WORKING PAPER 416

STRUCTURAL ANALYSIS OF SOCIAL DATA: A GUIDE TO HO's GALOIS LATTICE APPROACH AND A PARTIAL RE-SPECIFICATION OF Q-ANALYSIS

ABSTRACT

A pedagogic exposition of Ho's Galois lattice approach for analysing social data is given in this paper. This provides a basis for re-specifying Atkin's Q-analysis in a way that is believed to be more powerful, more economical and accessible to a wider audience. Concepts of traffic and eccentricity are re-examined, but more generally, the suggested re-specification may enhance the utility of Q-analysis and broaden its applicability, while remaining true to its underlying ethos.

ACKNOWLEDGEMENTS

I would like to thank Helen Couclelis, Steve Seidman, Tony Gatrell, Lionel March and Graham Chapman for their comments on an earlier draft of this paper. The paper was written as part of a research project sponsored by the Economic and Social Research Council.

S.M. Macgill
School of Geography
University of Leeds

First draft - February 1984
First revision - May 1984
Second revision - November 1984
1985 WORKING PAPER

INTRODUCTION

The main purpose of this paper is to provide an introductory guide to Ho's recently developed Galois lattice approach for interpreting social data. The need for such a guide arises from the consideration that the remarkable series of papers within which the approach appears (Ho 1982 a,b,c,d; see also March 1982,1983) may not be accessible to many for whom the approach may be relevant. The implication here is that the approach is at the same time simple and powerful, but, perhaps paradoxically, these telling qualities may not be sufficiently apparent from the original literature.

An understanding of Ho's Galois lattice approach provides a perspective for re-examining the structural features (language?) of Atkin's (1974) Q-analysis. In particular it appears to form a basis for reformulating these aspects of Q-analysis with greater precision, greater transparency, greater economy, and in a form that should be accessible to a wider audience. It is believed that the reformulation not only remains true to the underlying general theory and philosophy of Q-analysis but actually provides a basis for refining and extending the applicability of the traffic concept and associated theories of transmission: aspects that have proved troublesome to develop in empirical contexts. New definitions of parameters of structure are also suggested, to overcome weaknessess in these as currently developed in Q-analysis and to widen the utility and applicability of this style of approach. The suggested re-specification appears to be in agreement with Griffiths' (1983) re-working of Q-analysis, though the approach below has a different origin, emphasis and objective and requires less mathematical expertise on the part of the reader. The re-examination of Q-analysis is not, however, the primary purpose of the present paper and a fuller specification may be developed in later work.

The type of data for which Ho's approach and the structural "language" of Q-analysis is applicable is as given in Table 1 - a binary matrix or, what Ho would more appropriately term a logical array, representing a relationship between a set of objects (a,b,c,d,e,f) say, and a set of features (1,2,3,4,5,6,7), where the terms "objects" and "features" are afforded very wide interpretations. There are many contexts in which such data may arise - see Table 2 for possible examples of objects and features. The applicability of the method in the context of weighted data is also discussed below.

A prominent philosophy behind all but the most simplistic methods of interpreting data of the type given in Table 1 is that the features of interest lie in, or are based on an identification of, how similar and how distinctive individual objects or groups of objects are from others. If the array contained nothing but 0's or nothing but 1's, there would be no such interest - everything would be the same. Interest in social systems that can be sensibly represented in a binary relation derives from the fact that such patterns of data often have an intriguing complexity. What is needed is a gentle but efficient method of revealing and probing it.

Many classificatory algorithms and exploratory methods have been developed, seeking to consolidate similarities and identify distinctive patterns as a means of guiding data interpretation. With very few exceptions (see below) they all involve some loss of information. Ho's method suffers no such loss; the given data is unravelled so as to pinpoint precisely and definitively where similarities between objects lie in terms of the features they have in common, and the resulting aspects of similarity are reorganised in a particularly logical way and in the most compact possible form. This appears to be a very significant achievement and, according to March 1982 "has the widest significance in all those fields where classification and taxonomy are relevant" (p.424). As suggested above it also has the widest significance to those who have been attracted to the utility of the style of analysis made possible through Q-analysis for examining the internal structure of social data. The style and format of "output" is likely to be unfamiliar to many potential users, though one of the aims in providing the guide below is to prevent this unfamiliarity inhibiting a realisation of the methods utility.

AN OUTLINE OF THE METHOD

The essence of Ho's method can be conveniently conveyed in diagrammatic or graphical form, and for ease of comparison with the original literature the same pilot data set as in Ho 1982c and March 1982 is used (Table 1), this being done without loss of generality. The data set as a whole can be represented in graphical form (Figure 1a). From this, separate graphs can be drawn depicting which features each object displays, and then further graphs can be drawn depicting all the mutual similarities between objects in terms of sharing common features. For a relatively small data set, such as that now being handled, this can be approached simply and methodically. The features displayed by each object can be identified first (see Table 3), then mutual similarities between pairs of objects can be identified, then mutual similarities between triples of objects, then mutual similarities between quadruples. The procedure could be extended to quintuples and beyond, should further similarities arise, but this does not occur with the present data set.

In technical terms the resulting graphs would be called the complete subgraphs or cliques of Figure 1b and they are sketched in Figure 2. (Figure 1b has been constructed from Figure 1a by adding extra links to connect all objects to each other and all features to each other - this makes the graph technically complete, though does not materially alter it). It can be formally proved that the graphs given in Figure 2 not only capture all the mutual similarities between objects, in terms of sharing common features, but also necessarily capture all the mutual similarities between features in terms of being displayed by given objects.

In terms of identifying distinct similarities between objects and features a number of the graphs given in Figure 2 turn out to be redundant, being wholly contained within others. For example (e, 14) is wholly contained with (de, 14); the first of these parentheses identifies the fact that object e displays features 1 and 4; the second of the parentheses clearly also identifies this fact, as well as the additional fact that object d also displays these features. The parentheses that are redundant are marked x in Table 3. Those that are not redundant must be retained: the object-feature relations they represent are given in Figure 3. In technical terms these would be called the maximally complete subgraphs from Figure 2.

Three aspects in particular can be noted. First, that the graphs given in Figure 3 represent a definitive isolation of all the similarities from the original data relation, there being no redundancy whatever in this representation: the similarity structure of the data set, then, is contained uniquely and completely in the graphs of Figure 3, or, identically, in the sets of terms represented there - those that are not marked 'x' in Table 3.

Second, that Ho has developed an efficient algorithm (see Ho 1982d, explained in the Appendix to this paper) for isolating the information in these graphs. This algorithm achieves the same ends but follows a different procedure from that indicated above because, although the procedure above quite explicitly revealed the underlying logic of the method, it would become cumbersome for a rather larger data set, and the computations required would increase exponentially with the size of the data matrix.

Third, that the parentheses of Table 2 (or, equivalently, the graphs of Figure 3) can be ordered in a particularly useful way, bringing out structural interrelationships that are not immediately apparent from the set of graphs in Figure 3 or the list in Table 2. This ordering involves the construction of Ho's Galois lattice, and, for the example worked through above, is given in Figure 4.

The node (M,Ø) at the top of this lattice is included for completeness to close the structure. Below this is the only grouping from Table 2 (or Figure 3) revealing a mutual similarity between four objects. Below this are groupings involving mutual similarities between three objects, and below this again, mutual similarities between pairs of objects. Below again are features displayed uniquely by single objects and, for completeness, (Ø,N) closes the lattice.

These different groupings become nodes on a lattice by linking them, whenever, for any two nodes at consecutive "levels" within the lattice, the objects in the node at the level below are a subset of (are covered by) those in the node at the level above <u>and</u> the features in the node at the level above are a subset of (are covered by) those in the node at the level below. This is a particularly useful way of ordering the elements as it is important to know not only where the similarities and correspondences between the entities lie (ie the nodes of the lattice) but also how the resulting groupings relate to each other. In March's (1982) words (p.426) - "A designer or planner using this representation will always be aware of the most <u>complete</u> picture of the situation at whatever level the examination takes place in the lattice. Such a designer or planner will also know how to compose or decompose the problem locally by moving up and down the lattice". March's remarks assume, of course, that the designer or planner can "read" the lattice meaningfully. This reading will hopefully be aided in what follows in later sections.

In Figure 4b, an alternative lattice representation of the same results is given. Whereas in Figure 4a each level referred to a given number of objects, in Figure 4b each level refers to a given number of features. Apart from this difference, we can consider the two representations to be equivalent to each other. Ho's (1982d) algorithm (see Appendix) incorporates construction of the lattice edges as well as specification of the nodes, yielding a structure of the type of Figure 4a, though Figure 4b could be generated from a minor modification.

