE}
\end{array}
\]

(Graphic representation of desired result)

<table>
<thead>
<tr>
<th></th>
<th>ONE</th>
<th>TWO</th>
<th>THREE</th>
<th>FOUR</th>
<th>FIVE</th>
</tr>
</thead>
<tbody>
<tr>
<td>ONE</td>
<td>0 2 2 3 9</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>TWO</td>
<td>2 0 0 1 7</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>THREE</td>
<td>2 0 0 1 7</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>FOUR</td>
<td>3 1 1 0 6</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>FIVE</td>
<td>9 7 7 6 0</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

(NB: This problem contains no contradictions involving "trade-offs" where an element is desired to be two places at one time and hence can simply be graphically represented.)

We should expect the analysis of this problem by program
CAPLAN to arrange the elements ONE through FIVE in this predetermined order on the basis of the input distance matrix shown above. In fact, that is what occurs upon analysis. The arrangement of the primitive elements at the base of the hierarchical binary tree indicate the order expressed in the above diagram. (It should be noted that the technique is designed to be descriptive of the data, and hence the question of the quantitative distance between the elements in the output is excluded. However, it is a simple matter, having the pre-order to arrange the elements in the appropriate manner.) The solution to problems of this sort are obtained directly and with an economy of means, excluding any iterative "hill climbing" (trial & error) procedures involved in many other applications. This specific problem was one of many submitted with the same direct determination of a result.

As we shall see in the following case studies, CAPLAN generates the necessary interpretive and supportive analysis for solutions to problems to be readily extended to two dimensions and hence the schematic planar layout
One-dimensional Case Study Analysis.
of spaces (the problem as stated).

The above problem is a classical test of the efficacy of programs of this type. To the knowledge of the author, this is perhaps the only application for a problem of this type which determines a result in a direct manner. Needless to say, the relative power of the technique is increased by the complexity of the problem presented. Complexity in this case is more often than not a function of the number of spatial elements included in the study.

2. Case Study No. 1 - Domestic Suburban Dwelling

The purpose of the following case study example is to describe in some detail the application of this technique in the solution of a specific space allocation problem. The example chosen (hypothetical data) concerns a typical domestic suburban dwelling. It is an example that was deliberately chosen, not because of any inherent complexity (number of units) but rather as an example whose solution is within the immediate experience of most readers. It should be stressed that the intention of this example is not so much to develop a prototype layout of a single-family dwelling as to
**DESCRIPTION OF INITIAL INPUT DATA**

**ELEMENT LISTING - DESCRIPTION AND PROFILE**

**SPATIAL ELEMENT DESCRIPTION MATRIX PROFILE**

<table>
<thead>
<tr>
<th>Element</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 = KITCHEN</td>
<td></td>
</tr>
<tr>
<td>2 = DKGST/EATING</td>
<td></td>
</tr>
<tr>
<td>3 = CARPORT</td>
<td></td>
</tr>
<tr>
<td>4 = FORMAL ENTRY</td>
<td></td>
</tr>
<tr>
<td>5 = INFORMAL ENTRY</td>
<td></td>
</tr>
<tr>
<td>6 = OUTDR LATING</td>
<td></td>
</tr>
<tr>
<td>7 = FORMAL LIVING</td>
<td></td>
</tr>
<tr>
<td>8 = LIVING AREA</td>
<td></td>
</tr>
<tr>
<td>9 = DINING RM*</td>
<td></td>
</tr>
<tr>
<td>10 = BATHROOM</td>
<td></td>
</tr>
<tr>
<td>11 = LAUNDRY</td>
<td></td>
</tr>
<tr>
<td>12 = GUEST BDRM*</td>
<td></td>
</tr>
<tr>
<td>13 = PARENT BDRM*</td>
<td></td>
</tr>
<tr>
<td>14 = CHILD BDRM* (2ND)*</td>
<td></td>
</tr>
<tr>
<td>15 = POWDER RM*</td>
<td></td>
</tr>
<tr>
<td>16 = STUDY</td>
<td></td>
</tr>
<tr>
<td>17 = TV RM*</td>
<td></td>
</tr>
</tbody>
</table>

