AD-785 268

THE CAT!J TECHNIQUE: SOME DESCRIPTIVE TESTS OF ITS ADEQUACY

H. Russell Bernard, et al

West Virginia University

Prepared for:

Office of Naval Research

10 September 1974

AD785268

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

AD785 268

REPORT DOCUMENTATION PAGE		READ INSTRUCTIONS BEFORE COMPLETING FORM
1. REPORT NUMBER BK-104-74	2. GOVT ACCESSION NO.	3. RECIPIENT'S CATALOG NUMBER
4. TITLE (and Subtitle) THE CATIJ TECHNIQUE: Some Descriptive Tests of Its Adequacy		5. TYPE OF REPORT & PERIOD COVERED INTERIM
		6. PERFORMING ORG. REPORT NUMBER
7. AUTHOR(s) Peter D. Killworth (Cambridge University) H. Russell Bernard (West Va. University)		8. CONTRACT OR GRANT NUMBER(s) N00014-73-A-0417-0001 177-949
9. PERFORMING ORGANIZATION NAME AND ADDRESS H. Russell Bernard Department of Soc/Anthro West Virginia Univ. - Morgantown, W. Va. 26506		10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS
11. CONTROLLING OFFICE NAME AND ADDRESS ONR - Code 452 Arlington, Virginia 22217		12. REPORT DATE September 10, 1974
		13. NUMBER OF PAGES 36
14. MONITORING AGENCY NAME & ADDRESS (if different from Controlling Office) West Virginia University Morgantown, W. Va. 26506		15. SECURITY CLASS. (of this report) UNCLASSIFIED
		15a. DECLASSIFICATION/DOWNGRADING SCHEDULE

16. DISTRIBUTION STATEMENT (of this Report)

Approved for public release, distribution unlimited.

17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report)

DDC

18. SUPPLEMENTARY NOTES

19. KEY WORDS (Continue on reverse side if necessary and identify by block number)

sociometric, group structure

20. ABSTRACT (Continue on reverse side if necessary and identify by block number)

The catij matrix for describing group structure is critically examined. Comparisons are made with direct factoring of distance matrices and with traditional clique-finding techniques in sociometry.

DD FORM 1 JAN 73 1473 EDITION OF 1 NOV 65 IS OBSOLETE
S/N 0102-014-6601

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

ONR Code 452
Contract #N00014-73-A-0417-0001

Report #BK-104-74

THE CATIJ TECHNIQUE:
Some Descriptive Tests
Of Its Adequacy

Peter Killworth
Department of Applied Math
Cambridge University
Cambridge, England

H. Russell Bernard
Department of Soc/Anthro
West Virginia University
Morgantown, W. Va. 26506

Research supported by the Office of Naval Research, Organizational Effectiveness Division (Code 452), and the Federal Bureau of Prisons.

This paper was delivered at the meetings of the American Sociological Association, Montreal, Canada, August 27, 1974.

ia

Recently we have been working on some aspects of sociometry which we believe require attention. First of all, it is painfully obvious that traditional sociometric techniques are neither "analyses" nor even adequate descriptions of the structure of human groups. The sociomatrix and the sociogram are clearly no more or less than the original sociometric data, molded into an impressive shape, subject to all the measurement error one imagines one gets by asking people to list the top 3 or 4 people they trust, or love, or want to work with. Beyond this, sociomatrices are typically manipulated using clique finding algorithms to show chunks of people who all like or trust or admire one another in a way which makes them a clique by some definition. Clearly, it is the definition of a clique which one uses which forces particular chunks to appear. In 1946 Forsythe and Katz recommended the factoring of sociomatrices as a chunk finding method, but this never really caught on. (See Macrae 1960; Wright, 1961)

Second, sociometric data is virtually all affective; the history of the field is rooted in the psychiatric tradition of Moreno, and social psychology in general has dealt with affective relations: trust, admiration, fear, respect, love or what have you.

Third, whatever a sociogram says about some group at time t, it says nothing about change in the group. Dynamic sociometric studies are rare indeed, and virtually nothing exists on <u>how</u> structures change over time. A very recent model by Jack Hunter at MSU is a welcome step in this direction (Hunter, in press).

Our work has, for the past 2 and 1/2 years, addressed these three areas of concern. The catij device was developed to collect data of greater richness and less measurement error than traditional sociograms. At this stage we can only say that we hope we have been successful.

In our original article on the catij techniques we demonstrated that artificially induced perturbations in the data were largely filtered out within catij matrix (Bernard and Killworth, 1973). However, we are aware that catij may be measuring something different from the sociogram, and thus we must reserve judgment on the extent to which it removes error from the original sociomatrix. A study is under way to examine the accuracy of the sociogram versus catij in defining communication networks. This involves an elite deaf community in Washington who communicate via teletypes. After taking sociometric and catij data, we are logging their communications via the carbon sheets on the teletype and comparing actual communication patterns with those predicted from the catij data.

We will assume 2 things here: 1) Anyone interested in the mathematical proof that catij does filter perturbations from data and that if a significant matrix, different from a sociomatrix, will consult our earlier work, (1973). 2) Anyone interested in the use of catij for applied work will consult the paper under discussion at this session.[1] We will address ourselves at this time to the adequacy and power of the techniques compared with earlier sociometrics. First, a definition of the catij matrix:

> data is obtained in the form of a "distance matrix" dij whose ij entry is the ranking by i of j. This ranking runs from o for i himself to N-1 for the person he ranks lowest, according to the question asked.