COMPARISON WITH THE Q-ANALYSIS APPROACH

We can shed a different light on what Ho's method produces by looking at it alongside Atkin's (1974) Q-analysis, an approach developed earlier for aiding the interpretation of key features of social systems which can be sensibly represented in an array of binary data.
A number of authors (Gould 1981, Beaumont and Gatrell 1981) have been attracted to Q-analysis due to its capacity for exposing key interpretative structural features of binary data without distorting or transforming the data in any way: attention is drawn to the tendency in other methods to aggregate or to transform, and, in avoiding such devices, Q-analysis and the data exploration afforded through its characteristic algorithm has been termed "data friendly", revealing only the "natural structure" of the data. Ho's method would appear to be an alternative approach for revealing a "natural structure" of data, - again no aggregation or distortion is encountered - though in this case a somewhat different representation of "natural structure" is produced, one that is more precise and economical and in turn sheds new and revealing light on Q-analysis.

That a binary relation between two sets (in the present case, between objects and features) has a latent structure can perhaps be most readily appreciated by observing how two geometrical figures can be constructed to represent the original data. In the present case row a in Table 1 is represented by the triangle (simplex) with vertices 2,3 and 6 in Figure 5 b, row b by the four sided figure (simplex) with vertices 1,2,3 and 7, and so on. Correspondingly, the columns in Table 1 can be identified as individual parts (simplices) of Figure 5a. Part of the driving philosophy in Q-analysis is that distinctive or interesting features of such structures (holes, protruding or isolated parts, dense parts) may well correspond to "interesting" or "significant" aspects of the original data. (The fact that different edges in this figure are of different lengths is of no consequence whatever here.) A characteristic algorithm for representing and exploring this structure is used.

For comparative purposes, the data given in Table 2 will be filtered through the characteristic algorithm of Q-analysis. The technical procedure in this algorithm involves identifying how many features each object has in common with each other object, Table 4, and then forming a ranking of the dimensions of objects by the number of features that can be counted, identifying a linkage or connectivity between objects at a given dimensional level, n say, whenever they have n features in common with each other (Figure 6 a). A diagrammatic representation of the resulting ranking and connectivity structure is given in Figure 6 b, and in Figure 6 c so-called q-nearness graphs depict precisely where the connectivities within each group at each level lie.

It is not suggested that the nature of Atkin's (1974) Q-analysis can be grasped simply from a mechanical application of this algorithm. The algorithm is but one aspect of the approach and to some, not the most interesting. It is simply a device for filtering out connectivities at successive dimensional levels these being considered to represent important aspects of structure. Note that the algorithm has simply filtered out the different dimensional levels of the geometrical structure in Figure 5 a, drawing out the connectivities between objects at each level. In a similar way, the same algorithm could be used to reveal the connectivity between the features in the conjugate simplicial complex, in terms of being held in common by different objects. Gould (1980) uses the analogy of looking at a system through spectacles of different dimensional strengths. At level 2, we should see only objects and faces of objects that are of dimension 2, and if there are connectivities of this dimension between objects we should see a space or a tunnel of dimension 2 linking them together. Similarly a different strength of spectacles, at level 4, say, would enable us to see only 4 dimensional objects, and interconnections, if any exist, of that dimension. Thus one way of viewing the written output of the Q-analysis algorithm is as a depiction of the number of pieces the full geometry falls into at each dimensional level, along with the composition of each piece. In the general theory of Q-analysis the connectivities (and lack of them) are regarded as being of paramount importance as enabling and constraining spaces for Atkin's distinctive concept of traffic.

It may be somewhat to the dismay of the purist that many authors have utilised the Q-analysis algorithm with little or no explicit attempt to invoke "traffic" or other aspects of the approach (see Macgill and Springer 1985 for a review), and have intimated immense benefit from doing so in terms of gaining interpretative insights into data sets, in some cases suggesting a gestalt appreciation of the fabric of the data (Chapman 1983) and an assurance that the structure being revealed is a natural structure. As discussed elsewhere (Macgill 1984), such authors are in fact indulging in a new interpretation of the single link method of cluster analysis, for the algorithm used in Q-analysis is technically no more than that, using a between-object similarity coefficient derived by counting common features. The interpretation of what this simple technical procedure can offer is, however, quite different in the different traditions of Q-analysis and cluster analysis.

What is of interest at this stage is to compare the style of output from the Q-analysis algorithm with that given from Ho's lattice representation, in terms of the structural insights into the data that are yielded. The aspects exposed by the Q-analysis algorithm are the number of features displayed by each object (seen most vividly from the respective dots in Figure 6b - reflecting the highest Q-levels at which they appear), and the bonds or connectivities (or absence of bonds or connectivities) between objects at each level - seen most vividly from the q-nearness graphs in Figure 6c - though it would appear to be a serious omission in many analyses that such graphs are not explicitly invoked (see Springer and Macgill?). In both cases, Figure 6a provides a summary picture, but in neither case has it been revealed which features are involved. (ie we know how many features, but we do not know which they are). This could only be found from a careful and time consuming examination of the original data. Those familiar with the approach will appreciate that the appearance of, for example, two objects in the single component at level 2 is not necessarily a reflection of the fact that they have two common features. Objects a and d, for example are not directly linked, but are only chained together via b and c as intermediaries. This chaining is not in itself an undesirable feature of the approach: it may reflect significant channels for "traffic" (see the core literature on Q-analysis, but then also the additional arguments below). It does mean, however, that the components given in Figure 6a cannot necessarily be interpreted as groupings of <u>similar</u> objects; only groupings of <u>connected</u> objects.

The features exposed in the lattice representation are the precise similarities between objects in terms of the actual (named) features they display. Figure 4b reflects some of the pattern portrayed in Figure 6 - objects b and c at the four "feature" level, objects a, b,c and d at the three "feature" level, and so on - such as a simple ranking algorithm would also indicate. However, the lattice representation also pinpoints which specific features are involved in each case, and in doing so it identifies similarities between objects in terms of these features at each level. Nodes only appear on the lattice if they pick up similarities[1] that cannot be picked up at other nodes: it is in this sense that there is no redundancy in the approach.

Footnote [1] The term similarity is used above in a very precise sense, for as March (private communication) has remarked, the object - feature correspondences that define each node on the lattice can be thought of as sub-universes (the original logical array replete with 1's being their "universe"). For a creature - or machine - with the ability just to sense those pervasive features which characterise every object in such

(Continuation of footnote)
a sub-universe, no individual object is distinguishable from the sub-universe itself. The objects in such a sub-universe are not merely similar, they are the same within these terms of reference. The only reason for the distinguishability of the objects lies outside the limiting confines of that particular sub-universe (ie at other nodes and levels in the lattice). Additional sensors will be required to pick out the other distinctive features to discriminate among the objects.]

An alternative way of understanding how the lattice diagram, Figure 4, can be read is to appreciate its relationship to the geometrical (simplicial complex) portrayal of the original data given in Figure 5a and 5b. It gives a two dimensional representation of those multi-dimensional forms, capturing all the key structural features, as can be seen in more detail in what follows.

Nodes <b, 1237> and <c, 1257> in Figure 4a, representing the four-dimensional objects b and c correspond to the named simplices in Figure 6b with vertices 1,2,3,7 and 1,2,5,7 respectively. Lattice nodes <a, 236> and <d, 1457> similarly correspond to the appropriately named simplices in Figure 5b, defining objects a and d. The other named simplices (objects) in Figure 5b are all faces of these simplices and are not separately identified in the lattice. The next four lattice nodes <ab, 23>, <bc, 127>,<cd, 15> and <de, 14> correspond to the three 2-dimensional faces in Figure 5b defined by edges 23, 15, and 14 and the 3-dimensional face 127. Although there are other 2-and 3-dimensional faces in Figure 5b, these others are not individually identified in the lattice because they are simply faces of individual simplices that have already been identified at earlier levels in the lattice. This exclusion reflects a lack of redundancy in the lattice representation. The final three nodes in the lattice <abf, 3>, <abc, 2>, and <bcde, 1> correspond to vertices 1,2 and 3, respectively, in each case specifying on which simplices these vertices are defined. The other four vertices, 4,5,6 and 7, are not identified explicitly because these are defined wholly on simplices that have already been accounted for in the lattice, reflecting again the lack of redundancy in the lattice representation.

Had the correspondence been based on Figure 4b instead of Figure 4a, it would merely have identified the dimensions (of objects and faces for which objects are vertices) in a more orderly way (dimension 4 at level

4, dimension 3 at level 3, dimension 2 at level 2 and dimension 1 - the vertices themselves - at level 1), but this would have occurred at the expense of clearly identifying the four distinct objects a,b,c and d at a single (the lowest) level in the lattice (these objects not all being of the same dimension).[1]

Similar (and dual) correspondences can be drawn for the conjugate simplicial complex in Figure 5a. What is especially worthy of notice is the relative ease with which the lattice representation can successfully capture the key structural features in a manner that - unlike the multi-dimensional complex - can be readily drawn on the page. Moreover, the two parts of Figure 4 (a and b) are so similar that it is as if a single diagram (Figure 4a or 4b) is sufficient where previously there were two - the simplicial complex and its conjugate (Figure 5a and 5b).

