Social Networks 7 (1985) 203–224
North-Holland

NETWORK MACROSTRUCTURE MODELS FOR THE DAVIS–LEINHARDT SET OF EMPIRICAL SOCIOMATRICES *

Eugene C. JOHNSEN

University of California, Santa Barbara, and Harvard University **

Davis, Holland and Leinhardt have studied various microlevel and macrolevel models for group social structure in terms of certain two-valued relations on the group. These models are essentially defined in terms of permitted and forbidden triad types at the microlevel which imply and are implied by specific ordered clique structures at the macrolevel. Upon testing these models against a large collection of empirical sociomatrices, it was found that these models fit the data moderately well; however, not perfectly, since one or two triads forbidden by the models occur in over half of the sociomatrices at frequencies greater than or equal to chance expectation. These discrepancies have encouraged further theorizing and elaboration of the models, but there has heretofore been no macromodel put forth which fits the total data exactly. In this paper we obtain such a model, which allows hierarchy within cliques. In addition, we give an exact data-fitting macromodel for each class of group sizes as well, except for a somewhat less precise description for the class of largest sizes. In each case the general macromodel exhibits all of the permitted triads and none of the forbidden ones. Similar data from a study of Hallinan not only support the macromodel for the total data when the forbidden triads are taken individually but also when they are taken as a set, thus showing that the data actually support this particular macromodel and not another.

1. Introduction

In an extended and cumulative line of research, coming out of earlier substantive work of Heider (1946), Newcomb (1953, 1968), Homans (1950) and Davis (1963), various models have been proposed, studied and tested for the social macrostructure which is produced by an affect

* The main results of this paper were presented at the Third Annual Social Network Conference in San Diego, California, February 13, 1983. The author wishes to thank the members of the Analysis of Social Networks Seminar at UCSB and of the Mathematical Sociology Seminar at Harvard for their helpful discussions, with special thanks to Dorwin Cartwright, James Davis and Harrison White for their valuable comments and suggestions.
** Department of Mathematics, University of California, Santa Barbara, CA 93106, U.S.A.

relation (e.g., liking, friendship) which may exist between the members of a human group. Heider's social-psychological formulation of *balance* in triads, which was elaborated in a particular direction by Newcomb, was formalized and analyzed by Cartwright and Harary (1956) in graph-theoretic terms, which resulted in establishing a rigorous logical connection between the group's microstructure (the types of triads which can appear) and its macrostructure (the cliques which can appear and the relation between them) when the mathematical results are interpreted substantively. Taking their cue from Homans that social groups are structured by a combination of cliquing and ranking, Davis (1967, 1970), Davis and Leinhardt (1972) and Holland and Leinhardt (1971) progressively generalized the models for this macrostructure to *clustering, ranked clusters of cliques* and *transitivity* in order to accommodate the empirically observed departures of real social groups from the unduly restrictive macrostructure prescribed by the balance model. A colorful account of this development has been rendered by Davis (1979). A main conclusion of this work was that the set of permitted triads in the closest data-fitting model, namely transitivity, accounts for all except one of the triads which are relatively common in the data from a large collection of empirical sociomatrices (Davis 1970). Alternatively, all except one of the seven triads forbidden by this model are relatively rare in this data. As related by Davis (1979: 57):

> Six down, and one to go.
>
> But it would not go. That blasted 2-1-0 triad (two mutual positive relationships and one asymmetric) was clearly not rare (in fact, it tends to occur at higher than chance levels in our data bank matrices), and we have simply been unable to find a plausible ideal type group structure where 2-1-0s exist.

In this paper we account for the occurrence of the 210 triad and obtain the macromodel which fits the data for the total set of these sociomatrices exactly. In the macrostructure for the transitivity model the cliques are homogeneously linked internally. The surprise here is that this "errant" 210 is the essential triad needed to produce hierarchy within cliques, and in fact that is where it must occur. Thus, our macromodel accommodates differentiation (in the form of hierarchy) within cliques, a substantive structural property both needed and anticipated (Hallinan 1974). We also obtain the macromodels which exactly fit the data for each class of group sizes given by Davis (1970), except for a somewhat less precise characterization of the macromodel corresponding to the class of largest sizes. After a section of formal preliminaries, these results are presented in the following three sections. In the last section we relate our results to some aspects of the substantive theory and discuss the rather strong support for our main result given by the independent data of Hallinan (1974).

Since the relationship between the microstructure and macrostructure is precise and sometimes rather sensitive to modification (i.e. small variations in permitted microstructure may result in large changes in possible macrostructure), these results need to be carefully demonstrated; hence the presentation of full mathematical proofs. It may be noticed that our method of relating a group's macrostructure to its microstructure is quite general and can be applied to data for other dyadic relations besides affect.

2. Preliminaries

We shall formally represent a social group of n persons by a set of n vertices $V = \{v_1, v_2, \ldots, v_n\}$ and the two-valued relation R on the group by the two subsets D^+ and D^- of ordered pairs of distinct vertices from V, where every such pair (u, v) in $V \times V$ is labelled either $+$ or $-$, written $R(u, v) = +$ or $R(u, v) = -$ and drawn as a $+$ or $-$ signed arrow from u to v, where (u, v) is in D^+ if $R(u, v) = +$ and (u, v) is in D^- if $R(u, v) = -$. For the moment we do not specify R for ordered pairs on the same vertex (u, u). Note that D^+ and D^- are disjoint sets which together exhaust all the ordered pairs of distinct vertices in $V \times V$. Since D^- is determined by D^+ we can completely describe this structure by the loop-free digraph (V, D^+).