---

[1]The paper under discussion was "Catij, a new sociometric and its application to a prison living unit." It appeared first as Report BK-102-73 under this ONR contract and it will appear in Human Organization, winter, 1974.

The catij entries are obtained by seeking minimal distance routes from person to person using intermediaries. An entry of 1 in catij indicates that there was no shorter route from i to j than the direct distance dij. An entry of 2 indicates that there was one intermediary involved in the fastest route; and an entry of n indicates that (n-1) intermediaries were used. We believe (for effective networks) that a person's catij-1 entries represent his perceived direct communicants, his 2's as less direct, and so on.

The data are processed using the KBPAK, described in Report BK-101-73 under this contract, and available on tape from the authors. The output of the KBPAK is as follows: a distance matrix, a minimal distance matrix, and the catij matrix, followed by a series of the usual statistics; mutual links (i.e., catij 1's), number of catij 1 links to and from everybody (a crude "popularity" or centrality measure). Then follows, purely for comprehension purposes, a factoring of the catij matrix using a Varimax rotation. Experimentation shows the Varimax rotation to give the most intuitively useful results in this technique. Stingily, people with at least a 0.6 factor loading are chosen for inclusion in a "group." Catij 1 links are then used to show the ties between groups (and their members with all other groups and their members).

"Groups" are defined as people who have similar views on their relationships to the universe. This may result because of cliqueing or not and in general ethnographic evidence must be relied upon to decide the meaning of underlying factors.

Figure 1 is a graphic representation of an entire female living unit at Kennedy Youth Center, a coeducational, federal prison in Morgantown, West Virginia. The unit is called Cottage C. Figure 2 shows the groups in Cottage C and their connections with other groups. Brokers provide links between groups where there are none. A recent article by Killworth (1974) has demonstrated that, in general, groups of order size 40-80 will have some brokers who: a) do not belong to a group and b) do not know each other.

Figure 1 and 2 Here

These data, which show strong connectivity in the group and virtually no isolates, are rather different from what might be expected from more traditional sociometrics. Several questions have been raised regarding the handling of our data, and the manipulation of sociometric data in general. 1) Is factoring a reasonable thing to do to a sociogram or to a catij matrix? 2) Is a clique-finding (using traditional techniques) a reasonable thing to do to a sociogram or a catij matrix? 3) To what extent do the sociogram and catij yield similar results? We are aware that others have been concerned with finding the best way to treat sociometric data. Since we believe catij to be something rather different from a traditional sociometric we have tried to test its power and usefulness in the field. Also, we performed some direct comparative experiments between catij and earlier sociometrics. Here are the results. For consistency, Cottage C data is used throughout.

1) FACTORING

Consider first the results of a sociogram concocted from our data (a 100% ranked sociogram) by selecting the first k choices, where k is an integer. All other rankings are construed as "no choice". Factoring the resulting matrix (we always factor by rows, for reasons explained in Killworth and Bernard, 1974), we obtained some groups which made no sense from any perspective: sociogram, catij, or ethnography.

If instead of placing 1's in the sociogram for "choice" and zeros for no choice, we replace the zeros by a large number (51 = N-1 in this case), then we obtain results somewhat less abysmal, but nothing to write home about. There are more catij links and more reciprocated sociogram links between "group" members than in the previous example, but the groups still look totally unlike any plausible definition of a group.

If we set the number of choices up to 7, the results are more plausible, the more so if zeros are replaced by N-1. For example, the whole of the case management group (i.e. counseling staff in the cottage) were factored into one clump of people. However, the failing of this case and also when we factored the entire matrix dij of inter-personal rankings was that the brokers between cliques (all and only the butch homosexuals, incidentally) were placed in a large "clique." This is counter ethnographic; counter to the insight of the case management staff of Cottage C; and counter to theoretical expectations as noted above (cf. Killworth, 1974). As it turns out, three dominant aggressive homosexuals maintained their own groups ("Harems") and acted in competition with one another rather than as a clique in the ordinary sense of the term.

Figure 3 Here

What all of this says, of course, is that the larger the number of choices, the more rich (ethnographically and numerically) are the results. Why should one throw away data, after all? We have been told again and again in conversation with sociometricians that limited choice sociograms are a thing of the past; unfortunately all available data indicate that this is not true.

Now, the question is: is factoring a reasonable technique for treating sociograms? The answer appears to be, on most effective, limited choice sociograms, no. Various other factoring techniques have been used (McRae, 1960; Hubbell, 1965) but have not been overly successful. Even in the latter cases considered here, very poor sociogram connectivities within the predicted group is the rule; also much important ethnographic detail is lost by the use of the sociogram. For example, catij turned up a strong bond between a staff member and an inmate which was not found using basic sociometric tests, but which was a conspicuous and known factor within the group because of its rarity. In another case, 3 inmates turned up as a group where no apparent social links were discernible. This bothered us until we found, a week later, that the three had attempted to escape together.

2) CLIQUE-FINDING AS A METHOD OF COMPREHENSION OF THE SOCIOGRAM

The most telling point suggesting that catij may be superior to a simple sociogram comes from comparing the factoring of catij with clique finding techniques on sociograms concocted as already described. For this experiment we chose one of the most commonly accepted clique finding algorithms, given by Harary and Ross in 1956, and the definition of a clique given by Harary in countless articles: for our purposes, it is isomorphic to the complete graph of order n.