To summarise, then, the nodes on the lattice either represent distinctive objects - objects having a combination of features displayed by no other object - these are the nodes with a single letter on the left hand side of the comma - or they represent those cases where features are held in common by collections of objects, giving a complete but most economical specification of the latter - these are the nodes with collections of letters on the left hand side of the comma.[2] In terms of a corresponding simplicial complex, the identification of commonly held features is equivalent to identifying distinctive shared faces and vertices (and the simplices on which they are defined) of appropriate dimension.

The difference between the above and the Q-analysis algorithm will now be seen: at level 2, for example, the Q-analysis algorithm gives a single component with objects a,b,c,d and e. In the lattice approach, the same objects appear at the corresponding level, but in three components, the nodes ab, cd and de each having a different pair of features in common. The Q-analysis connectivity implicit in the single component at level 2

---

1 It might be useful to add to the lattice representation by underlining each object in the highest dimensional (in terms of features) node at which it appears - effectively identifying its full set of features ie. identifying the simplex it defines. Correspondingly it might be useful to mark each feature in the highest dimensional (in terms of objects) node.

2 "As we ascend the lattice, objects become more distinguishable as qualitative information about their emanations increases; as we descend the lattice, qualities become more distinguishable as objectual information about their embodiments increases". March 1983 p .

derives in fact from connectivities at higher levels: of b's four characteristics, two are shared by a and three by c, but a and c only have one in common with each other. The edges in the lattice structure show precisely when and how such links between different levels can (and should?) be made.

Whether the relatively favourable light in which Ho's lattice approach has been indicated in this paper is born out in realistic empirical contexts must await further work. While recognising Atkin's pioneering achievement it is my belief that the output produced from Ho's lattice approach is considerably more informative than that from the usual Q-analysis algorithm, though it is not necessarily more extensive, there being a lot of redundancy in the Q-analysis output. The informativeness of the lattice and the relative redundancy of the latter go hand in hand: many of the individual terms in the Q-analysis listing do not have interpretative significance (they are abstract entities) in terms of the original data, and indeed it is not possible to construct the simplicial complex from the information in the listing of output. On the other hand, each node in the lattice does have a direct interpretation, and the simplicial complex can be constructed solely from the information given in the lattice. Recognition of these apparently superior qualities provides a basis for enhancing and not for denigrating Q-analysis.

It may be sensible to stimulate a reader's intuitive appreciation of what is produced from Ho's method, particularly for those not closely acquainted with Q-analysis, by replacing the abstract terms "objects" and "features" in the pilot example worked through above by one of the examples given in Table 2 (or other choices). Figure 4, for example, can be reinterpreted taking the objects as "shopping centre locations" and the features as "retailing activities". In this case the original data in Table 1 would show the incidence of different retailing activities at different shopping centre locations in some region.

Of the six retail locations, four a,b,c and d, have distinct combinations of activities (see the four nodes at the lowest level in the lattice) the other two, e and f, only providing activities that are also found elsewhere (e providing activities 1 and 4 that are also found at d, and f providing activity 3 that is also found at a and b). The locations providing distinct combinations of activities do have some in common with each other - b and c both provide 1,2 and 7; a and b both provide 2 and 3 (see the nodes on the left at level 3). Moreover, locations a,b and c all have a single activity in common with each other, namely activity 2 (see node abc, 2 at level 2). These and all other correspondences are reflected in the nodes that are defined in forming the lattice. There are no correspondences that are not reflected in the nodes: in this sense the representation is complete or, in the terminology used earlier in the paper, a definitive isolation of all the similarities in the original data relation has been achieved. The exact interrelationships between the lattice nodes are specified through the given edges.

The lattice can be seen to provide a very distinctive and detailed structured organisation of, in this case, data on retail activities according to location. The level of detail is greater than can perhaps be conveniently assimilated in one go, but it may be a pious hope to expect to achieve simple representations of inherently complex systems. The level of detail given in the lattice would also seem to be the minimum that must be handled if some of the original data is not to be lost. Utilising the approach alongside others may be revealing in order to identify where information losses have occurred, and perhaps offering guidance into particular parts of the lattice. The identification of second order structural features - many possibly distinctive to particular types of context - and the development of suitable computer software, both seem worthy of further exploration in view of the size of lattice to be produced (and its consequent swamping potential) from a large data set.

TOWARDS A RE-SPECIFICATION OF Q-ANALYSIS

(i) Traffic on a lattice. The concept of traffic has been developed in Q-analysis and exploited as a powerful and evocative metaphorical concept. The simplicial complex is taken to be a relatively static backcloth which can support, or inhibit "activity" (ie. traffic) - according to its local and global structural configurations. Such traffic could be monetary expenditure, physiological stress, disease, verbal argument, traffic congestion or one of a myriad of other possibilities. The distinction between backcloth and traffic would appear to have some rough correspondence with the longer standing distinctions between pattern and process (pattern here acting as some kind of constraint on process), structure and function, or evidence and argument, and although the idea has been developed in Q-analysis with far greater specificity than hitherto and has been found to yield useful insights in particular contexts, there have been difficulties in translating the concept into a general operational form with a powerful cutting edge (see also, Seidman 1983). As Johnson, a prominent exponent and practitioner of Q-analysis has remarked, "When I wrote my paper on q-transmission I found it very difficult to find convincing examples with q>1, that is, I found it difficult to find systems with the q-transmission property." (Johnson 1983, page).

Part of the difficulty in developing the theory of q-transmission may lie i the restriction that has been (artificially?) imposed that q-transmission can only occur within q-connected components (Johnson 1982). This may be an artificial restriction because these components do not represent readily interpretable features. Moreover they are not such significant features of the simplicial complex (and therefore of the original data) as the lattice nodes and edges (components being derived, not direct features). As such they (q-connected components) may be less successful in providing a basis for the development of applicable theory. Also, the (again artificial) restriction of considering traffic transmission grade by grade, rather than freely allowing changes of grade may be a further restriction that can be usefully overcome.

The possible transmission of traffic from simplex 3 to simplex 2 then to 7, then to 1 then to 5 and/or 4 can be visualised in Figure 6a. At risk of failing to stimulate the interest of readers unsympathetic to the following specific illustration, "traffic" might represent following a line of argument in a simple but effective literary text, beginning with themes a,b and f (simplex 3 - chapter 1?) moving to themes a,b and c (simplex 2 - chapter 2?) and then dropping theme a to dwell on b and c (simplex 7 - chapter 3?) and so on through themes bcde, then cd, ending in the final(chapter?) with just d, and e recalled from earlier. A badly structured text would correspondingly have a disconnected or contorted structure.

In terms of the alternative lattice representation, this would correspond to transmission from node <abf,3> to <abc,2> via <ab,23>, and then via <bc,127> to <bcde,1> and then to <cd,15> and <de,14>. This might be seen even more clearly, with apologies to lattice theorists, if nodes $\langle M, \emptyset \rangle$ and $\langle \emptyset, N \rangle$ were ommitted from the representation and the remaining nodes rearranged in a different order, though preserving linkages . Just as the configuration and shared faces of the simplicial complex depict allowable channels of transmission,(though it is not precisely these that are faithfully picked up in the Q-connected components in the algorithm), so too do the edges and nodes in the lattice. Although the above literary illustration involved traffic being transmitted along quite a long path, in other cases shorter, more local, transmission may be of interest. What is important to note is that the traffic changes dimension during transmission, a characteristic that cannot easily be picked up in the existing formal literature on Q-analysis, the latter being pre-occupied with traffic of a given and fixed dimension (Johnson 1982).

The development of all this in a more precise notational form awaits further work; the suggestion being that, if possible, this would allow a more flexible operational construction of traffic transmission, and more realistic examples of traffic. Whether such development is warranted will depend on whether the basic inter-set binary array is in itself a sufficiently important underlying structural feature of the system of interest. This needs careful consideration context by context.

In cases where it is, the utilisation of the lattice nodes and edges in place of the more traditional components yielded from the Q-analysis algorithm as the operational indicator of "backcloth" would add to and not detract from the basis so far established in Q-analysis. For a context in which traffic would be a more overtly dynamic entity (monetary flows, for example) parallels with feedback mechanisms and pulse processes could be explored, the justification for developing the new lattice basis then depending on its ability to pick up effects that could not otherwise be handled at all, or only more cumbersomely, on the basis of more traditional "systems theory" representations (Forrester 1969, Roberts 1976, Wilson 1981).

(ii) Eccentricity. The concept of eccentricity has been developed in Q-analysis as an indicator for capturing the relative distinctiveness of individual simplices (objects) from the complex (collection) as a whole. This concept can also be re-examined here, and a new counterpart will be suggested which may be more suitable than its predecessor because it overcomes the current weakness of the definition of eccentricity in making no discrimination between a simplex that is joined directly to one other simplex at the appropriate dimensional level (q) and a simplex that is joined directly to many at that dimensional level. In other words, the indicator of eccentricity as currently used does not discriminate as to whether an object has a given set of features in common with one or with many other objects.