Let u, v be distinct vertices in V. If $R(u, v) = +$ and $R(v, u) = +$ then u and v are *positively* or *mutually* related (connected, linked) and we write uMv. If $R(u, v) = -$ and $R(v, u) = -$ then u and v are *negatively* or *null* related and we write uNv. Finally, if $R(u, v) = +$ and $R(v, u) = -$ then u is *asymmetrically* related to v and we write uAv. Note that if uMv then vMu and if uNv then vNu, but if uAv then it is not so that vAu. Thus M and N are symmetric relations (connections, links) and A is an asymmetric relation on ordered pairs of distinct vertices in V. We now extend M, N and A to full relations on V by taking as a convention that uMu and uAu but not uNu for all u

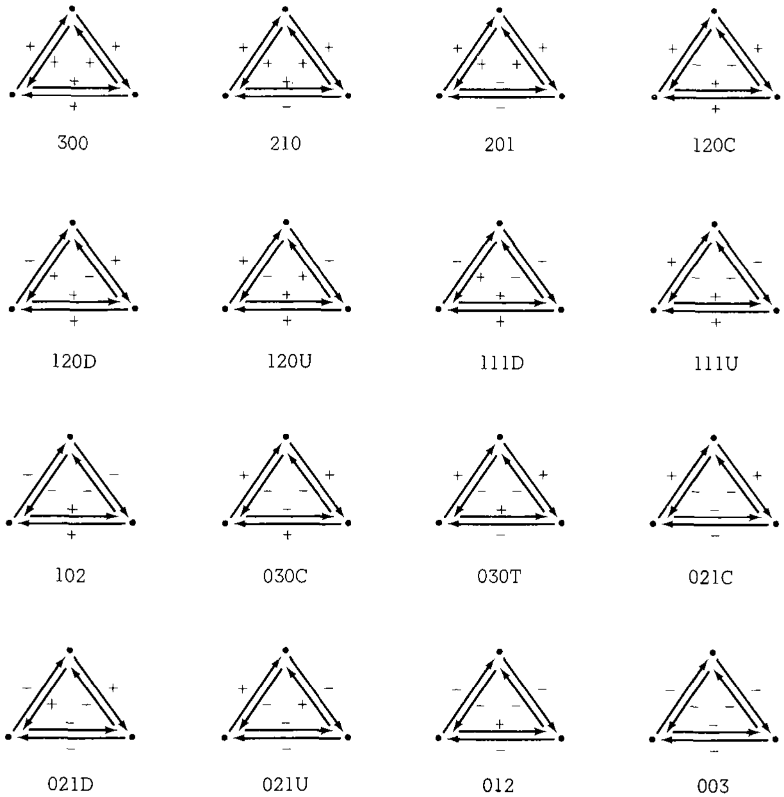

Fig. 1. The 16 triad types.

in V, which makes M and A but not N reflexive on V as well. Now, among a set of three distinct vertices from V there are 16 different combinations of M, A and N connections between the pairs of vertices in the set. Each combination is a *triad type* expressed as an ordered triple of nonnegative integers, $m:a:n$, where m, a and n are the numbers of M, A and N relations, respectively, and $m+a+n=3$, together with a special letter C, D, T or U standing for "cyclic", "down", "transitive", or "up". The set Θ of the 16 different triad types is given in Figure 1.

Now, various group structure models X can be defined in terms of the subset P_X of Θ, of all triads permitted to appear in the structure, and its complementary subset $P_X^c = \Theta - P_X$ of all triads forbidden to appear. Clearly, only P_X or P_X^c need be specified in order to define the

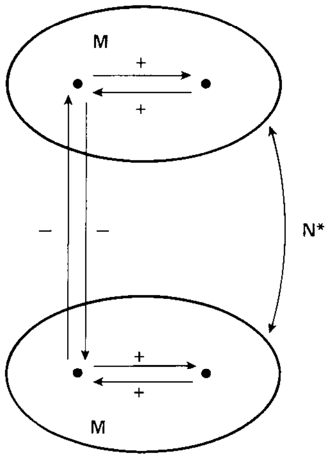

Fig. 2. Macrostructure for balance model: $P_{BA} = \{300, 102\}$. M stands for M-clique.

structure, and either one is called the *micromodel* corresponding to X. For an arbitrary subset P_X it may not be possible to exhibit all of the triads of P_X in a single general macrostructure. For example, $P_X = \{300, 003\}$ cannot produce a macrostructure exhibiting both triads 300 and 003 simultaneously. However, in the examples to follow and in the general macromodels to be developed here this situation does not arise.

Example 1. $P_{BA} = \{300, 102\}$. Here the A relation does not appear and the M and N connections can be viewed as single undirected signed edges. The resulting macrostructure is that of the *balance model* of Cartwright and Harary (1956) for a complete graph, consisting of at most two M-cliques (maximal subsets of vertices, pairwise connected by the M relation) which are related by N^* (i.e. completely interconnected by the N relation). See Figure 2.

Example 2. $P_{CL} = \{300, 102, 003\}$. As in Example 1 we view the M and N connections as single unidirected signed edges. The resulting macrostructure is that of the *clustering model* of Davis (1967) for a complete graph, consisting of any number of M-cliques which are pairwise related by N^*. See Figure 3.

Example 3. $P_{RC} = \{300, 102, 003, 120D, 120U, 030T, 021D, 021U\}$.

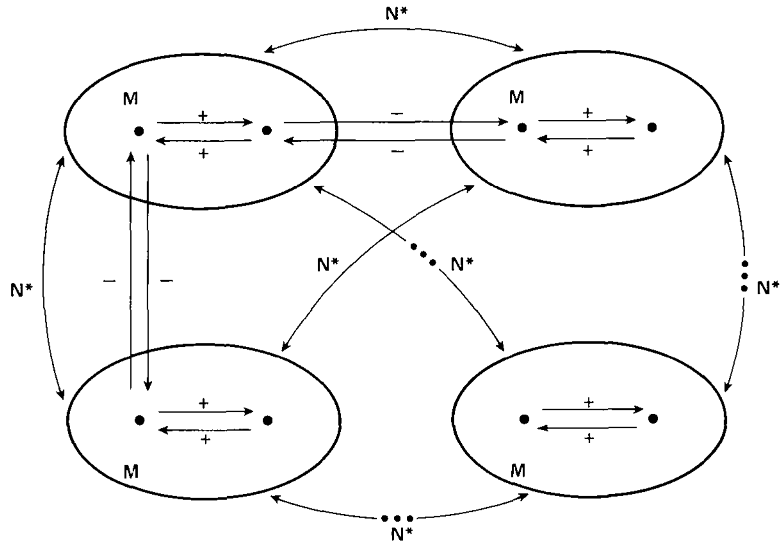

Fig. 3. Macrostructure for clustering model: $P_{CL} = \{300, 102, 003\}$. M stands for M-clique.

The resulting macrostructure is that of the *ranked clusters of M-cliques model* of Davis and Leinhardt (1972), consisting of a single hierarchy of clique levels, with cliques at the same level pairwise related by N^* and cliques at different levels pairwise related by A^* (i.e. completely interconnected by the A relation, all A relations going in the same direction) from the lower clique to the higher one. See Figure 4.