A general description of the results is simple: Brokers are indicated weakly, and they are not the brokers found by 1) factoring catij; 2) taking intuitive evidence from members of the staff of C cottage; 3) taking ethnographic evidence by observation and by interviewing inmates in C cottage. Again, we must conclude that clique finding is all right in a pinch, so long as one has data for at least 7 choices. However, we are neither pinched nor forced to use only 7 choices.

The results are as follows: For four choices there are 10 3-cliques as shown.

Figure 4 Here

There is no connectivity among them, however. We have noted elsewhere the counter-intuitive notion of a group of 52 people split into unconnected groups. At the risk of repetition, we emphasize that these anomalous results (isolates) occur because of the intrinsic nature of the forced-choice sociogram and the manipulations performed on them.

Things get better, of course, as the number of choices increases. For seven choices, all but 2 of the 52 people are in a clique, brokering between cliques, or hanging on to a clique. However, the network produced by the algorithm, even for seven choices, is very weakly connected. (A version of this network will be shown in figs. 5-18).

We agree with Alba (1973), who has recently given a better graph-theoretic definition of a clique, that the completely connected definition frequently used in the literature is counter-intuitive and counter-productive. We have not had the opportunity to try later algorithms on our data, specifically algorithms which conform to Alba's definition. The important point is that the cliques we derive look like well defined sub-groups, have strong connectivity between them and are intuitively more useful than chunks produced by factoring dij at any level of choice.

3) A COMPARISON BETWEEN CATIJ AND THE SOCIOGRAM

We now consider the connectivity of the entire group if we use limited choice sociograms compared with catij. Consider the following 14 diagrams, figs. 5-18, produced by a Calcomp.

Figure 5-18 Here

Each sociogram has been produced from our data by setting the number of choices to some integer K = 1, 2, ..., 7 and all other choices as well. The first seven show the network for reciprocated choices (i.e. a link is drawn between a and b only if a chooses b and b chooses a), and the second group shows the same but with unreciprocated choices (a link is drawn if either of a and b chooses the other). Not all individuals are shown: we have illustrated only those used in the catij diagram for Cottage C shown in figure 1. The placements of the individuals correspond to that diagram also; they are not "ideal" placements, in terms of reducing the number of crossovers on the graph, but the best we could achieve for catij. However, the placements make the sociograms look more complex than they really are. For example, the connectivity for the seven-choice reciprocated sociogram is very weak.

Note that non-reciprocated links are far more complex than reciprocated links, as one might expect. In general, however, reciprocity is important only in affective relations. For effective links, reciprocity is important only by its irrelevance. (Patently, in terms of time, I talk to you as long as you talk to me.) This raises the issue of affective vs. effective structures in groups. At this time we can only say that our anthropological bias resulted in the development of our instruments on effective data, while the history of sociometry is almost completely dominated by affective data. This may be one reason for the lack of progress beyond the reformulation of sociometric data into matrices and graphs. It is difficult to imagine what an affective network might look like or mean. Effective networks are more easily imagined, and construction of networks rather than structures has been our goal.

It is clear that reciprocity is a very tough requirement; it reduces connectivity radically. We point out again that for effective data (or people's cognition about effect, to be more honest), reciprocity is too stringent (even irrelevant). If A ranks B 2nd, and B ranks A 10th, then (a) this says significant things about communication and (b) it is lost in a reciprocal sociogram. We have countless examples of this occurring.

By the time we get to seven unreciprocated choices, the links are approaching the complexity of our own catij diagram. There are, in fact, more links in an unreciprocated 7-choice sociogram (364 = 7N) than in the catij-1 matrix (335). The overlap is, of course, reasonable. 282 of the 364 links (77%) are catij-1's and 23% are not--some are even catij-4's. 53 catij-1's (16%) were missed by using the sociogram. It is precisely these links which form vital between-group links and which are omitted by the sociogram (cf. Granovetter, 1974 for a discussion of the importance of weak ties in the structuring of groups). If we had to choose a limited choice sociogram, we would insist on at least 7 unreciprocated choices. However, the point is that we do not have to make this choice. There is no direct proof that the catij matrix is better than a sociogram, but it is surely different. To date we have used the catij to describe school classrooms, prison living units, ships at sea, and institutional bureaucracies up to 143 persons. In all cases the results conform to observation and expert knowledge by members of the groups. In several cases the results were used as direct aids to management decision making. Thus, although, we have no direct numerical proof that the catij is superior to traditional sociograms, we are encouraged by its apparent usefulness in field studies of complex groups.

Bibliography

Alba, Richard D.
1973 "A graph-theoretic definition of a sociometric clique," J. of Math. Soc., 3:113-26.

Bernard, H.R., and Killworth, P.D.
1973 "On the social structure of an ocean-going research vessel and other important things" Social Science Research, 2:145-84.

Forsyth, E. and Katz, L.
1946 "A matrix approach to the analysis of sociometric data," Sociometry, 9:340-47.

Harary, F. and Ross, I.C.
1957 "A procedure for clique detection using the group matrix," Sociometry, 20:205-16.

Hubbell, C.H.
1965 "An input-output approach to clique identification," Sociometry, 28:377-99.

Killworth, P.D.
1974 "Intransitivity in the structure of small closed groups," Social Science Research, 3:1-23.

MacRae, D.,Jr.
1960 "Direct factor analysis of sociometric data", Sociometry, 23:360-71.

Wright, Benjamin, and Evits, Mary
1961 "O-technique and direct factor analysis in sociometry," Sociometry, 24:82-98.