Aspects suitable to incorporate in capturing more fully the characteristic of eccentricity for each object appear to be: the number of features it possesses; how many of these are also possessed by other objects; the number of "other" such objects there are. In fact all this information is given in the shared face matrix (equivalently, Table 4) and, taking A_{ij} to represent the item in row i and column j of this table, a new definition of eccentricity for object i can be given by

$$ecc_i = \sum_{j \neq i} \frac{(A_{ii} - A_{ij})}{n}$$

Above, n is the total number of objects. This compares with the previous formula.

$$ecc_i = \frac{A_{ii} - \sup_{j \neq i} A_{ij}}{\sup_{j \neq i} A_{ij} + 1}$$

$$= \frac{\hat{q} - \check{q}}{\check{q} + 1}$$ in standard Q-analysis notation.

It is interesting to note that the new definition of eccentricity could also be used directly in the context of traditional Q-analysis, for although how to translate recognition of the need to incorporate additional detail into something operational was prompted by exploration of the lattice representation, the resulting formula does not require any information that is not readily given in mainstream Q-analysis.

(iii) <u>Generic indicators of other key structural features</u>. The structure vector in Q-analysis records the number of components the simplicial complex falls into at each dimensional level, on this vector, for the example represented in Figure 5 would be

$$(\overset{3}{2}\ 3\ 1\ \overset{c}{1})$$

It is my belief that this vector summarises information of very limited utility, this somewhat negative belief bound up with the reservations expressed earlier about the interpretative significance of q-components, and of traffic of a fixed dimension.

Notwithstanding limitations in this particular case, the rationale for seeking some generic indicators of structure remains sound. In empirical contexts involving a large number of objects and features, the extensiveness of output both from a traditional Q-analysis and a lattice representation in itself suggests the desirability of developing more manageable generic (second order) indicators of structural features. These would be designed to complement, not replace, the extensive detail, and provide further signposts for interpretation. They may in turn have a higher interpretative significance in their own right (as was the case, to some extent, with the structure vector). In cases where such constructs involve the compression of detail, it would be desirable to keep track of what it is that is being compressed (ie. to recover uniquely the compressed detail - a facility that is denied in the case of the structure vector). At a much simpler level, something based on a count of the number of nodes at each level in the lattice may be meaningful.

Another type of structural feature that has attracted particular attention in some of the theoretical literature on Q-analysis is that of a hole in a simplicial complex - a hollow closed area in the complex, bounded by simplices. It has proved troublesome to develop an algorithm for detecting these features in practice, but it is interesting to consider whether such fea ures could be readily identified from a lattice representation. In fact, a recent paper by Johnson (1983) suggests that the foci of interest chosen by Ho in developing his lattice approach are likely to be <u>more powerful</u> than those hitherto considered in the Q -analysis literature as regards the detection of holes in a complex. Johnson (1983) argues for the use of the concept of 'stars' to detect holes, and here the relevance of Ho's lattice approach is immediate, for stars and the nodes on Ho's lattice representation are one and the same thing. Before developing further the mathematical formalism for the representation of holes as structurally significant features, further consideration of their correspondence with important characteristics of the 'real world' may be warranted. More generally, Johnson's (1983) suggestion that stars are likely to emerge as a more useful focus for Q-analysis than q-connected components, reinforces the present argument of advocating a reorientation for Q-analysis.

<u>SUPPLEMENTS TO HO's GALOIS APPROACH</u>

In referring in the above only to aspects of the "structural language" of Q-analysis, consideration of other aspects of Atkin's approach have inevitably been omitted, notably consideration of the manner in which the objects and features - the subjects of the binary relation whose structure is examined - are themselves embedded within respective hierarchies of objects and features. Though Ho does not make reference to this, consideration of these hierarchies along the lines suggested in the Q-analysis literature could be as much a precedent to structural exploration using the Galois lattice approach as they are to explorations using the traditional Q-analysis algorithm.

To restrict structural analysis to binary data is also deserving of comment. The recognised procedure in Q-analysis for handling weighted relation is to reduce it to a binary relation through slicing. Again, this could be readily invoked prior to deriving a Galois lattice, just as it is with a number of other methods. There does remain controversy over the amount of information that is omitted and then sometimes never sufficiently acknowledged in the structural analysis (see Macgill 1983).

It should not be surprising, that it is not immediately clear how best to handle weighted relations in general for it would appear that it implies the attempt to represent continuous data (the weighted relation) in what are essentially discrete ways - Atkin's simplicial complexes, Ho's Galois lattices and structural equivalents.[1] Before searching and even extending the mathematical universe for an appropriate style of approach for handling weighted relations, careful consideration is warranted as to the meaning of the weights in the relation being put forward for examination. In some cases they may merely constitute numerical (quantitative) representations of what are really (discrete) qualitative phenomena: here there may be grounds for viewing each weight on the scale corresponding to a distinct feature, transforming a weighted relation with no structure to a binary (and fairly) sparse relation with much. This raises a more general point about the different types of phenomena that may exist, and therefore the different types of representation that will be required, a far-reaching point though one that will not be pursued further in the present paper.

---

1 Although March (1982 p) comments that "it is obvious how a weighted relatio relation might be represented in this form", the type of weighted relation being referred to here may well be restricted to a particular kind, this point will not be elaborated further here.

CONCLUSIONS

1. Ho's Galois lattice approach would appear to provide a simple but powerful way to gain a particular type of structural analysis of a binary relation. Specific qualities of the approach have been brought out in the pedagogic guide given above.

2. The approach appears to provide a basis for refining and extending Atkin's Q-analysis. This may be the subject of further work by the author, to be presented in its own right in due course. The suggestion that the focus in Q-analysis on connectivity should be replaced by a focus on similarity may not appeal to all who have been attracted to Q-analysis. However, until convincing examples of chains of connection of given dimensions are found, then such a shift in focus may be what is required to add new life and applicability to this style of approach. If those using Q-analysis are genuinely interested in exploring the structure of the simplicial complex corresponding to a binary relation, then Ho's lattice would appear to provide a far more suitable guide than the listings of components and shared face matrices from the traditional Q-analysis algorithm.

3. There would appear to be a case for reviewing and possibly re-working empirical Q-analyses that have been undertaken in the past in the light of the approach reviewed above, with a view to gaining additional structural insights. Such re-working may in some cases even suggest that significant features pointed to by the conventions of Q-analysis are not what they have seemed.

4. There remains scope for much further work on ways of reducing structural representations of the type considered above to more compact forms while preserving or isolating important structural features, possibly alongside the development of second order indicators,and scope also for developing structural analyses of weighted relations. There is also more generally a need to build on the essentially descriptive focus of the material presented above towards a more explanatory angle, this being approached with the greatest sensitivity for the nature of the social phenomena that it may be suitable to depict in this type of representation. Couclelis (1983) suggests some severe limitations in the utility of approaches that, as with Ho's approach, represent objects simply as collections of features and not as different kinds of functions of these features. March(1983) on the other hand suggests a wide domain of applicability for Ho's representation and powerful generalisations that can follow. There would appear to be rich territory underlying each of these views, which has scarcely been touched by social scientists.

REFERENCES

R.H. Atkin (1974) Mathematical structure in Human Affairs, Heineman.

J.R. Beaumont and A.C. Gatrell (1982) An Introduction to Q-analysis C.A.T.M.O.G. 34 Geo Abstracts, Norwich.

G. Chapman (1983) A structural analysis of two farms in Bangladesh, in T.P. Bayliss-Smith and S. Wanmali (eds.) Understanding Green Revolutions, Cambridge University Press.

H. Couclelis (1983) On some problems in defining sets for Q-analysis, Environment and Planning B 10, 423-438.

J.W. Forrester (1969) Urban dynamics, M.I.T. Press, Cambridge, Mass.

P. Gould (1980) Q-analysis: an introduction for social scientists, geographers and planners. International Journal of Man-Machine Studies 12, 169-99.

P. Gould (1981) A structural language of relations, ch 17 in R.G. Craig and M.L. Labovitz (eds) Future trends in Geomathematics.

H.B. Griffiths (1983) Using mathematics to simplify Q-analysis, Environment and Planning B 10, 403-422.

Y.S. Ho (1982a) The planning process: mappings between state and decision spaces, Environment and Planning B 9, 153-162.

Y.S. Ho (1982b) The planning process: a formal model, Environment and Planning B 9, 377-386.

Y.S. Ho (1982c) The planning process: fundamental issues, Environment and Planning B 9, 387-395.

Y.S. Ho (1982d) The planning process: structure of verbal descriptions, Environment and Planning B 9, 397-420.

J.H. Johnson (1982) Q-transmission in simplicial complexes, International Journal of Man-Machine Studies 16, 351-371.

J.H. Johnson (1983) Q-analysis: a theory of stars, Environment and Planning B 10.

S.M. Macgill (1983) Q-analysis: issues and non-issues, Environment and Planning B 10, 371-380.

S.M. Macgill (1984) Cluster analysis and Q-analysis, International Journal of Man-Machine Studies 20, 595-604.

S.M. Macgill and T. Springer (1985) An appraisal of Q-analysis through its applications (in preparation).