Example 4. $P_{TR} = \{300, 102, 003, 120D, 120U, 030T, 021D, 021U, 012\}$. The resulting macrostructure is that of the *transitivity model* of Holland and Leinhardt (1971), consisting of a collection of M-cliques partially ordered by the A^* relation (by convention, every M-clique is in relation A^* to itself) where incomparable M-cliques are pairwise related by N^*. See Figure 5.

3. Initial results

Davis (1970) tested the ranked clusters of M-cliques model against the empirical data from a collection of 742 sociomatrices (hereinafter called the Davis–Leinhardt data set). Using sampling distributions for triads

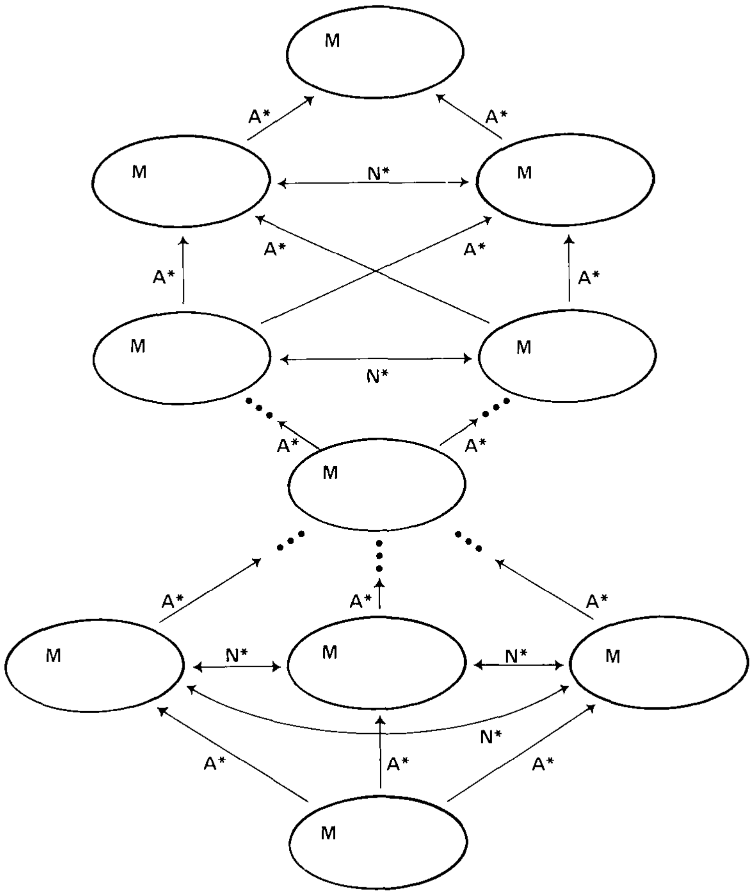

Fig. 4. Macrostructure for ranked clusters of M-cliques model: $P_{RC} = \{300, 102, 003, 120D, 120U, 030T, 021D, 021U\}$. M stands for M-clique; all interclique A^* relations implied by transitivity are suppressed.

in random graphs derived by Holland, he tabulated the percentages of sociomatrices in which each of the 16 triads occurred with frequency less than chance expectation for the total set of sociomatrices and for the subsets of sociomatrices corresponding to group sizes 8–13, 14–17, 18–22, 23–38 and 39–79 (*op. cit.*: Table 1). Now, for each set of sociomatrices an empirical model X is defined in terms of a set of permitted triads P_X consisting of all triads occurring with frequency less than chance expectation in (i) less than 50 percent, or (ii) less than

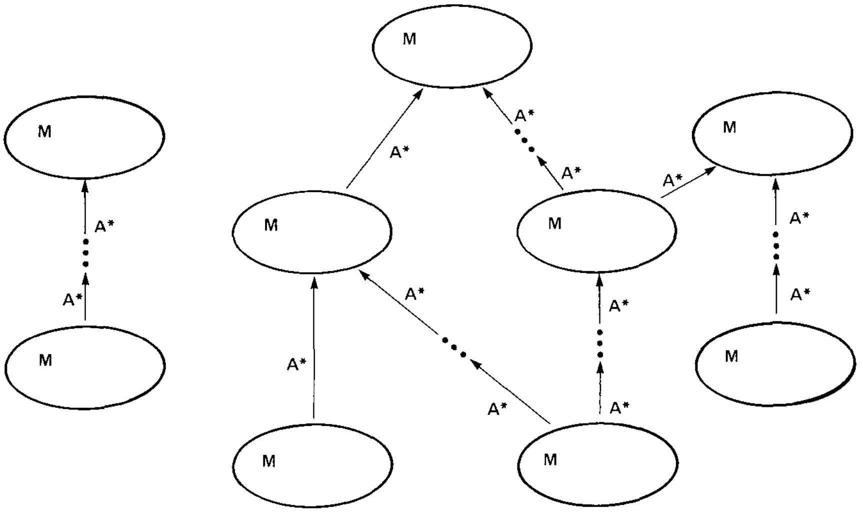

Fig. 5. Macrostructure for transitivity model: P_{TR} = {300, 102, 003, 120D, 120U, 030T, 021D, 021U, 012}. M stands for M-clique; all interclique A^* relations implied by transitivity are suppressed and all other missing interclique relations are N^*.

or equal to 50 percent, of the sociomatrices in the set, the remaining set of triads from the set of 16, P_X^c, being considered forbidden. In the data the distinction between criteria (i) and (ii) only occurs for group sizes 8–13 and triad 003; otherwise these criteria are the same.

A problem with Davis' tabulation is that the percentages for the triads 111D, 111U, 021D, 021U, 120D and 120U are give only for the aggregated pairs 111D and 111U, 021D and 021U, and 120D and 120U. Evidently, the percentages for the individual triads in these pairs are not readily obtainable (Samuel Leinhardt, private communication, May 10, 1983). Nevertheless, since none of the aggregated pair percentages is close to 50 percent (they are either ≤ 41 percent or else ≥ 60 percent over the five size classes, and either ≤ 27 percent or else ≥ 78 percent for the total set) and since similar independent evidence indicates that the triads in each aggregated pair tend to have frequencies falling on the same side of chance expectation (Holland and Leinhardt 1981: Table 8; Hallinan 1974; Table 3) we shall assume here that the percentages for these aggregated pairs are representative of the individual triads in the pairs.