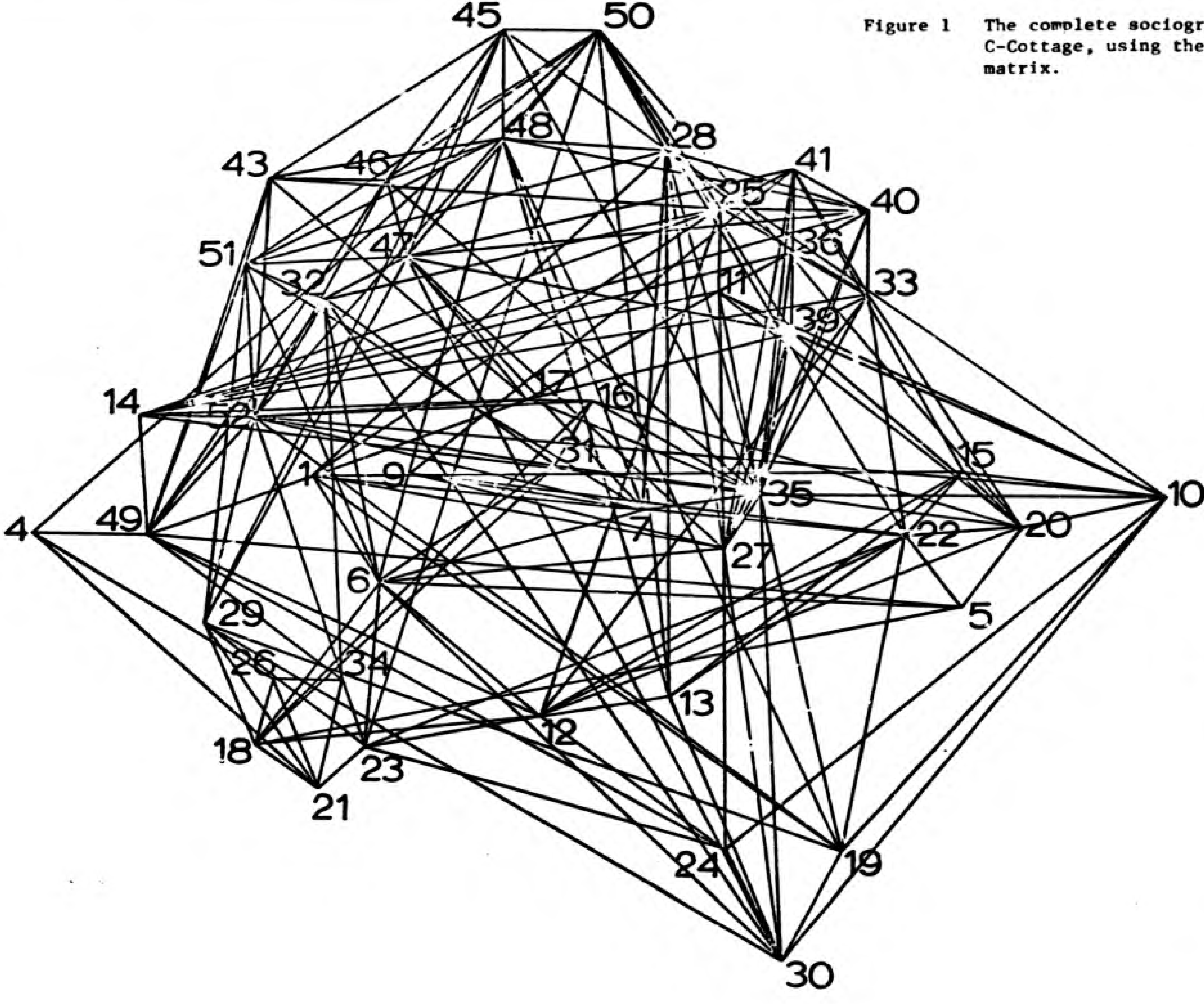
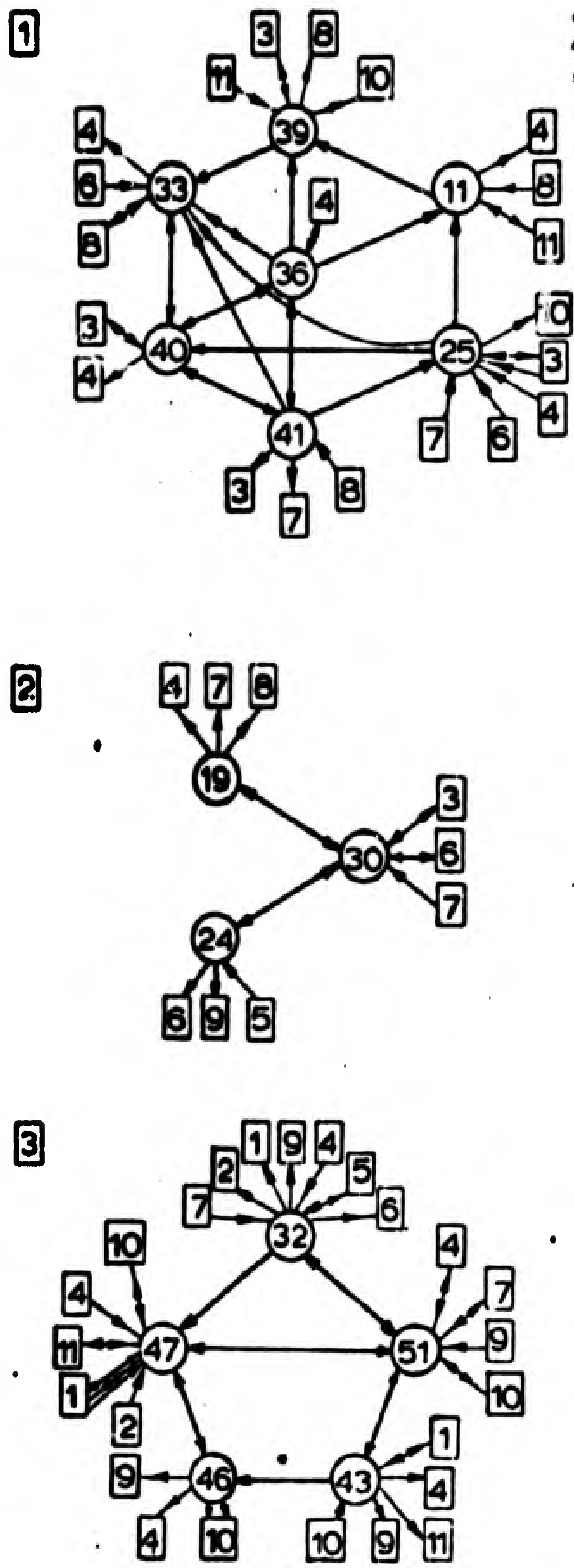
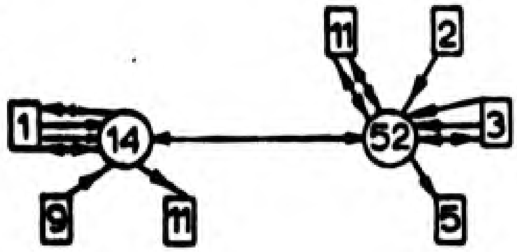