L. March (1982) On Ho's methodological approach to design and planning Environment and Planning B 9, 421-427.

L. March (1983) Designs in a universe of chance: a Piercian construction Environment and Planning B 10, 471-484.

F.S. Roberts (1976) Discrete Mathematical Models, Prentice-Hall Inc.

S. Seidman (1983) Re-thinking backcloth and traffic: perspectives from social network analysis and Q-analysis Environment and Planning B 10, 439-456.

A.G. Wilson (1981) Geography and the Environment: Systems Analytical Methods, John Wiley and Sons.

Table 1

features / objects	1	2	3	4	5	6	7
a	0	1	1	0	0	1	0
b	1	1	1	0	0	0	1
c	1	1	0	0	1	0	1
d	1	0	0	1	1	0	0
e	1	0	0	1	0	0	0
f	0	0	1	0	0	0	0

Table 2 Some "objects" and associated "features"

People	Socio-economic characteristics
Shopping centres	Types of retail outlet
Buildings	Constructive features
People	Events in which they are engaged
Spatial areas	Plant species type
Towns	Socio-economic attributes
Farmers	Farming characteristics
Community areas	Occupational groups
Routes in a road system	Links from which routes are composed

Table 3a Features displayed by each object

(a, 236)
(b, 1237)
(c, 1257)
(d, 145)
(e, 14) x
(f, 3) x

Table 3b Mutual similarities between pairs of objects

(ab, 23)		(be, 1)	x
(ac, 2)	x	(bf, 3)	x
(af, 3)	x	(cd,15)	
(bc, 127)		(ce, 1)	x
(bd, 1)	x	(de, 14)	

Table 3c Mutual similarities between triples of objects

(abc, 2)
(abf, 3)

Table 3d Mutual similarities between quadruples of objects

(bcde, 1)

Table 4 Similarities between objects (subtract 1 from all entries to derive shared face matrix)

	a	b	c	d	e	f
a	3	2	1	0	0	1
b	2	4	3	1	1	1
c	1	3	4	2	1	0
d	0	1	[illegible]	3	2	0
e	0	1	1	2	2	0
f	1	1	0	0	0	1

Figure 1 : Graphical representation of data in table 1

Figure 1a

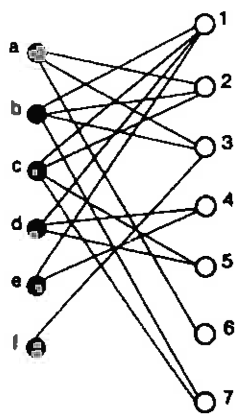


Figure 1b

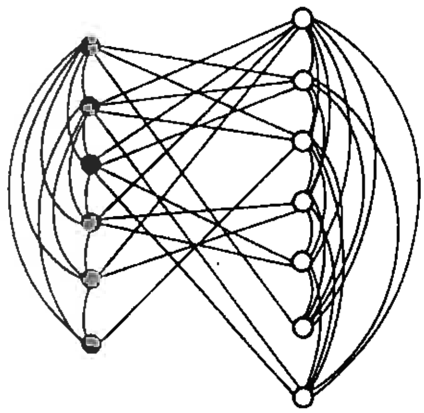
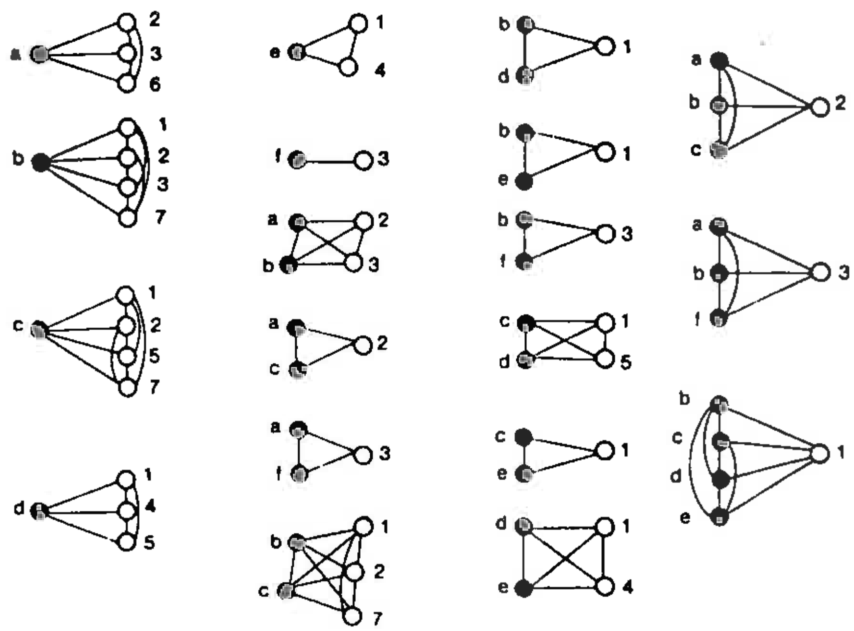


Figure 2 : Cliques of figure 1b



<u>APPENDIX</u>

<u>Ho's Efficient Algorithm for Deriving the Lattice</u>

The purpose of this Appendix is to provide a simplified account of Ho's fast algorithm. However, it should be noted that some discrepancies in the detail of what is required have been encountered, leaving a couple of outstanding issues, and the present account should not be assumed to be definitively authoratitive.

One way of finding the elements defining the nodes of the lattice is to follow the procedure summarised in the text. However, the number of operations involved in so doing would increase exponentially with the number of objects (or features) that are specified. This explosion in the number of operations as the number of data points increases is a problem that has long been encountered in searching for cliques in the graph-theoretic literature. Ho (1982d) presents a different algorithm for deriving the elements of the lattice, one which does not suffer such explosive effects, (and, as an aside, a version of this algorithm may also be applicable for clique identification more generally?)

Before outlining the structure of the algorithm, a comment on its context is in order. The majority of the paper in which it appears, Ho (1982d), is concerned with establishing the fact that the <u>complete</u> set of terms (or elements) of the form $(A,\lambda(A))$, where A is a subset of objects and $\lambda(A)$ is the subset of features all the elements of A have in common with each other, is a lattice. In other words, the set of terms that identify the definitive similarities between objects and features define a lattice. The significance of this is that they must conform to the algebra of lattices. This means that there are operations that when applied to any pair of elements of the lattice will yield again an element of the lattice. There are two of these operations and these are called the <u>meet</u> $\wedge$ (defined as the greatest lower bound of the pair of elements being handled) and the <u>join</u> $\vee$ (defined as the least upper bound of the pair of elements being handled).

If, then, we are presented with only some of the elements of the lattice, we might be able to find more of them by applying the operations $\wedge$ and $\vee$ (and, in particular, $\vee$) to the elements we have. In fact, by starting with what in a sense are the leanest of the terms (the minimal elements) the $\vee$ operation (similar to the set union operation $\cup$) can be used to generate any additional terms that can possibly exist on the lattice.

The minimal terms are those which characterise each object in turn (either as a distinctive set of features, or as a set of features that are also displayed by certain other objects) and those which characterise each feature in turn (either as possessed uniquely by a given set of objects, or as possessed in conjunction with other features by a given set of objects). Any duplicate terms are discarded. In fact, for the worked example, (see table A1 for the basic data) the minimal terms turn out to be ten of the eleven terms on the whole lattice. In other cases we should not necessarily expect such a high proportion.

The first stage of the algorithm is to generate the minimal elements of the lattice. These are found by working in turn on each individual object and then on each individual feature.

From object a, we get the term (a,236), for object a and only object a displays features 2, 3 and 6.

From object b, we get the term (b,1237), for object b and only object b displays features 1, 2, 3 and 7.

From object c, we get the term (c,1257), for object c and only object c displays features 1, 2, 5 and 7.

From object d, we get the term (d,145), for object d and only object d displays features 1, 4 and 5.

From object e, we get the term (de,14), for object e displays features from 1 and 4, and these are also displayed by object e.

From object f, we get the term (abf,3) for object f displays feature 3, and this is also displayed by objects a and b.

From feature 1 we get the term (bcde,1) for feature 1 is displayed by each of b, c, d and e.

From feature 2 we get the term (abc,2) for feature 2 is displayed by each of a, b and c.

From feature 3 we get the term (abf,3) for feature 3 is displayed by each of a, b and f.

From feature 4 we get the term (de,14) for feature 4 is displayed by d and e, and these also both display 1.

From feature 5 we get the term (cd,15) for feature 5 is displayed by c and d, and these also both display 1.

From feature 6 we get the term (a,236) for feature 6 is displayed by a, and this also displays 2 and 3.

From feature 7 we get the term (bc,127), for feature 7 is displayed by b and c, and these also both display 1 and 2.

The inspection procedure used to generate each of these terms can be readily mechanised. Row i of table A1 will be denoted by λ_i, column j by λ^j, and cell ij by λ_i^j. We form the arrays σ and α where

$$\sigma_i = \prod_{j:\lambda_i^j \neq 0} \lambda^j \quad \text{and} \quad \alpha_i = \prod_{j:\lambda_j^i \neq 0} \lambda_j$$

$\prod$ denotes termwise Boolean multiplication on the vectors λ^j and λ_i. This multiplication follows the rules: 0.0=0, 0.1=0, 1.0=0. 1.1=1; so, for example, since λ^2 = (111000) and λ^3 = (110001) then $\prod_{j=2,3} \lambda^j = \lambda^2 \lambda^3$ = (110000).