The empirical micromodels to which the data conform are as follows.

A. Total set:
$P_A = \{300, 102, 003, 120D, 120U, 030T, 021D, 021U, 021, 210\}$.
B1. Sizes 8–13, criterion (i):
$P_{B1} = \{300, 102, 120D, 120U, 030T, 021D, 021U\}$.
B2. Sizes 8–13, criterion (ii):
$P_{B2} = \{300, 102, 003, 120D, 120U, 030T, 021D, 021U\}$.
C. Sizes 14–17:
$P_C = \{300, 102, 003, 120D, 120U, 030T, 021D, 021U, 012\}$.
D. Sizes 18–22:
$P_D = P_A$.
E. Sizes 23–38:
$P_E = P_A$.
F. Sizes 39–79:
$P_F = \{300, 102, 120D, 120U, 030T, 021D, 021U, 012, 210, 120C\}$.

Davis observed that the set of forbidden triads for the total set, $P_A^c = \{201, 030C, 111D, 111U, 021C, 120C\}$ does not match the set of forbidden triads (what he called "critical" triads) which define the proposed ranked clusters of M-cliques model $P_{RC}^c = \{201, 030C, 111D, 111U, 021C, 120C, 210, 012\}$, and on the basis of 4310 individual sociomatrix tests on these critical triads he noted about a 70 percent success rate for that model. It should be mentioned, however, that this model is fully supported by the data for group sizes 8–13 under criterion (ii), since $P_{B2} = P_{RC}$. This we state as the first result.

Theorem 3.1. Under criterion (ii) the empirical triad data for group sizes 8–13 in the Davis–Leinhardt data set conform exactly to the micromodel P_{RC} corresponding to the ranked clusters of M-cliques macromodel.

Since $P_{B1} = P_{RC} - \{003\}$ and the permitted triads in P_{RC} can all appear together in its macromodel, the macromodel corresponding to P_{B1} may be obtained from that for P_{RC} by suppressing those parts of the macrostructure which exhibit the 003 triad. This can easily be done. The 003 triad appears in the macromodel for P_{RC} (see Figure 4) only among vertices coming from three different M-cliques at the same level of the hierarchy. Thus, forbidding the 003 triad merely limits the number of N^*-related M-cliques at each level of the hierarchy to at most two and does not prevent any other triad of P_{RC} from appearing in the macromodel. The macromodel, then, looks like Figure 4 but with at most two M-cliques at each level. We call the resulting macromodel, corresponding to P_{B1}, the *ranked 2-clusters of M-cliques model.* We thus have the following result.

Theorem 3.2. Under criterion (i) the empirical triad data for group sizes 8–13 in the Davis-Leinhardt data set conform exactly to the micromodel corresponding to the ranked 2-clusters of M-cliques macromodel.

We also note here that the transitivity model is fully supported by the data for group sizes 14–17, since $P_C = P_{TR}$, which we state as the next result.

Theorem 3.3. The empirical triad data for group sizes 14–17 in the Davis–Leinhardt data set conform exactly to the micromodel P_{TR} corresponding to the transitivity macromodel.

It remains to characterize the formal macromodels corresponding to the empirical micromodels P_A for the total set of sociomatrices and P_F for group sizes 39–79 since, as observed above, the empirical micromodels for group sites 18–22 and 23–38 are identical to that for the total data set. We present an exact characterization for the total set in the next section and a rather comprehensive characterization for group sizes 39–79 in the section following.

4. The macromodel for the total set

For the total set of sociomatrices the empirical micromodel to which the data conform is given by the set of forbidden triads

$P_A^c = \{201, 030C, 111D, 111U, 021C, 120C\}$.

In order to determine the macrostructure corresponding to P_A^c we will need to consider a new relation on V.

Definition. The relation $\tilde{M}$, called *M-connectedness*, is defined on V as follows:

(a) $u\tilde{M}u$ for all u in V, and

(b) $u\tilde{M}v$ for $u \neq v$ in V if and only if for some $r \geq 2$ there exist vertices $x_1 = u$, x_2, ..., $x_r = v$ for which $x_i M x_{i+1}$ for all $i = 1, 2, \ldots, r-1$.

If $u\tilde{M}v$ we say that u is *M-connected* to v.

We note that $\tilde{M}$ is reflexive, symmetric and transitive, which means that $\tilde{M}$ is an equivalence relation on V. Hence V is partitioned into equivalence classes under $\tilde{M}$, called *$\tilde{M}$-cliques*, where every pair of vertices in the same $\tilde{M}$-clique are M-connected and every pair in different $\tilde{M}$-cliques are not M-connected.

Before obtaining the main theorem of this section we need the following results.

Lemma 4.1. If a group structure model X has the property that {201, 111D, 111U} $\subseteq P_X^c$, then for every pair of vertices u, v in an $\tilde{M}$-clique we must have either uMv, uAv or vAu.

Proof. Let W be an $\tilde{M}$-clique of the model X. If W has only one vertex then the lemma is true by definition. So let u and v be any two distinct vertices in W. Since $u\tilde{M}v$ there exist vertices $x_1 = u$, x_2, ..., $x_r = v$ in V such that $x_i M x_{i+1}$ for all $i = 1, 2, \ldots, r-1$ for some $r \geq 2$, and since $\tilde{M}$ is an equivalence relation all of these vertices are in W. Now, if $r = 2$ then uMv and we are done. Suppose $r = 3$. Then $x_1 M x_2$ and $x_2 M x_3$, and since triad 201 is forbidden we must have either $x_1 M x_3$, $x_1 A x_3$ or $x_3 A x_1$, and again we are done. Now suppose $r \geq 4$ and, inductively, assume that either $x_1 M x_{r-1}$, $x_1 A x_{r-1}$ or $x_{r-1} A x_1$. Then, since $x_{r-1} M x_r$ and since triads 201, 111D and 111U are forbidden we must have either $x_1 M x_r$, $x_1 A x_r$ or $x_r A x_1$, and we are again done. Thus, by induction on r, we have the lemma.

If W and W' are two $\tilde{M}$-cliques for which uAv, respectively uNv, for every pair of vertices u in W and v in W' then W is defined to be in the relation A^*, respectively N^*, to W' and we write WA^*W', respectively WN^*W'.