*Case Study No. 1 - Matrix.*
demonstrate the efficacy of a particular approach to the problem. Then let us assume that the logic involved in the problem is that which might be found in a typical case of its actual application in Suburban U.S. Again, for the purpose of describing a technique rather than a unique prototype solution, the nomenclature defining the spatial element descriptions has the generally accepted meaning.

a. General Characteristics of the Data

The input data for this case study analysis (as shown) consists of seventeen spatial elements descriptions and proximal affinity interaction matrix with the scoring system previously discussed in the body of this paper (i.e., ++ = 4, + = 3, 0 = 2, - = 1, -- = 0).

b. Problem Analysis

Selected aspects of the problem analysis are shown here, including the binary "hierarchical graph of associative relationships" graphically describing the pattern of cluster formation and
Case Study No. 1 - Hierarchical Graph of Associative Relationships.
more detailed cluster descriptions at certain thresholds, including a breakdown of various clusters structural statistics.

c. Interpretation of Results

The conventions to be adhered to in interpreting the results from CAPLAN have been outlined previously. To reiterate, possibly the most central issue is the degree to which a particular cluster is connected and consequently proximally how strongly related is its element membership. This can be generally given by the length of the linkage line (shorter = more connected) and the per cent connectivity on the ordinate axis of the graph. The significance of the other general structural statistics shall be referred to in the course of the study.

(i) The first step involves discriminating the presence of any tightly connected, relatively independent clusters in the middle range of the hierarchy, for example, those contained in subsets 110, 111, and
Case Study No. 1 - Graph Showing Zones.
112, and to a much lesser extent, 113 (c. 95% of 89% connectivity - see diagram).

It is interesting to note that all the elements contained in subset 112 (bathroom, bedrooms, guest bedroom) constitute a specific general function of the building type, namely the sleeping unit. Similarly, the other subsets describe formal entertainment (111); family outdoor living and service (110); and the "loosest" of the four, the informal family indoor living area (113). This is not necessarily a coincidental result. In most typical dwellings of this type, these general physical units correspond to the set of general and separable activities. (The interaction decision-making, in keeping with the intended logic of the situation, reflected this aspect of domestic activity cycles.) Furthermore, at this level, there can be seen to be a split (see diagram) among those spaces concerned
based on the more private interfamily activities and the more public face of formal entertainment. The relative weakness of subset 113 suggests that it is less of a physical unit than the other three, and as such it tends to pervade areas of the rest of the study more closely associated with it (110 and 112) (i.e., the number of nodes of separation is at a minimum).

It should also be noted that the built physical enclosure represented by subset 110 is minimal (really only the laundry and the carport).

In our first spatial organization of the problem, the following zonal relationship diagram might be generated (including relative areas).

1. Informal family living
2. Outdoor/service
3. Sleeping Unit
4. Formal entertainment

Analysis of each zone can then be proceeded with, in a similar manner. For
example, consider the sleeping unit shown by Zone 3.

Bathroom
P. Bedroom
C. Bedroom
G. Bedroom

SUBGRAPH 3.

Subset 104 is a tightly connected cluster and its relationship, as represented, to 109 is of interpretive significance as it illustrates a powerful aspect of the present graphic representation.

Clearly, on the basis of the analysis information, the parent bedroom (13) and child bedrooms are required to be in close proximity. Spatial Element 10 can be seen to pervade the conjunctive subset of these
two elements and hence have a (functional) proximal dependency upon them or vica versa (sequence I through III). Hence, the following planning consideration may be made:

Just as clearly the guest bedroom should be located nearer the bathroom than necessarily next to the other bedrooms as it only joins the cluster 109 to form 112 after the addition of the bathroom unit.

Clearly, element 12 belongs to this cluster, as the information loss with respect to link 1 (see output between 109 and 112) is considerably less than that between 112 and 115, i.e., between the sleeping zone 3 and the rest of the study.
Zone 4: (The formal entertainment area)

SUBGRAPH 4.
Consider the most tightly connected section of this subgraph.

Reference to the detailed description of cluster or subset 106 yields the fact that the element most strongly defining the cluster is element 7 (formal living).