See tables A2 and A3 for the arrays σ and α respectively. The minimal terms being sought are then found by pairing row of σ with each row of λ, and by pairing each column of λ with each row of α. So, for example, pairing row σc with row λc gives (c,1257) and pairing column λ^6 with row α_6 gives (a,236). Whether derived from inspection or using the mechanised procedure, any duplicate terms should be deleted. Since there are 6 objects and 7 features, in this case thirteen terms are generated, but 3 turn out to be duplicates. The ten remaining terms (the minimal terms) are:

(a,236)	(bcde,1)
(b,1237)	(abc,2)
(c,1257)	(cd,15)
(d,145)	(bc,127)
(de,14)	
(abf,3)	

Note that these are all of the form (A, λ (A)) where A is a subset of objects, and λ(A) is the subset of features displayed by each and every member of A.

It is now necessary to determine whether further lattice terms are generated when the join operation is applied to the above terms. Ho does this by starting in round 1 with those with just one object, then in round 2 with those with two objects, then in round 3 with those with three, and so on. The meet operation is used in each round to determine which elements from previous rounds need not be looked to to generate further terms (being contained within other terms).

The join $\vee$ and meet $\wedge$ operations which are used need to be explained.

For two lattice elements $(A, \lambda(A))$ and $(B, \lambda(B))$ the join is defined as:

$$(A, \lambda(A)) \vee (B, \lambda(B)) = (\lambda^{-1}(\lambda(A) \cap \lambda(B)), \lambda(A) \cap \lambda(B))$$

and the meet as:

$$(A, \lambda(A)) \wedge (B, \lambda(B)) = (A \cap B, \lambda(A \cap B))$$

$\cap$ is the usual set intersection operation.
Note that for the join, the term $\lambda(A) \cap \lambda(B)$ is the subset of features possessed both by the objects in A and by the objects in B. $\lambda^{-1}(\lambda(A) \cap \lambda(B))$ is then the subset of objects that each possesses this set of features. For the meet, $A \cap B$ is the subset of objects common to A and B, and $\lambda(A \cap B)$ are the features common to all members of this subset of objects.*
The remainder of the algorithm now follows.

Let S = the initial set of 10 terms, and add $(M,\emptyset)$**, for completeness. This set will be expanded by adding to it any further terms that are generated as lattice nodes.

Let L = $(\emptyset,N)$**. L will serve as a kind of "lid" on terms already laid down in the lattice that need be considered no further.

Let S (n) = all the terms in S with n objects.

Set n = 0 initially and n = n + 1 on each successive round.

Round n:

Step 1 (a) Form the join between each pair of elements in S (n) and if a term is generated that is not already an element of S, it should be added to S.

1 (b) Form the join between each element in S (n) and each element in L and if a term is generated that is not already (or has not already been) an element of S, it should be added to S.

** Terms $(M,\emptyset)$ and $(\emptyset,N)$ can be read (abcdef,0), (0,1234567), saying that there are no features common to all objects, and no object that displays all features.

* $A \cap B$ would give a subset of terms $x_1, x_2, \ldots x_n$.

$$\lambda(A \cap B) = \lambda(x_1, x_2, \ldots x_n) = \lambda(x_1) \cap \lambda(x_2) \cap \lambda(x_3) \ldots \cap \lambda(x_n)$$

as in Ho 1982d equation 8.

2*[1] Terms S (n) will be the terms for level n in the lattice. Each element of S (n) should be linked by an edge to any terms at a lower level if the element at level n <u>covers</u> the element at the lower level*[2].

3 (a) Form the meet between each pair of elements in S (n) and if the meet is an elemnt of L, delete this meet from L.

(b) Form the meet between each element in S (n) and each element in L, and if the meet is an element of L, delete this meet from L.

4 Add the elements of S (n) to L. Set n = n + 1 and go back to step 1 (a).

There is a minor (I think) difference between the above explanation of the algorithm and the detail given in Ho (1982d) pp 411-414. A modification has been introduced into step 2 above in order to take account of a numerical discrepancy in Ho's worked example. This discrepancy arises when undertaking step 3 (b) in round 3. (This should be equivalent to step 2 (5) of the third pass in Ho's explanation; Ho 1982d, p. 414). By this stage in the algorithm we have:

$S = \{(abc,2),(abf,3),(bcde,1),(M,\emptyset)\}$

$L = \{(ab,23),(bc,127),(cd,15),(de,14)\}$ and the portion of the lattice given in Figure A1.

$S\ (3) = \{(abc,2),(abf,3)\}$

For round 3

Step 1 (a) $(abc,2) \vee (abf,3) = (M,\emptyset)$. The term $(M,\emptyset)$ has already been in S and is therefore not added to S again.

1 (b) $(abc,2) \vee (ab,23) = (abc,2)$

$(abc,2) \vee (bc,127) = (abc,2)$

$(abc,2) \vee (cd,15) = (M,\emptyset)$

$(abc,2) \vee (de,14) = (M,\emptyset)$

---

*1 The account of this step differs from that given by Ho.

*2 $(A, \lambda(A))$ is said to cover $(B, \lambda(B))$, if $(B, \lambda(B)) < (A, \lambda(A))$ and there exists no other element $(X, \lambda(X))$ such that $(B, \lambda(B)) < (X, \lambda(X)) < (A, \lambda(A))$.

(abf,3) $\vee$ (ab,23) = (abf,3)
(abf,3) $\vee$ (bc,127) = (M,$\emptyset$)
(abf,3) $\vee$ (cd,15) = (M,$\emptyset$)
(abf,3) $\vee$ (de,14) = (M,$\emptyset$)

There are no terms here that have not already been in S so nothing further need be done at this step. i.e. S = S

2 See Figure A2

3 (a) (abc,2) $\wedge$ (abf,3) = (ab,23). The term (ab,23) which is currently in L, should be deleted from L. i.e. L = {(bc,127),(cd,15),(de,14)}

3 (b) (abc,2) $\wedge$ (bc,127) = (bx,127)
(abc,2) $\wedge$ (cd,15) = (c,1257)
(abc,2) $\wedge$ (de,14) = ($\emptyset$,N)
(abf,3) $\wedge$ (bc,127) = (b,1237)
(abf,3) $\wedge$ (cd,15) = ($\emptyset$,N)
(abf,3) $\wedge$ (de,14) = ($\emptyset$,N)

Since the term (bc,127) that has been generated here is an element of L, it should be deleted from L*. None of the other terms generated here are elements of L. i.e. now L = {(cd,15) (de,14)}

4 Adding the elements of S (3) to L gives
L = {(cd,15), (de,14). (abc,2), (abf,3)}. We now Set n = 4 and go back to step 1 (a) for the next round.

Two issues left outstanding from the above interpretation of the algorithm are:

(i) Whether, if correct in principle, step 2 above may be expressed in an operationally more efficient way.

(ii) Whether the ramifications of the discrepancy would more generally be confined to step 2 (where L cannot be used as Ho had intended), or whether they would be of wider significance.

---

* Ho does not do this. The fact that the above working of this step differs from Ho's is the reason for changing the account of step 2. The end result is not affected in this case, though it may be in other cases.

Table A1. The basic logical array λ

		N						
		1	2	3	4	5	6	7
M	a	0	1	1	0	0	1	0
	b	1	1	1	0	0	0	1
	c	1	1	0	0	1	0	1
	d	1	0	0	1	1	0	0
	e	1	0	0	1	0	0	0
	f	0	0	1	0	0	0	0

Table A2. Array for $\sigma = \lambda^{-1} \cdot \lambda$.

		M					
		a	b	c	d	e	f
M	a	1	0	0	0	0	0
	b	0	1	0	0	0	0
	c	0	0	1	0	0	0
	d	0	0	0	1	0	0
	e	0	0	0	1	1	0
	f	1	1	0	0	0	1

Table A3. Array for $\alpha = \lambda \cdot \lambda^{-1}$.