Lemma 4.2. If a group structure model X has the property that {111D, 111U, 120C} $\subseteq P_X^c$, then for every pair of $\tilde{M}$-cliques W and W' we must have either WA^*W', $W'A^*W$ or WN^*W'.

Proof. If W and W' each have only one vertex then the lemma is trivially true. So suppose one of the $\tilde{M}$-cliques, say W', has at least two vertices (the other $\tilde{M}$-clique W having one or more vertices). Let u and v be two vertices in W' and let s be a vertex in W. Now there exist r vertices $x_1 = u$, $x_2, \ldots, x_r = v$ in W' such that $x_i M x_{i+1}$ for all $i = 1, 2, \ldots, r-1$ for some $r \geq 2$. Suppose sAu. Since $x_1 M x_2$ and since triads 111D and 120C are forbidden we must have sAx_2. This argument can be continued inductively and we conclude that sAv. Thus, if sAu for some vertex u in W' then sAv for every other vertex v in W'. If W has only the single vertex s then WA^*W'. Now suppose W has another vertex t. Then there exist q vertices $y_1 = s$, $y_2, \ldots, y_q = t$ in W such that $y_j M y_{j+1}$ for all $j = 1, 2, \ldots, q-1$ for some $q \geq 2$. Since sAu and $y_1 M y_2$ and since the triads 111U and 120C are forbidden we must have $y_2 A u$. This argument can be continued inductively and we conclude that tAu. Then by the previous argument we have tAv. Thus sAu implies that tAv for every t in W and v in W' and again we have WA^*W'. Now if, instead of sAu, we assume that uAs, then a similar argument to the above shows that vAt for every v in W' and t in W which means that $W'A^*W$. Finally, if sNu then, since triads 111D and 111U are forbidden, we have inductively that sNx_i and $x_i M x_{i+1}$ imply sNx_{i+1} for $i = 1, 2, \ldots, r-1$ or sNv, and likewise $y_j Nu$ and $y_j M y_{j+1}$ imply $y_{j+1} Nu$ for $j = 1, 2, \ldots, q-1$ or tNu, whence tNv. Thus, sNu implies that tNv for every t in W and v in W' and we have WN^*W'. This proves the lemma.

Lemma 4.3. If a group structure model X has the property that $\{030C, 021C\} \subseteq P_X^c$, then for every three $\tilde{M}$-cliques W, W', W'', if WA^*W' and $W'A^*W''$ then WA^*W'', i.e., A^* is transitive on $\tilde{M}$-cliques.

Proof. Suppose for three $\tilde{M}$-cliques W, W' and W'' we have WA^*W' and $W'A^*W''$. Let s be any vertex in W and u any vertex in W'', and let t be some vertex in W'. Then sAt and tAu and, since the triads 030C and 021C are forbidden, we must have sAu. So, for every s in W and u in W'' we have sAu, which means that WA^*W''. Thus A^* is transitive on $\tilde{M}$-cliques.

Lemma 4.4.. If a group structure model X has the property that $\{030C, 021C, 120C\} \subseteq P_X^c$, then A is a transitive relation on V. Thus, under the relation A, V is a partially ordered set.

Proof. Let u, v and w be vertices in V where uAv and vAw. Then, since triads 030C, 021C and 120C are forbidden, we must have uAw. Thus A is transitive on V. Since A is also reflexive and antisymmetric, V is a partially ordered set under A.

Corollary 4.5 If a group structure model X is defined by $P_X^c = \{201,$ 030C, 111D, 111U, 021C, 120C$\}$ then
(a) under the relation A, V is a partially ordered set, and
(b) under the relation $\tilde{M}$, V is partitioned into $\tilde{M}$-cliques such that
  (i) every $\tilde{M}$-clique has only M and A links and is partially ordered with respect to A, and
  (ii) the set of all $\tilde{M}$-cliques form a partially ordered set under the relation A^* induced by A, where $\tilde{M}$-cliques which are incomparable with respect to A^* are in the relation N^* induced by N.

Proof. By Lemma 4.4 V is a partially ordered set under the relation A. Under the relation $\tilde{M}$, V is partitioned into $\tilde{M}$-cliques. By Lemmas 4.1 and 4.4 each $\tilde{M}$-clique has only M and A links and is partially ordered with respect to A. By Lemma 4.2 every two $\tilde{M}$-cliques are in either the A^* relation induced by A or the N^* relation induced by N. By Lemma 4.3 the relation A^* is transitive on the set of all $\tilde{M}$-cliques and, since A^* is also reflexive (by convention) and antisymmetric, the set of all $\tilde{M}$-cliques form a partially ordered set under this relation. As a result, $\tilde{M}$-cliques which are incomparable with respect to A^* are in the relation N^*.

The macrostructure of the model given by Corollary 4.5 is an amalgam of hierarchy and cliquing which we shall call the *hierarchical $\tilde{M}$-cliques model*. The set of permitted triads for this model is

$$P_{HC} = \{300, 102, 003, 120D, 120U, 030T, 021D, 021U, 012, 210\}.$$

This macrostructure is illustrated in Figure 6a where the structure internal to an $\tilde{M}$-clique is illustrated in Figures 6b and 6c. Note that each $\tilde{M}$-clique W contains a unique, possibly empty, M-subclique W_M consisting of all the vertices of W which are in the M relation to every other vertex of W. Since $P_A = P_D = P_E = P_{HC}$ we have the following main result.

Theorem 4.6. The empirical triad data for the total set and for group sizes 18–22 and 23–38 in the Davis–Leinhardt data set conform exactly to the micromodel P_{HC} corresponding to the hierarchical $\tilde{M}$-cliques macromodel.