Hence, the relationship can be represented by the partial affinity diagram.

Examination of the decision sequence
(profile) associated with cluster 107 shows that element 16 is loosely associated with this cluster (however the same argument holds for element 12 above); the strength of the relationship is directed towards 7 and 9:

Also, by the same procedure, element 4, formal entry, which pervades the cluster (111)
can be seen from the decision sequences to be most closely associated with element 7, the formal living area. Hence, the subgraph reveals the following "affinity diagram" representation:
Zone 2: (Outdoor family living and service area)

Carport
Laundry
Informal Entry
Outdoor Eating

SUBGRAPH 2.

By a similar process of interpretation (cluster 110) element 5 (informal entry) is seen to form the fulcrum of the sub-graph 2:

II.

(There is clearly something intuitively correct about this arrangement as the "informal entry" (5) is the means by which traffic circulation external elements 3, 11, and 6 are physically associated with the rest of the study.)
As we have said, 113 in comparison to the other zone subsets (2, 3, and 4) is more loosely connected. Clearly 105 and particularly 102 are closely interrelated spatial units. Information contained in cluster 113 (their conjunctive subset) allows interpretation to be made about the nature of their interrelationships. The decision sequence reveals that element 2 (breakfast/eating) is "central" to this cluster.

(ii) Interrelationships Between Zones

So far we have described the general decomposition of the hierarchical graph
into zones and resolved the spatial relationships of each zone in the form of 2-D affinity diagrams. However, the task still remains to study the interrelationships between these diagrams to fully describe the study in the desired manner.

Proceeding in order, consider cluster 114 (the conjunction of 113 and 110). The detailed cluster information at this level reveals no elements held in common. However, inspection of the decision sequence (profile) reveals that the closest connection between 113 and 110 occurs through elements 5 and 1 (informal entry and kitchen). Again this would seem intuitively the correct solution. (113 + 110)
A similar consideration of subset 115 (conjunction of 112 and 114) reveals that the information lost by 112 in combining in this manner is considerable (link No. 2) and that 112 still remains relatively isolated from the remainder of the study. Inspection of the decision sequences indicates a tentative relationship between 14 and 17. (113 + 110 + 112)

However, in one respect the decision sequence at level 116 (conjunction of 115 and 111), although the cluster is relatively unconnected, indicates a strong proximal linkage between elements 9 and element 1 (of 113). This would seem at first to be somewhat contradictory as 111 and 113 are on opposite sides of the hierarchical
graph. However, closer scrutiny shows that in consideration of the relative levels of the hierarchy shown here cluster 113 is some 33% less connected than the least connected of the other three zone subgraphs, and, in accordance with the prescribed logic of the interpretive process, to have a pervading effect on the entire study.

In addition, within this subgraph element 1 is relatively (101, 102, 103, 104) unconnected. This is due not so much to its direct relationship with 2 (value 4), but to the general "pull" of the rest of the study. (Inspection of the profile of 1 shows an unconfined (to one area) string of strong interactive decisions.) In fact, at the outset of the analysis during step 1 the subgraph associated with cluster 113 could have been seen to be the central
overall spatial unit in the study by the manner just described.

I. Family living area
II. Outdoor service/living
III. Sleeping area
IV. Formal entertainment

AFFINITY DIAGRAM.

This represents a graphical representation of the case study spatial relationships in the form of an affinity diagram, the required form of solution for the problem as stated. It clearly reveals the consequences of the decisions made with respect to proximal relationship and the space allocation problem. As such, it acts as a simulation of the designer's value structure in this aspect of the problem. In pursuing this case study, the object has not been to place emphasis on the decisions made in terms of developing some proto-
type housetype, but rather on the process involved in resolving this particular space allocation problem. (In all likelihood there could be some argument with the nature of the decisions made; the author is perhaps revealing some biases. However, it would seem that the result is fairly indicative of many typical examples of this typology - the intent.) In fact, the technique is intended to be used in an iterative manner in order to better resolve decision-making discrepancies. In this study it would be useful to review the decisions associated with the kitchen, the dining room, and the sleeping unit. The object is to focus attention on the precision of these decision profiles to perhaps gain better definition (from that revealed in this analysis) of the interrelationships between the units involved.
The technique clearly allows the problem to be systematically structured and broken down in a manner which allows a solution to be directly generated. The following is a summary of the steps involved and the prescribed rules for the interpretation of the hierarchical graph in the manner previously outlined in the example.