		N						
		1	2	3	4	5	6	7
N	1	1	0	0	0	0	0	0
	2	0	1	0	0	0	0	0
	3	0	0	1	0	0	0	0
	4	1	0	0	1	0	0	0
	5	1	0	0	0	1	0	0
	6	0	1	1	0	0	1	0
	7	1	1	0	0	0	0	1

Figure A1

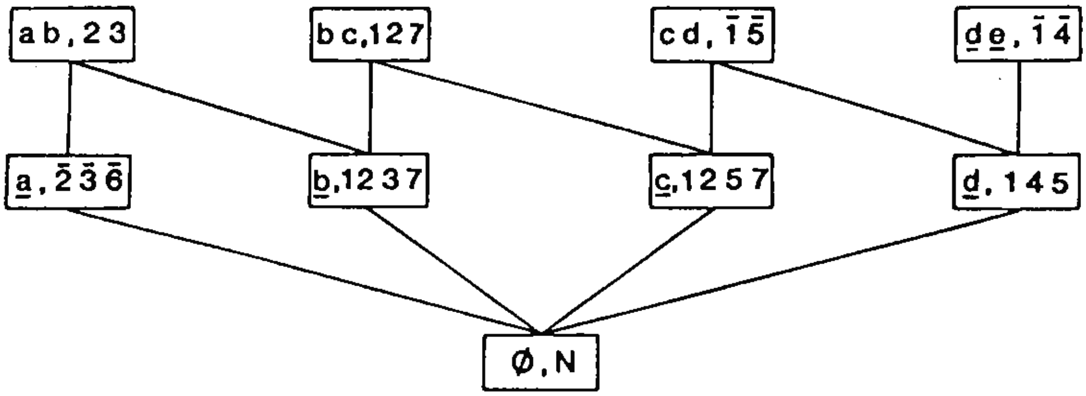


Figure A2

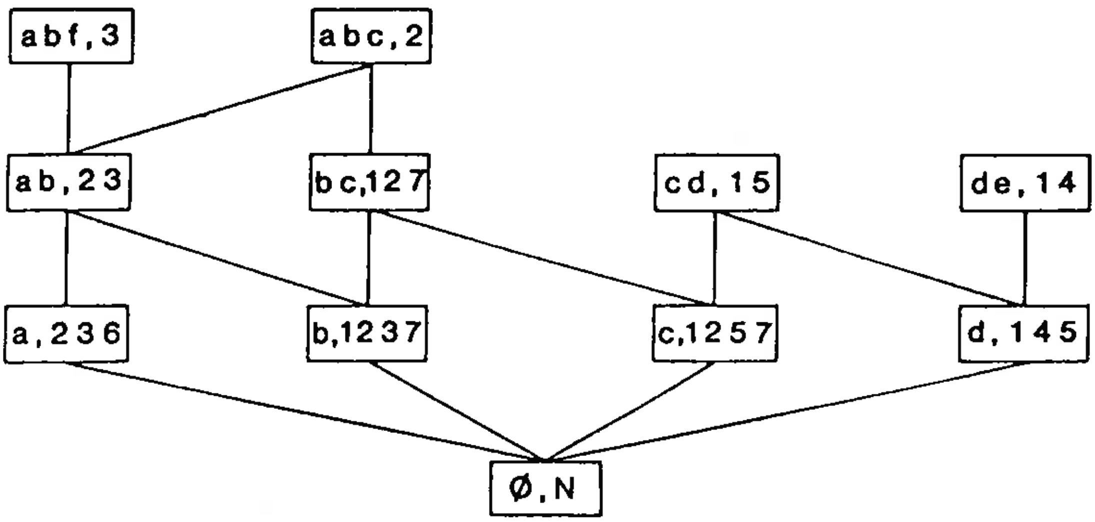


Figure 3: Definitive object-feature relations from table 1

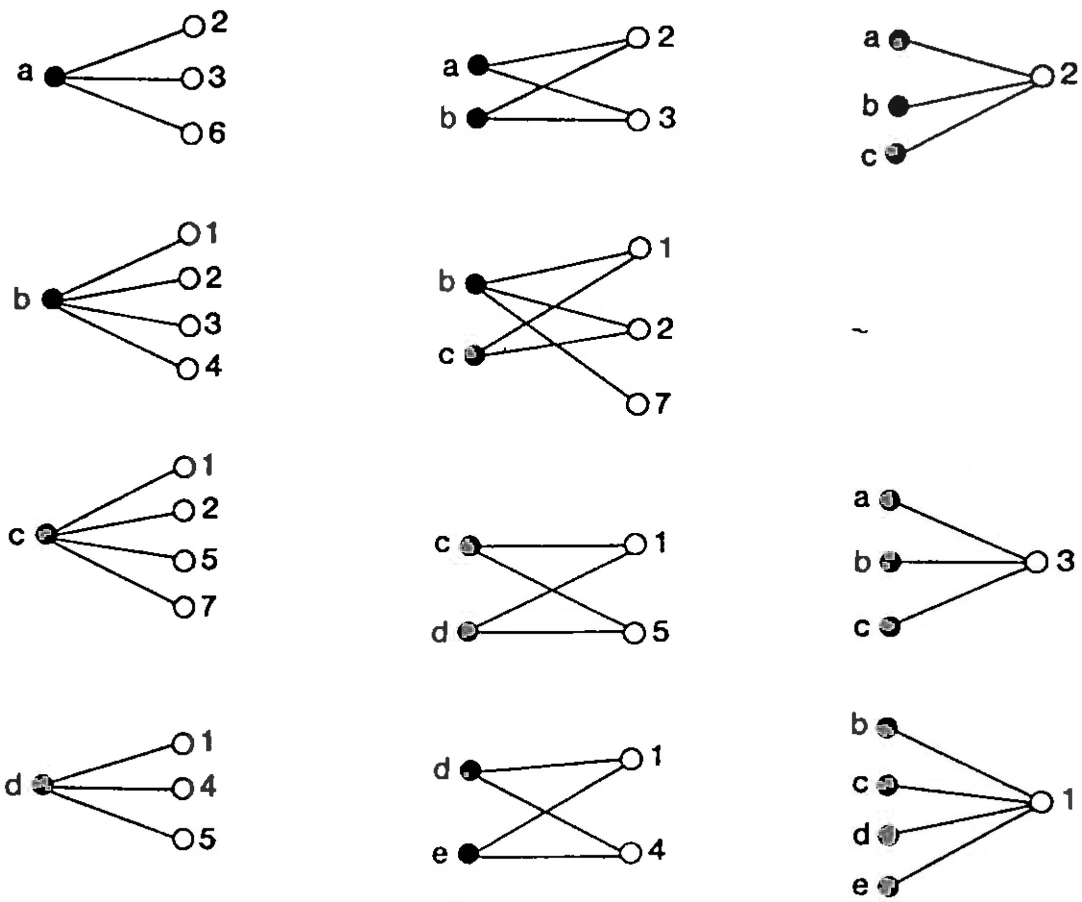


Figure 4a : Lattice derived from data in table 1

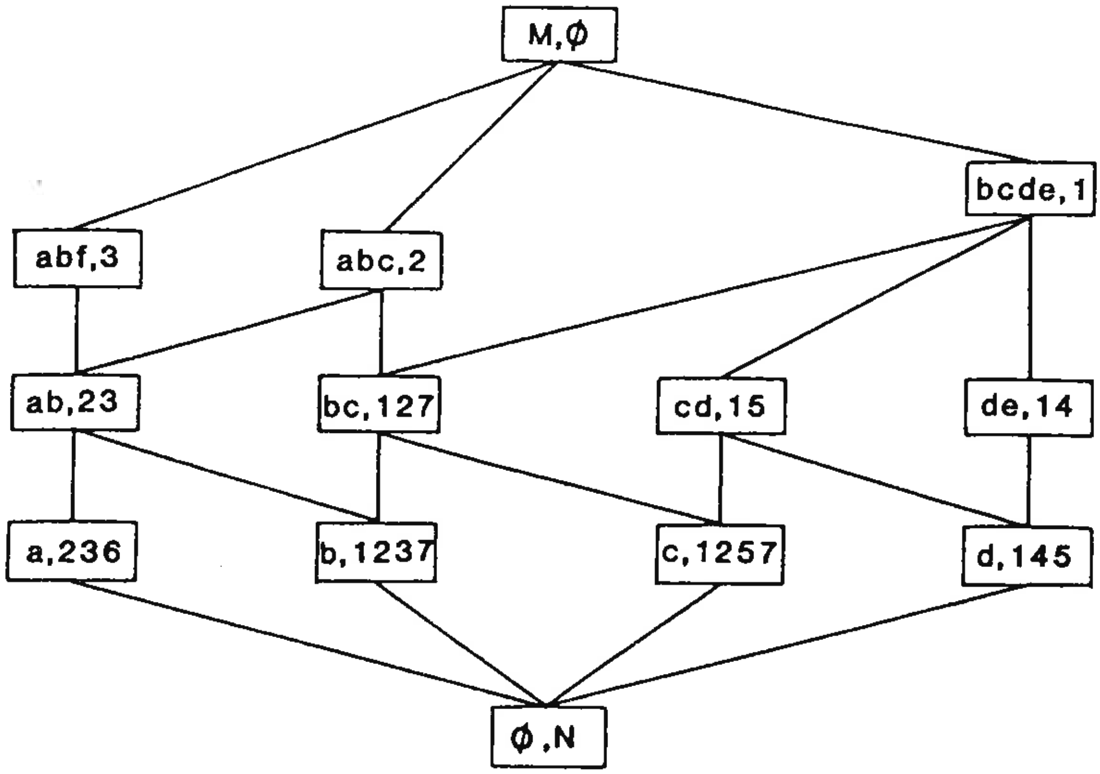


Figure 4b : Alternative configuration of the 'same' lattice

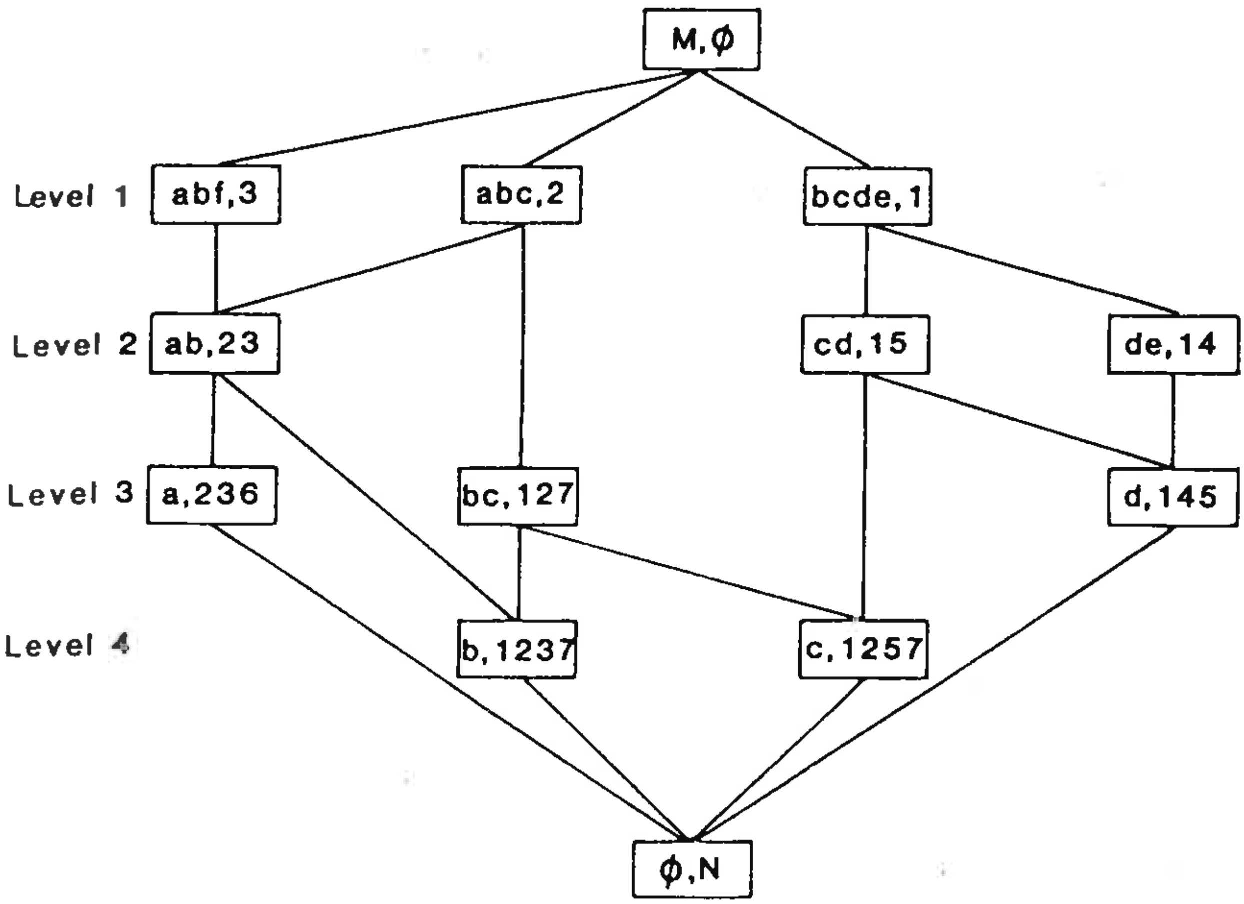


Figure 5 a : The simplicial complex derived from table 1 (Features of simplices, objects or vehicles)