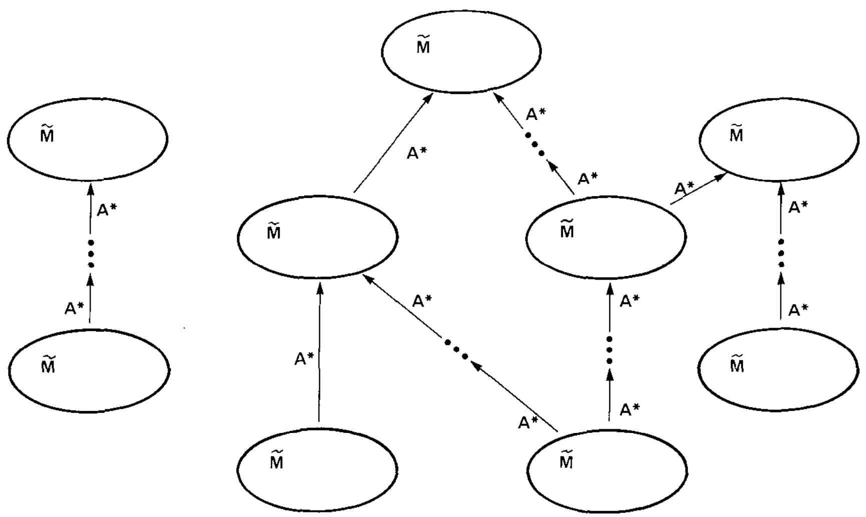

Fig. 6a. Macrostructure for hierarchical $\tilde{M}$-cliques model: P_{HC} = {300, 102, 003, 120D, 120U, 030T, 021D, 021U, 012, 210}. $\tilde{M}$ stands for $\tilde{M}$-clique; all interclique A^* relations implied by transitivity are suppressed and all other missing interclique relations are N^*.

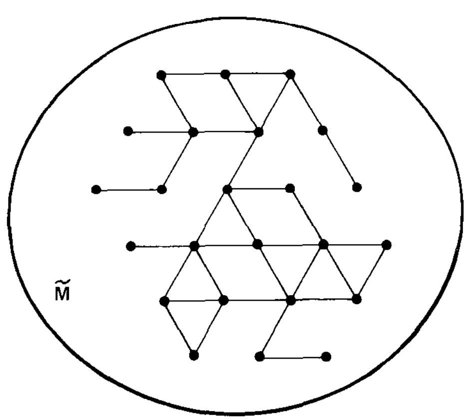

Fig. 6b. Hierarchical $\tilde{M}$-cliques model: M-connectivity of an $\tilde{M}$-clique. M links represented by unidirected edges; not all M links shown.

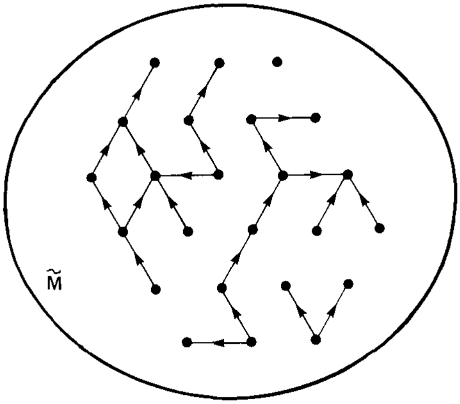

Fig. 6c. Hierarchical $\tilde{M}$-cliques model: Partial ordering under A of an $\tilde{M}$-clique. A links represented by directed edges; A links implied by transitivity are suppressed.

5. Group sizes 39–79

For group sizes 39–79 the empirical micromodel to which the data conform is given by the set of forbidden triads

$$P_{\mathrm{F}}^{\mathrm{c}} = \{201, 030\mathrm{C}, 111\mathrm{D}, 111\mathrm{U}, 021\mathrm{C}, 003\}.$$

Under the M-connectedness relation $\tilde{M}$ we again have V partitioned into $\tilde{M}$-cliques, and by Lemma 4.1 each $\tilde{M}$-clique has only M and A connections. Furthermore, by Lemma 4.3, A^* is transitive on the set of $\tilde{M}$-cliques, which means, as before, that the set of all $\tilde{M}$-cliques forms a partially ordered set under the relation A^*. However, because triad 120C is now permitted, the relation A may not be transitive within $\tilde{M}$-cliques and so neither V nor any $\tilde{M}$-clique need be partially ordered by A. In addition, the triad 120C may be present between pairs of $\tilde{M}$-cliques, which means that if two $\tilde{M}$-cliques are incomparable with respect to A^* they need not be related by N^*. Nevertheless, since triads 111D and 111U are forbidden, if any link between two $\tilde{M}$-cliques is an A relation then every link between them is an A relation. Thus, every two $\tilde{M}$-cliques W and W' are either completely A-linked (i.e. for every pair of vertices u in W and v in W' either uAv or vAu) or are completely N-linked (i.e. WN^*W').

If $\tilde{M}$-cliques W and W' are completely A-linked we define them to be in the relation $A^{\#}$ and write $WA^{\#}W'$. Note that $A^{\#}$ is a symmetric relation between $\tilde{M}$-cliques and that if $WA^{*}W'$ then $WA^{\#}W'$. Since $WA^{*}W$ for all $\tilde{M}$-cliques W, this implies that $A^{\#}$ is reflexive on $\tilde{M}$-cliques. We define the relation $\tilde{A}^{\#}$, called $A^{\#}$-*connectedness*, on the set of all $\tilde{M}$-cliques by stipulating that $W\tilde{A}^{\#}W$ for all $\tilde{M}$-cliques W, and that $W\tilde{A}^{\#}Y$ for two $\tilde{M}$-cliques W and Y if and only if for some $r \geq 2$ there are $\tilde{M}$-cliques $\tilde{M}_1 = W$, $\tilde{M}_2$, $\ldots$, $\tilde{M}_r = Y$ for which $\tilde{M}_i A^{\#} \tilde{M}_{i+1}$ for all $i = 1, 2, \ldots, r-1$. If $W\tilde{A}^{\#}Y$ we say that W is $A^{\#}$-*connected* to Y. We note that $\tilde{A}^{\#}$ is reflexive, symmetric and transitive, which means that $\tilde{A}^{\#}$ is an equivalence relation on the set $\{\tilde{M}\}$ of $\tilde{M}$-cliques. Hence $\{\tilde{M}\}$ is partitioned into equivalence classes under $\tilde{A}^{\#}$, called $\tilde{A}^{\#}$-*superclusters* or simply *superclusters*, where every pair of $\tilde{M}$-cliques in the same supercluster are $A^{\#}$-connected and every pair in different superclusters are in the relation N^{*}.

Now, because the triad 003 is forbidden, the macrostructure here cannot consist of more than two superclusters, and if it consists of two superclusters then every interclique link within each supercluster must be $A^{\#}$ (there cannot be any N^{*} links). Only in the case of one supercluster is it possible for two $\tilde{M}$-cliques to be related by N^{*}. In this case, because triad 003 is forbidden, the $A^{\#}$-distance between any two $\tilde{M}$-cliques in the supercluster must be at most 3. In summary, then, we have the following result.