Figure 1 The complete sociogram of C-Cottage, using the cati! matrix.

Figure 2 Groups in C-Cottage, obtained by factoring the catij matrix. Group members are shown in circles. Relationships to other groups are shown in squares.

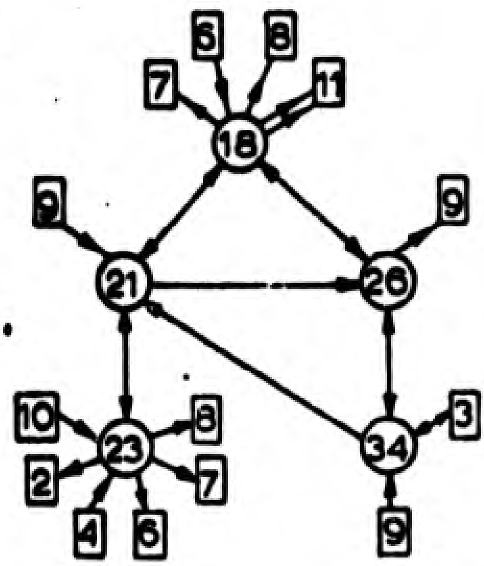


4



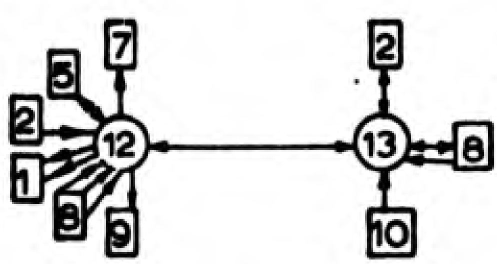
1
14
9
11
52
11
2
3
5

5



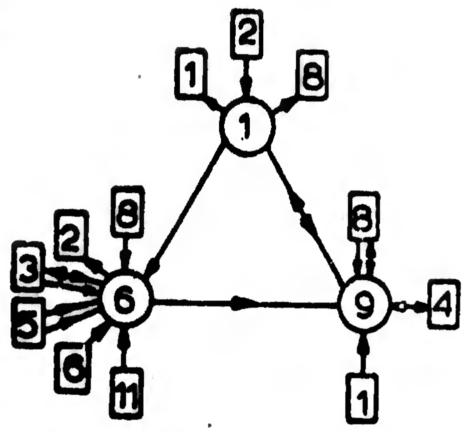
6
8
7
11
18
9
21
9
26
10
23
8
2
7
4
6
34
3
9

6

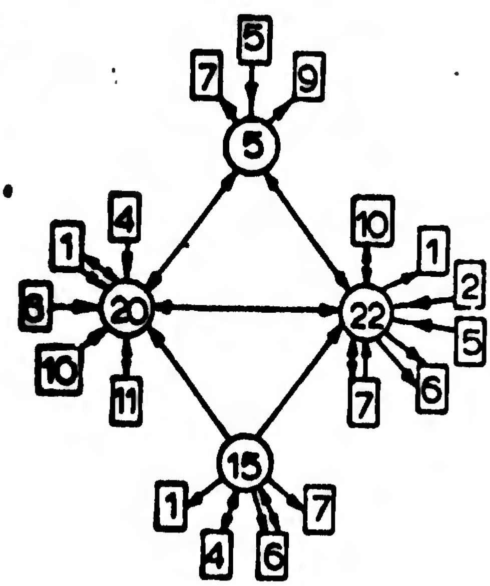


7
5
2
12
1
8
9
2
13
8
10

7


1
2
8
1
8
2
3
5
6
11
6
8
9
4
1

8


5
7
9
5
4
1
8
20
10
11
10
1
2
22
5
6
7
15
1
4
6
7

9

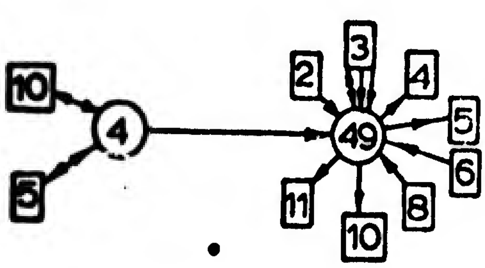

10
4
5
2
3
4
5
49
6
11
10
8

Figure 3a Results of factoring dt! on C-Cottage. Links are reciprocal: $\leq$ 3 links.