(i) Decomposition of the graph into well inter-related subgraphs at an intermediary level of the hierarchy.

(ii) Inspection of the relative levels of the representative subgraph subsets to discern any pervading spatial units of the study, e.g., 113 Centralization of their location.

(iii) Examination of subgraphs individually to form affinity diagrams. This involves examination of the successive conjunctive subsets with respect to the following information: elements strongly describing
the study, per cent connectivity of cluster (strength of relationship), examination of decision sequences for strong dependencies.

(iv) Combination of "affinity diagrams" into an overall schemata. Examination of conjunctive subsets in the manner of point iii.

(v) General scrutiny of study outcome. Further detailed inspection of specific decision sequences thought to be in error where necessary, including errors in impact, etc... 

(vi) A further iteration of the process where necessary.

A statistical check of the interaction matrix against the resolved affinity diagram reveals a 100% (10/10) resolution of the interactions of strength 4. On the other end of the scale (value 0) is also a 100% (5/5) resolution. A spot check also reveals that the values of 1 and 3 have in most cases been resolved. The result would seem to be representative of the initial information.
Further information content can be added to this "affinity diagram" by the addition of individual area specifications for the units involved. Program CAPLAN supplies this information directly for each aggregation of spaces.

3. Case Study No. 2 - Hospital Complex Master Plan

The following case study example of the technique was performed on behalf of the architectural firm of Caudill Rowlett and Scott (CRS) during the programming phase of the design activity associated with the Texas Tech Medical Complex.

a. General Data Characteristics

The data set consisted of twelve spatial elements. Each corresponded to a large medical school facility as can be seen from their descriptions. This is a case in which the technique was used at a level of generality above that of a single elemental activity unit prescribed in a previous section. The problem was to investigate the spatial arrangement of the master-plan units on the basis of the relationship specified. An
### Spatial Element Description Matrix Profile

<table>
<thead>
<tr>
<th>Element Description</th>
<th>Matrix Profile</th>
</tr>
</thead>
<tbody>
<tr>
<td>School of Pharmacy</td>
<td>411212343332</td>
</tr>
<tr>
<td>School of Allied Health</td>
<td>142303333333</td>
</tr>
<tr>
<td>School of Vet Med</td>
<td>124213232332</td>
</tr>
<tr>
<td>School of Dentistry</td>
<td>232412222323</td>
</tr>
<tr>
<td>Student Housing</td>
<td>101143201223</td>
</tr>
<tr>
<td>Health + Activities Center</td>
<td>233234330122</td>
</tr>
<tr>
<td>Medical School</td>
<td>332223423312</td>
</tr>
<tr>
<td>Teaching Hospitals</td>
<td>433203241311</td>
</tr>
<tr>
<td>Ambulant Pat Housing</td>
<td>332310314231</td>
</tr>
<tr>
<td>Mental Health Center</td>
<td>333221332422</td>
</tr>
<tr>
<td>Rehab Center</td>
<td>333322113241</td>
</tr>
<tr>
<td>Student Health</td>
<td>232332211211</td>
</tr>
</tbody>
</table>

Case Study No. 2 - Matrix.
Case Study No. 2 - Hierarchical Graph of Associative Relationships.
interval scale (1 - 4) and the interaction matrix were supplied by the programming staff of CRS-I with the following interpretation:

4 - Must be next to.
3 - Should be adjacent.
2 - Weak affinity.
1 - No association

b. Results

The program CAPLAN produced the following results, expressed in the form of a hierarchical binary tree and some sections of the detailed description of the cluster characteristics from the accompanying printout.