Theorem 5.1. The empirical triad data for group sizes 39–79 in the Davis–Leinhardt data set conform exactly to the micromodel which defines a macromodel consisting of either

(a) two superclusters of $\tilde{M}$-cliques, where each supercluster is a complete graph on its $\tilde{M}$-cliques with respect to the $A^{\#}$ relation and the $\tilde{M}$-cliques of each supercluster are partially ordered with respect to the A^{*} relation, or else

(b) one supercluster of $\tilde{M}$-cliques, of diameter at most 3 with respect to the $A^{\#}$-distance between $\tilde{M}$-cliques, in which the $\tilde{M}$-cliques are partially ordered with respect to the A* relation.

The macrostructure for this micromodel is illustrated in Figure 7, where the internal microstructure of the $\tilde{M}$-cliques is like that shown in Figures 6b and 6c, except that A need not be transitive within the $\tilde{M}$-cliques.

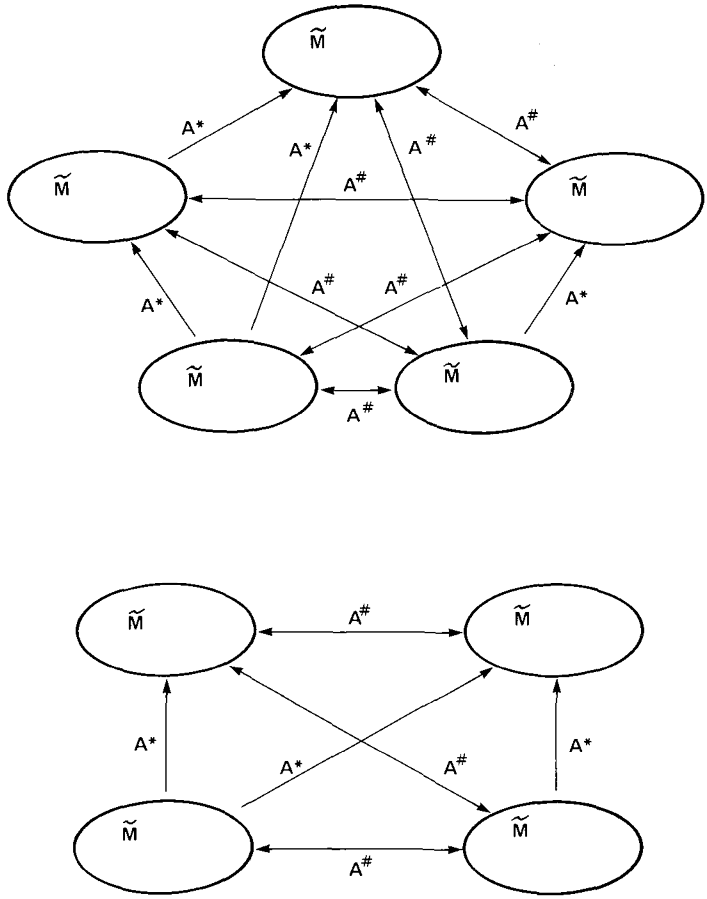

Fig. 7. Macrostructure for group sizes 39–79: $P_F = \{300, 102, 120D, 120U, 030T, 021D, 021U, 012, 210, 120C\}$. $\tilde{M}$ stands for $\tilde{M}$-clique; all interclique relations between superclusters are N^*; when only one supercluster, some interclique relations may be N^*.

6. Discussion

The underlying aim of this paper is to present an analytic method for relating the microlevel and macrolevel structures of a social group under a two-valued relation R. The method enables us to obtain an exact characterization of the macrostructure determined by the micromodel of permitted and forbidden triads specified by the full

Davis–Leinhardt data set, an unsolved problem in social structure remaining from previous research in the balance theoretic tradition (Davis 1979). It also enables us to characterize the macrostructures for the micromodels given by the five different size classes specified by Davis in this set of sociomatrices. In a statistical sense the micromodels for group sizes 8–13 and 14–17 are rather close to that for the total set in that each of the triads which distinguish them, namely 012, 003 and 210, occur with frequency less than chance expectation in almost exactly half or less of the sociomatrices. Thus, a small perturbation in the data could change these micromodels to the micromodel for the total set.

In line with this view, we note the results of Hallinan (1974). Her data from a study of 51 school classes (*op. cit.*: Table 4) specify the micromodel P^{c}_{HC} and hence determine the hierarchical $\tilde{M}$-cliques macromodel for each class of group sizes 8–17, 18–22, 23–38 as well as the total set, without exception. Hallinan's data was collected so as to avoid some of the major problems associated with the Davis–Leinhardt set of sociomatrices, in particular the problems of heterogeneity of type of group and type of sociometric relation, and of structural biases in data obtained by the use of the fixed choice method. Thus, if there is a structure determined by an affect relation within a group, we would expect her data to give a clearer picture of it. We note that when the Davis–Leinhardt sociomatrices for group sizes 8–13 and 14–17 are aggregated the percentage of sociomatrices with frequency of a given triad less than chance expectation is just the weighted average of the corresponding percentages for the two classes of group sizes and we obtain for the critical triads

$$\{109(51\%) + 198(36\%)\}/(109 + 198) = 41\% \text{ for } 012,$$

$$\{103(50\%) + 169(27\%)\}/(103 + 169) = 36\% \text{ for } 003,$$

but

$$\{67(51\%) + 172(51\%)\}/(67 + 172) = 51\% \text{ for } 210.$$

Thus, the micromodel specified by the aggregated Davis–Leinhardt data for group sizes 8–17 corresponds to the transitivity model, which differs from the hierarchical $\tilde{M}$-cliques model, supported by the Hallinan data, only in forbidding the triad 210.

A further observation should be noted here. Hallinan's data specifies the micromodel P^c_{HC} not merely for each of the six forbidden triads separately but also for all six jointly. From her data in Table 4 we can conclude, by minimizing the common overlap of the six subsets of sociomatrices which exhibit the relative rarity of the six triads in P^c_{HC}, that the percentage of sociomatrices which simultaneously exhibit all six triad frequencies with less than chance expectation is bounded below by

$$100-(100-\%030C)-(100-\%201)-(100-\%021C)$$
$$-(100-\%111D)-(100-\%111U)-(100-\%120C)$$

or

$\{100-0-0-0-10-10-10\}\% = 70\%$ for group sizes 8–17,

$\{100-0-0-10-10-10-15\}\% = 55\%$ for group sizes 18–22,

$\{100-6-0-0-2-2-14\}\% = 76\%$ for group sizes 23–38,

and

$\{100-2-0-4-7-7-14\}\% = 66\%$ for the total set.