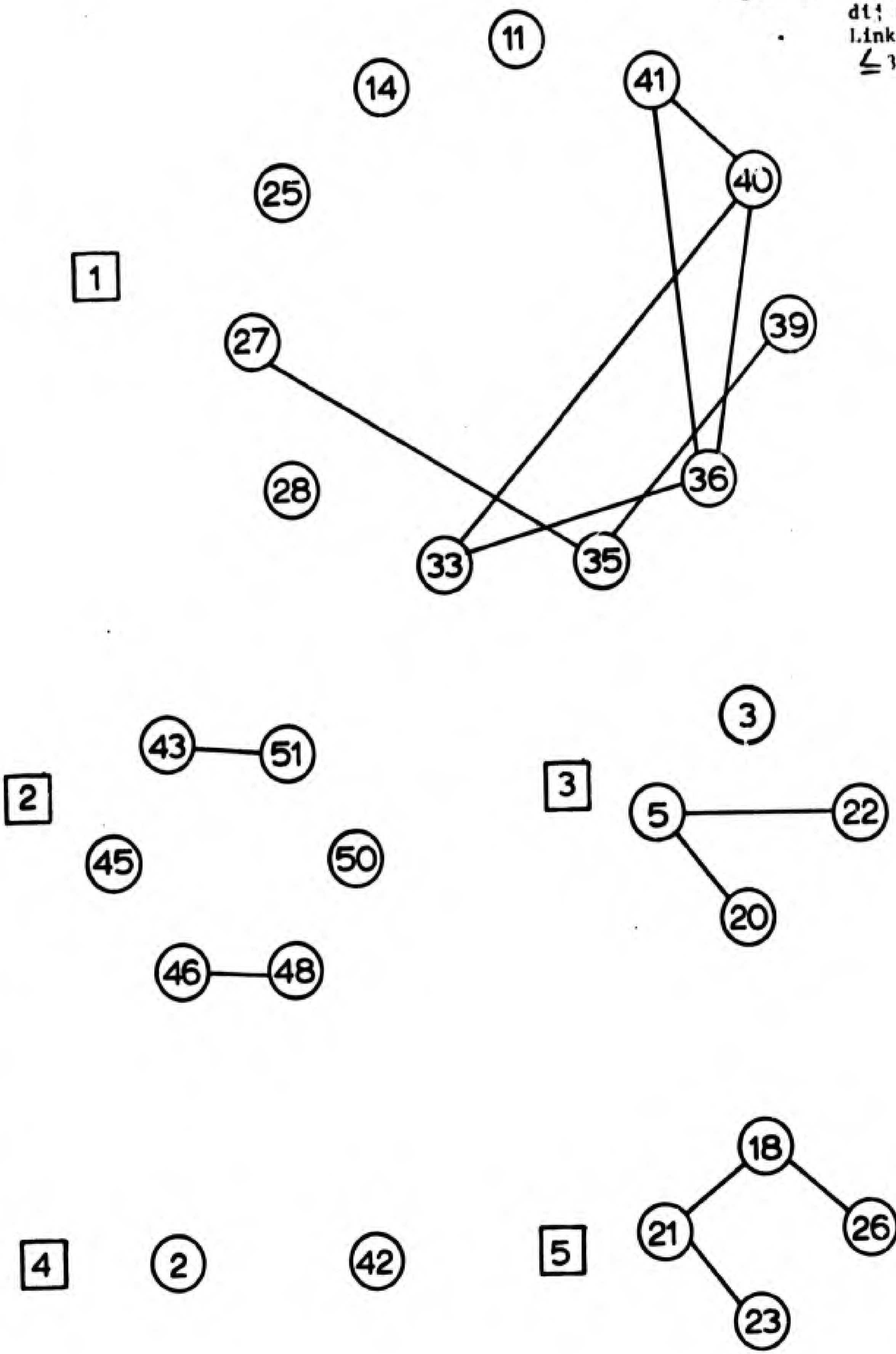
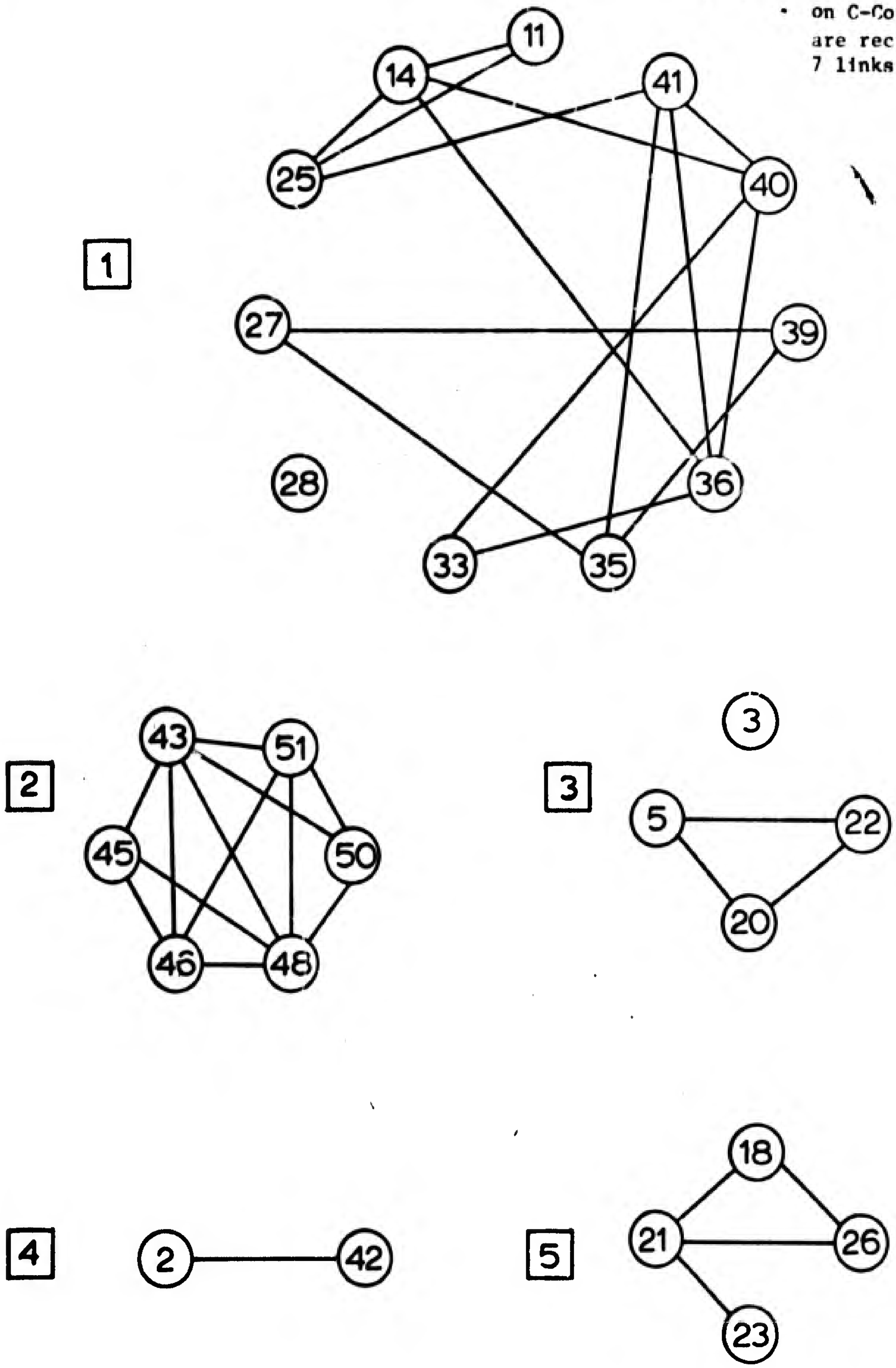