The specific intention behind the inclusion of this case study is to demonstrate the way in which alternative design concepts can be generated. Consider the hierarchical graph of associative relationships, particularly clusters 106 and 108.
Both clusters are relatively unconnected, indicating that their spatial relationship with other general areas of the study are strong. In terms of the possible design concepts involved in laying out this plan, an exclusive sets interpretation would require "student housing" and "teaching hospitals" to be centrally located as the focus of the complex.

However, an inclusive sets interpretation would suggest that the appropriate teaching hospitals and the associated student housing be treated as
decentralized spatial units.

DECENTRALIZED PLAN.

The solution actually followed this form to some extent.

4. Discussion

The application of this technique allows a space allocation problem of the type described to be systematically and comprehensively specified within the limits of the available information.

The selection of spatial elements and the making of interactive decisions is a highly subjective process. Different people concerned with the same problem are apt to have contradictory goals and requirements. The logic of a given situation can be a variable thing. The technique allows the consequences or results of a specific designer's assumptions, goals, and decisions to be made manifest. The technique becomes effectively a simulation device allowing the structured communication of the designer's decision-making biases.
As a consequence, it should be used iteratively as the designer can then explore different expressions of his decision-making and their effects upon the spatial relationships concerned. Having pursued this activity to a satisfactory conclusion, the designer then has a clear and useful document from which to continue his design activity and a standard against which to evaluate subsequent design development.

The technique would seem to yield the necessary analysis for a valid and quite direct solution to the space allocation problem. The results have been encouraging. The process does not have the disadvantage of some methods of solution by being largely a deterministic technique unsuceptable to the "local optimum" pitfalls likely in "hill climbing" procedures and the disconcerting possible phenomena of producing two different solutions to the same problem.

The case study examples shown here have been deliberately chosen to give a concise description of the technique. As has been suggested, perhaps the main practical value of the technique is in the development
of a prototype solution to a design allocation problem
from a fundamental description of the required infor-
mation and where necessary precedent does not exist
or is desired to be avoided. In any event, in a prob-
lem incorporating a large number of spatial elements,
it provides an easy and systematic approach to the
problem (as was the case in one of the CRS problems).

The analysis of the problem is descriptive of the infor-
mation and solution rather than attempting to describe
it in a precise geometrical form.

Given the stage of the design activity in which the prob-
lem is being at least initially resolved, this approach
would seem to have more inherent validity in the general
case to the pedantic dimensional and real distance speci-
fication of spatial elements and relationships. There is
usually not sufficient information for this to be done
with any real confidence. The problem of "fit and
misfit" would raise questions as to whether 100' distance
between two elements was really 100' or perhaps 70'
and so on.
One possible application that comes to mind after this development activity is the study of existing facilities. By analyzing these facilities in some detail in terms of their spatial elements and locational interrelationships, a comparable model of the facility can be established. At this point the designer has a detailed and useful descriptive model of the facility explaining much of the locational logic inherent in the design. This model is then suitable for further simulation, allowing the study of the consequences of different assumptions, goals, and decisions associated with an alternative logic or design strategy.
VII. SUMMARY

1. Conclusion

On the basis of this investigation, the general conclusion that can be drawn is that the technique outlined effectively performs the necessary analysis necessary for a direct solution to a space allocation problem to be found. It has few, if any, of the disadvantages of other techniques and is in a form that is readily usable, in both input and output modes, by a non-specialist designer. It can experimentally be shown to yield the optimum solution to the one-dimensional problem type, an example of which was outlined. In addition, a check of the case study example interaction matrix in comparison with the solution clearly shows that all the critical relationships had been appropriately dealt with. Further support is the computational capability of the technique. Probably its most powerful practical application is in the analysis of large studies and in the premeditated development of prototype space allocation solutions from fundamental information sources. It offers a means by which a designer's goals and
assumptions about a problem can be revealed and simulated. In addition, it potentially provides a powerful analytical and simulation device by which the inherent spatial logic of existing facilities can be described and analyzed.