This indicates that for each size class in Hallinan's data, a single majority of the sociomatrices specifies the micromodel P^c_{HC} as a total set of forbidden triads, rather than separate majorities of sociomatrices specifying the forbidden triads of P^c_{HC} individually. This is rather strong support for the proposition that the structure of a social group under an affect relation tends toward the full micromodel P^c_{HC} and its corresponding macromodel of hierarchical $\tilde{M}$-cliques (at least for school groups in the approximate size range of 10–40 members).

As shown in the previous section, the micromodel P^c_{HC} does not correspond to the Davis–Leinhardt data for group sizes 39–79. Unfortunately. Hallinan's study did not have any school classes in this size range. In order to transform the micromodel P^c_F into P^c_{HC} we would need to permit triad 003 and forbid triad 120C; however, the percentage of sociomatrices in which these triads occur with less than change expectation in the Davis–Leinhardt data are, respectively, 58 and 31 percent, neither of which is really close to 50 percent. Thus, either there is a rather sizeable error in the Davis–Leinhardt data for these group sizes or else the microstructure and macrostructure of such large groups are really different. Invoking the former puts us in the position of treating this data selectively, and perhaps two reasons for doing so are that the data base for group sizes 39–79 is somewhat smaller than that for each of the other classes of group sizes (Davis 1970: Table 1), and (along with group sizes 8–13) appears to be considerably less representative of the total set than are the other classes (extrapolating from the sample of 60 sociomatrices studied by Davis and Leinhardt (1972)). The latter alternative, however, has a certain appeal since it seems intuitively reasonable that in a group of size 39 or larger it would be difficult for a member to maintain a significant relation with any sizeable proportion of the group because of the effect of "saturation" on personal interaction. Thus, the appearance of a different type of microstructure and corresponding macrostructure should not be surprising.

The occurrence of a *proper* $A^{\#}$ relation (i.e. which is not A^*) between $\tilde{M}$-cliques in the macrostructure for group sizes 39–79 is of interest from another point of view. Such a connection between two $\tilde{M}$-cliques indicates that there is some friendship of each clique for the other and not a total lack of friendship of total presence of hostility. The accommodation of friendly cliques in a larger social structure has been developed by Boyle (1969), although not in the form given by Theorem 5.1.

As a final note, we mention that the hierarchical $\tilde{M}$-cliques model not only solves the macrostructure problem for the Davis–Leinhardt data but also (i) does so in a manner somewhat anticipated by Davis and Leinhardt, and (ii) in the process answers the substantive need for accommodating *intra*clique structure (cf. Hallinan 1974: 19–20 and Ch. VII, but see also Hunter (1975) regarding her treatment of this issue). Davis and Leinhardt state in their summary (1972: 250):

> Two of the seven specific hypotheses [that the 012 and 210 triads are relatively rare] have equivocal or negative outcomes. Whether this is random variation or a substantive flaw in the model is unclear. If it is a substantive flaw, the weakness seems to lie not in the idea of ranking or the idea of cliques, but in the assumptions about how the ranking system and the clique system are articulated.

In the hierarchical $\tilde{M}$-cliques model we can view this articulation in dual ways, depending on whether M or A connections set the pattern:

(a) Within the group M-connections form cliques which partition the group. Within the cliques and between members of different cliques a ranking relation A forms which partially orders each clique and partially orders the set of cliques (by inducing the interclique A^* relation) in a consistent manner so as to partially order the entire group.
(b) Within the group a ranking relation A forms which partially orders the entire group. Within the partial ordering M-connections form cliques which partition the group in a consistent manner so as to preserve the A relation between corresponding members of different cliques (thereby creating the interclique A^* relation).

It is possible, of course, that in the process of group structure formation both types of connections jointly set the pattern, each constrained only by consistency with respect to the currently evolved structure. In any such articulation we see that the cliques are internally differentiated by a hierarchy and that each intraclique hierarchy blends into the overall hierarchy of the group.

References

Boyle, Richard P.
1969 "Algebraic systems for normal and hierarchical sociograms". *Sociometry 32*: 99–119.
Cartwright, Dorwin and Frank Harary
1956 "Structure balance: A generalization of Heider's theory". *Psychological Review 63*: 277–293.
Davis, James A.
1963 "Structural balance, mechanical solidarity, and interpersonal relations". *American Journal of Sociology 68*: 444–462.
1967 "Clustering and structural balance in graphs". *Human Relations 20*: 181–187.
1970 "Clustering and hierarchy in interpersonal relations: Testing two graph theoretical models on 742 sociomatrices". *American Sociological Review 35*: 843–851.
1979 "The Davis/Holland/Leinhardt studies: An overview". In Paul W. Holland and Samuel Leinhardt (eds.), *Perspectives on Social Network Research*, New York: Academic Press. pp 51–62.
Davis, James A. and Samuel Leinhardt
1972 "The structure of positive interpersonal relations in small groups". In Joseph Berger et al. (eds.), *Sociological Theories in Progress Vol. 2*, Boston: Houghton Mifflin. pp. 218–251.
Hallinan, Maureen T.
1974 *The Structure of Positive Sentiment*, New York: Elsevier Scientific.

Heider, Fritz
1946 "Attitudes and cognitive organization". *Journal of Psychology 21*: 107–112.
Holland, Paul W. and Samuel Leinhardt
1971 "Transitivity in structural models of small groups". *Comparative Group Studies 2*: 107–124.
1981 "An exponential family of probability distributions for directed graphs". *Journal of the American Statistical Association 76*: 33–50.
Homans, George C.
1950 *The Human Group*. New York: Harcourt, Brace and World.
Hunter, John E.
1975 Review of *The Structure of Positive Sentiment*, by Maureen T. Hallinan, New York: Elsevier Scientific, 1974, in *Contemporary Sociology 4*: 504–507.
Newcomb, Theodore M.
1953 "An approach to the study of communicative acts". *Psychological Review 60*: 393–404.
1968 "Interpersonal balance". In Robert P. Abelson *et al.* (eds.), *Theories of Cognitive Consistency: A Sourcebook*. Chicago: Rand-McNally. pp. 28–51.