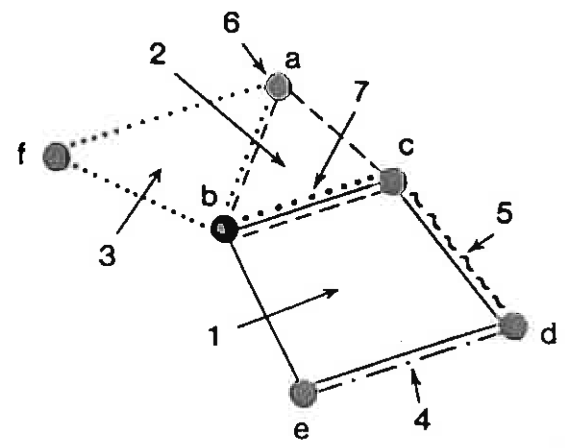


Figure 5 b : The conjugate simplicial complex

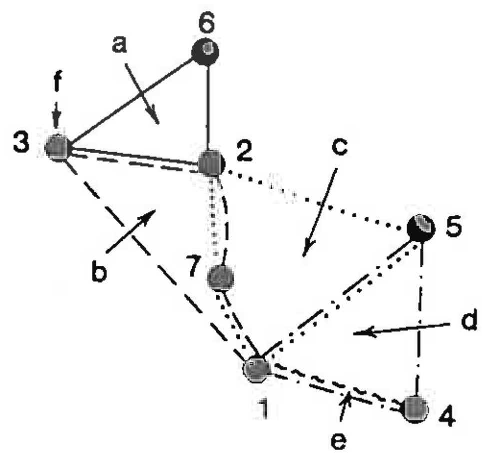


Figure 6 : Outputs from the Q-analysis algorithm

Figure 6a Listing of components at each dimensional level*

Level 4	(b) (c)
Level 3	(a) (bc) (d)
Level 2	(abcde)
Level 1	(abcdef)

Figure 6b Tree representation of Figure 6a.

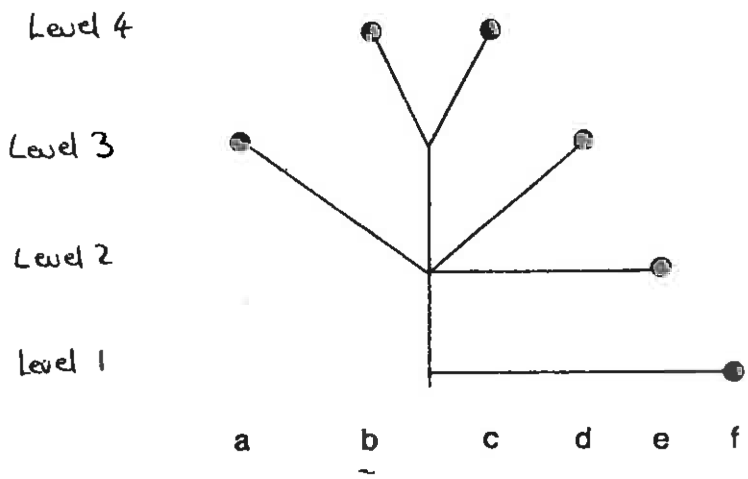
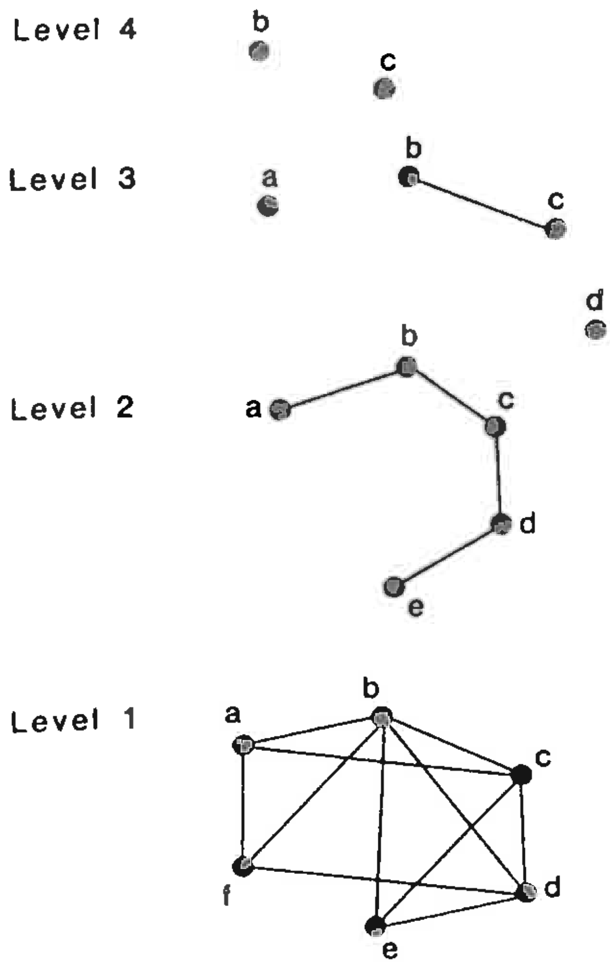


Figure 6c q-nearness graphs corresponding to Figures 6a and 6b.



* N.B. Level n refers to objects of dimension n-1