Figure 3b Factoring of d_{ij} on C-Cottage. Links are reciprocal; $\leq$ 7 links.



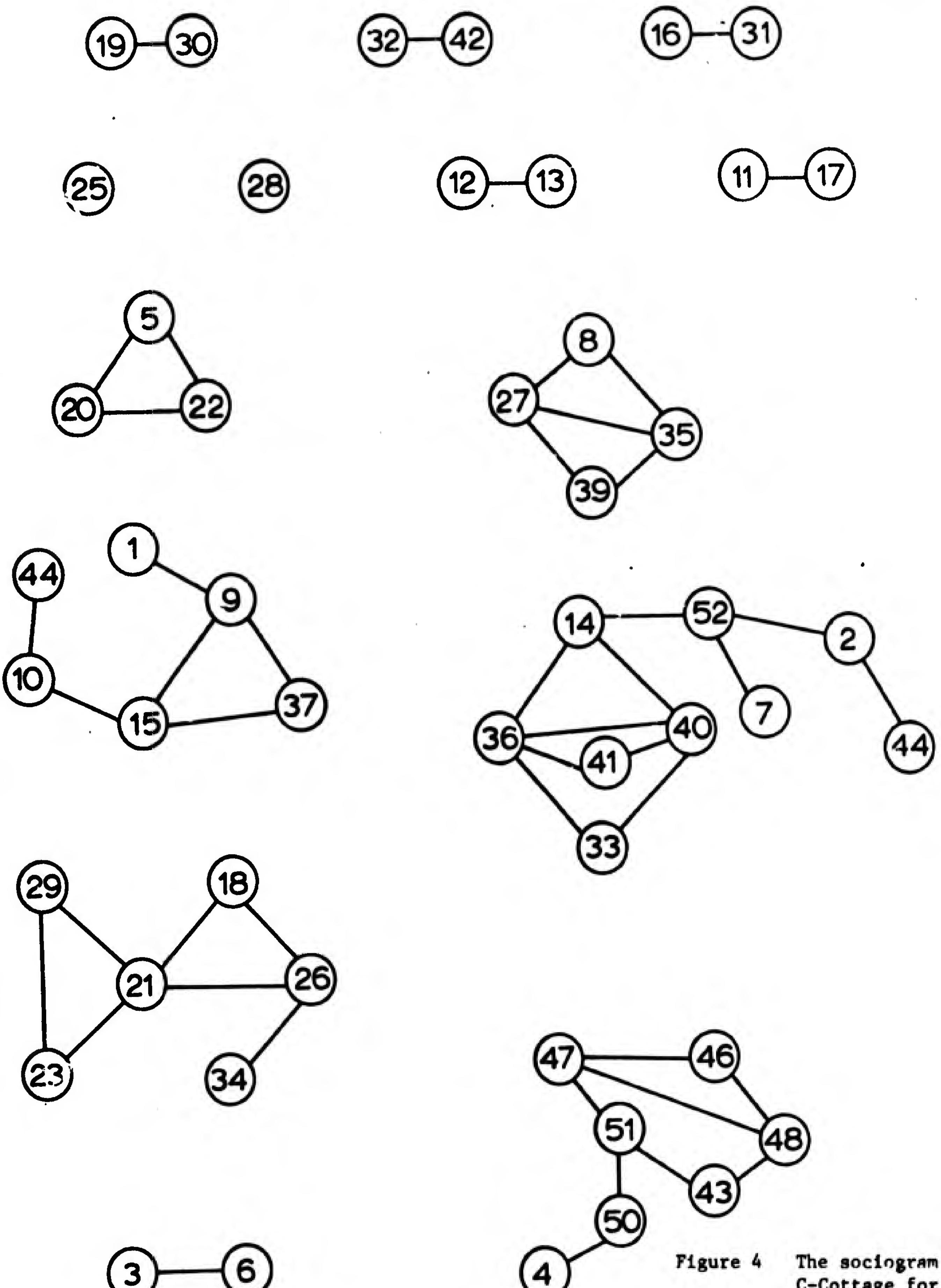