2. Future Work

General development work for investigating alternative methods for the solution of the space allocation problem should be undertaken. The satisfactory direct generation of a solution in completely spatially representative form has not been realized. Perhaps the approach to this problem in that manner is open to question. Stepwise solutions are probably more or less as effective. Specific to this investigation, further development should be undertaken in producing effective notations systems for interpreting the results. In addition, the use of this technique as a heuristic search strategy working with real distance data would enable the spatial elements to be pre-ordered by their interrelationships and a finite and feasible number of linear program objective functions to be specified and the optimum solu-
tion (1-dimension) to be directly solved including the factor of real distance. This could then be geometrically extended to the 2-dimensional case by manipulating the distances involved in a generate-and-test fashion.
ANALYSIS OF STUDY STRUCTURE - ELEMENTS

NEW CLUSTER

11 CLUSTERS REMAIN AFTER CLUSTERING CLUSTER 7 AND ELEMENT 0 TO FORM.

PRESENT STRUCTURE OF STUDY

CLUSTER 7 NUMBER OF ELEMENTS = 3

ELEMENT LISTING

Table: 7 FORMAL LIVING 2, 2, 1, 3, 0, 0, 4, 4, 1, 4, 2, 2, 1, 4, 3, 2
9 DINING RM. 4, 3, 1, 1, 1, 1, 2, 2, 1, 0, 1, 1, 1, 3, 2
14 POOL RM. 2, 2, 1, 3, 1, 1, 2, 3, 3, 1, 1, 2, 2, 1, 4, 2

TOTAL AREA OF CLUSTER = 370.00 SQ FT.

TOTAL NUMBER OF SIMILAR ATTRIBUTES = .4

STRAIGHTER STATISTICS OF CLUSTER AT THIS LEVEL

PERCENTAGE CONNECTIVITY OF CLUSTER = 93.

NUMBER OF CLUSTERS = 5.

NUMBER OF ELEMENTS REMAINING = 6.

PERCENTAGE SIMPLICITY GAINED BY CLUSTERING AT THIS LEVEL = 37.5000 PERCENT

PERCENT DATA LOSS LINK 1 FROM LOWER LEVEL = 3.

PERCENT DATA LOSS LINK 2 FROM LOWER LEVEL = 7.

PERCENTAGE DATA LOSS BY CLUSTERING AT THIS LEVEL = 6.616 PERCENT

COST BENEFIT RATIO = 5.667

Case Study No. 1.
ANALYSIS OF STUDY STRUCTURE ELEMENTS THRESHOLD NUM 109

MOTHER CLUSTER

8 CLUSTERS REMAIN AFTER CLUSTERING ELEMENT 10 AND CLUSTER 13 TO FORM CLUSTER NUMBER 10 OR 109

PRESENT STRUCTURE OF STUDY

CLUSTER 10 NUMBER OF ELEMENTS = 3

ELEMENT LISTING MATRIX PHUFILL

<table>
<thead>
<tr>
<th>ELEMENT</th>
<th>MATRIX PHUFILL</th>
</tr>
</thead>
<tbody>
<tr>
<td>10 BATHROOM</td>
<td>2,2,1,0,2,2,0,1,1,4,2,4,4,1,1,2.</td>
</tr>
<tr>
<td>13 PARENT BURN</td>
<td>1,1,1,1,1,1,2,1,4,1,2,4,4,2,3,2.</td>
</tr>
<tr>
<td>14 CHILD BURN (2ND)</td>
<td>1,1,1,1,2,2,1,3,1,4,1,2,4,4,1,1,3.</td>
</tr>
</tbody>
</table>

TOTAL AREA OF CLUSTER = 4000.00 SQ FT.

5 ELEMENT STRONGLY DEFINING CLUSTER 10 BATHROOM

5 ELEMENT STRONGLY DEFINING CLUSTER 13 PARENT BURN.

5 ELEMENT STRONGLY DEFINING CLUSTER 14 CHILD BURN (2ND).

TOTAL NUMBER OF SIMILAR ELEMENT DIMS. = 5

STRUCTURAL STATISTICS OF CLUSTER AT THIS LEVEL

PERCENTAGE CONNECTIVITY OF CLUSTER = 92.

NUMBER OF CLUSTERS = 6.

NUMBER OF ELEMENTS REMAINING = 7.

PERCENTAGE SIMPLICITY GAINED BY CLUSTERING AT THIS LEVEL = 56.2500 PERCENT

PERCENT DATA LOSS LINK 1 FROM LOWER LEVEL = 4.

PERCENT DATA LOSS LINK 2 FROM LOWER LEVEL = 8.

PERCENTAGE DATA LOSS BY CLUSTERING AT THIS LEVEL = 8.088 PERCENT

COST BENEFIT RATIOS BY CLUSTERING AT THIS LEVEL = 6.995.

Case Study No. 1.
ANALYSIS OF STUDY STRUCTURE - ELEMENTS

NEW CLUSTER

7 CLUSTERS REMAIN AFTER CLUSTERING CLUSTER 3 AND CLUSTER 5 TO FORM CLUSTER NUMBER 3 OR 115

CLUSTER 3 NUMBER OF ELEMENTS = 4

ELEMENT LISTING

PATRICK PROFILE

3 CAMPUS
4 2 3 4 5 6 7 8 9 10

5 INFORMAL ENTRY
2 3 4 5 6 7 8 9 10

6 OUTOUR ENING
3 4 5 6 7 8 9 10

11 LAUNDRY
3 4 5 6 7 8 9 10

TOTAL AREA OF CLUSTER = 420,000 SF

TOTAL NUMBER OF SIMILAR ATTRIBUTES = 5

STRUCTURAL STATISTICS OF CLUSTER AT THIS LEVEL

PERCENTAGE CONNECTIVITY OF CLUSTER = 92.

NUMBER OF CLUSTERS = 5.

NUMBER OF ELEMENTS REMAINING = 2.

PERCENTAGE SIMPLICITY GAINED BY CLUSTERING AT THIS LEVEL = 62.5000 PERCENT

PERCENT DATA LOSS LINK 1 FROM LOWER LEVEL = 3.

PERCENT DATA LOSS LINK 2 FROM LOWER LEVEL = 6.

PERCENTAGE DATA LOSS BY CLUSTERING AT THIS LEVEL = 7.904 PERCENT

COST BENEFIT RATIO BY CLUSTERING AT THIS LEVEL = 7.907

Case Study No. 1.
ANALYSIS OF STUDY STRUCTURE - ELEMENTS

NEW CLUSTER

6 CLUSTERS REMAIN AFTER CLUSTERING ELEMENT 4 AND CLUSTER 7 TO FORM
CLUSTER NUMBER 4 OR 111

PRESENT STRUCTURE OF STUDY

CLUSTER 4 NUMBER OF ELEMENTS = 5

ELEMENT LISTING - MATRIX PROFILE

<table>
<thead>
<tr>
<th>ELEMENT LISTING</th>
<th>MATRIX PROFILE</th>
</tr>
</thead>
<tbody>
<tr>
<td>4 FORMAL LB.</td>
<td>0.1 0.4 0.0 2.3 0.0 1.1 1.3 3.1</td>
</tr>
<tr>
<td>7 FORMAL LIVING</td>
<td>2.0 1.4 0.0 2.1 2.0 1.4 3.2</td>
</tr>
<tr>
<td>9 DINING RM.</td>
<td>4.3 1.3 1.1 2.4 4.1 0.1 1.3 3.2</td>
</tr>
<tr>
<td>15 PANTRY RM.</td>
<td>2.0 1.3 1.1 4.3 3.3 1.3 2.1 4.2 2.2</td>
</tr>
<tr>
<td>16 STUDY</td>
<td>2.0 1.3 1.1 4.2 2.1 3.1 1.2 4.2</td>
</tr>
</tbody>
</table>

TOTAL AREA OF CLUSTER = 420.00 SQ. FT.

TOTAL NUMBER OF SIMILAR ATTRIBUTES = 1

STRUCTURAL STATISTICS OF CLUSTER AT THIS LEVEL

PERCENTAGE CONNECTIVITY OF CLUSTER = 91.

NUMBER OF CLUSTERS = 5.

NUMBER OF ELEMENTS REMAINING = 1.

PERCENTAGE SIMPLICITY GAINED BY CLUSTERING AT THIS LEVEL = 48.7400 PERCENT

PERCENT DATA LOSS LINK 1 FROM LOWER LEVEL = 2.

PERCENT DATA LOSS LINK 2 FROM LOWER LEVEL = 9.

PERCENTAGE DATA LOSS BY CLUSTERING AT THIS LEVEL = 9.314 PERCENT

COST BENEFIT RATIO BY CLUSTERING AT THIS LEVEL = 7.382

Case Study No. 1.
ANALYSIS OF STUDY STRUCTURE - ELEMENTS

NEW CLUSTER

5 CLUSTERS REMAIN AFTER CLUSTERING CLUSTER 10 AND ELEMENT 12 TO FORM CLUSTER NUMBER 10 OR 112

PRESENT STRUCTURE OF STUDY

CLUSTER NUMBER OF ELEMENTS = 4

ELEMENT LISTING

10 BATHROOM

12 GUEST BATH

13 PARENT BATH

14 CHILD BATH

TOTAL AREA OF CLUSTER = 500.00 SQ FT

5 ELEMENT STRONGLY DEFINING CLUSTER - BATHROOM

TOTAL NUMBER OF SIMILAR ATTRIBUTES = 3

STRUCTURAL STATISTICS OF CLUSTER AT THIS LEVEL

PERCENTAGE CONNECTIVITY OF CLUSTER = 0.0,

NUMBER OF CLUSTERS = 5,

NUMBER OF ELEMENTS REMAINING = 0.

PERCENTAGE SIMPLICITY GAINED BY CLUSTERING AT THIS LEVEL = 75.0000 PERCENT

PERCENT DATA LOSS LINK 1 FROM LOWER LEVEL = 1

PERCENT DATA LOSS LINK 2 FROM LOWER LEVEL = 10

PERCENTAGE DATA LOSS BY CLUSTERING AT THIS LEVEL = 95.00 PERCENT

COST BENEFIT RATIO BY CLUSTERING AT THIS LEVEL = 0.000

Case Study No. 1.
### Analysis of Study Structure - Elements

**Threshold Num: 113**

<table>
<thead>
<tr>
<th>New Cluster</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>4 Clusters remain after clustering Cluster 1 and Cluster # to form Cluster number 1 or 113</strong></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Present Structure of Study</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Cluster 1</strong> Number of Elements = 4</td>
</tr>
</tbody>
</table>

**Element Listing**

<table>
<thead>
<tr>
<th><strong>Element</strong></th>
<th><strong>Matrix Profile</strong></th>
</tr>
</thead>
<tbody>
<tr>
<td>Kitchen</td>
<td>2.4.6.0.2.3.4.2.4.2.3.1.1.2.2.2</td>
</tr>
<tr>
<td>Breakfast/Family</td>
<td>2.4.6.2.1.2.2.3.1.1.1.1.2.2.3</td>
</tr>
<tr>
<td>Living Area</td>
<td>2.4.1.2.3.1.4.2.1.2.2.1.3.2.3</td>
</tr>
<tr>
<td>TV, HU, etc.</td>
<td>2.3.1.1.2.2.3.2.1.2.2.3.2.2</td>
</tr>
</tbody>
</table>

**Total Area of Cluster = 480.00 SQ FT**

**Total Number of Similar Attributes = 1**

**Structural Statistics of Cluster at This Level**

| **Percentage Connectivity of Cluster = 86** |
| **Number of Clusters = 4** |
| **Number of Elements Remaining = 0** |
| **Percentage Simplicity gained by clustering at this level = 81.0500 PERCENT** |
| **Percent Data Loss Link 1 from lower level = 7** |
| **Percent Data Loss Link 2 from lower level = 12** |
| **Percentage Data Loss by clustering at this level = 12.040 PERCENT** |
| **Cost Benefit Ratio by clustering at this level = 6.748** |

---

Case Study No. 1.
VIII. BIBLIOGRAPHY


51. Wirth, M., Eastbrook, G. F. and Rogers, D. J. "A Graph Theory Model of Systematic Biology with an Example for the Oncidiinae (Onchidaceae)" Systematic Zoology, Vol. 15, Number 1., pp. 32-47.