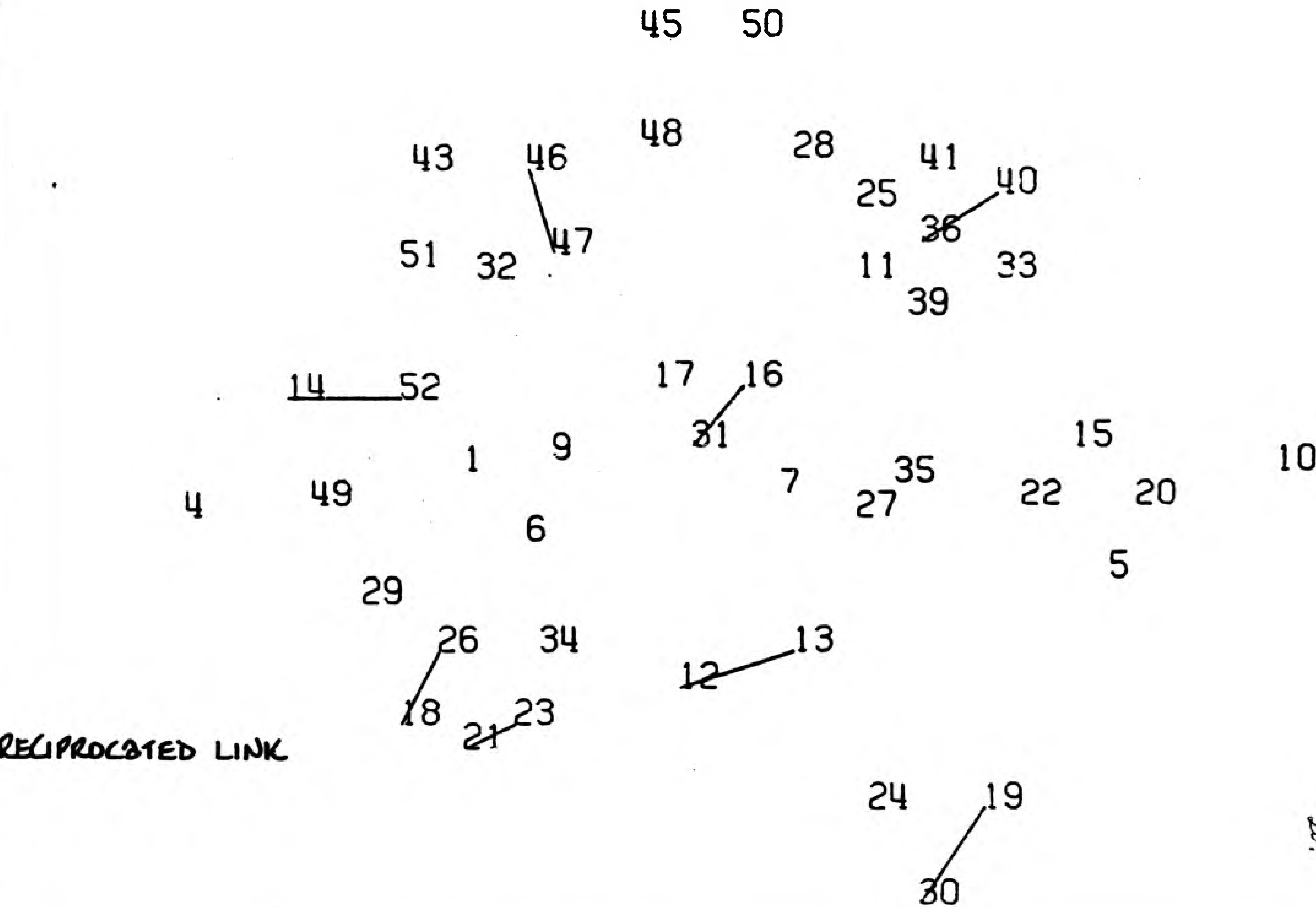


Figure 4 The sociogram of C-Cottage for four choices, reciprocated

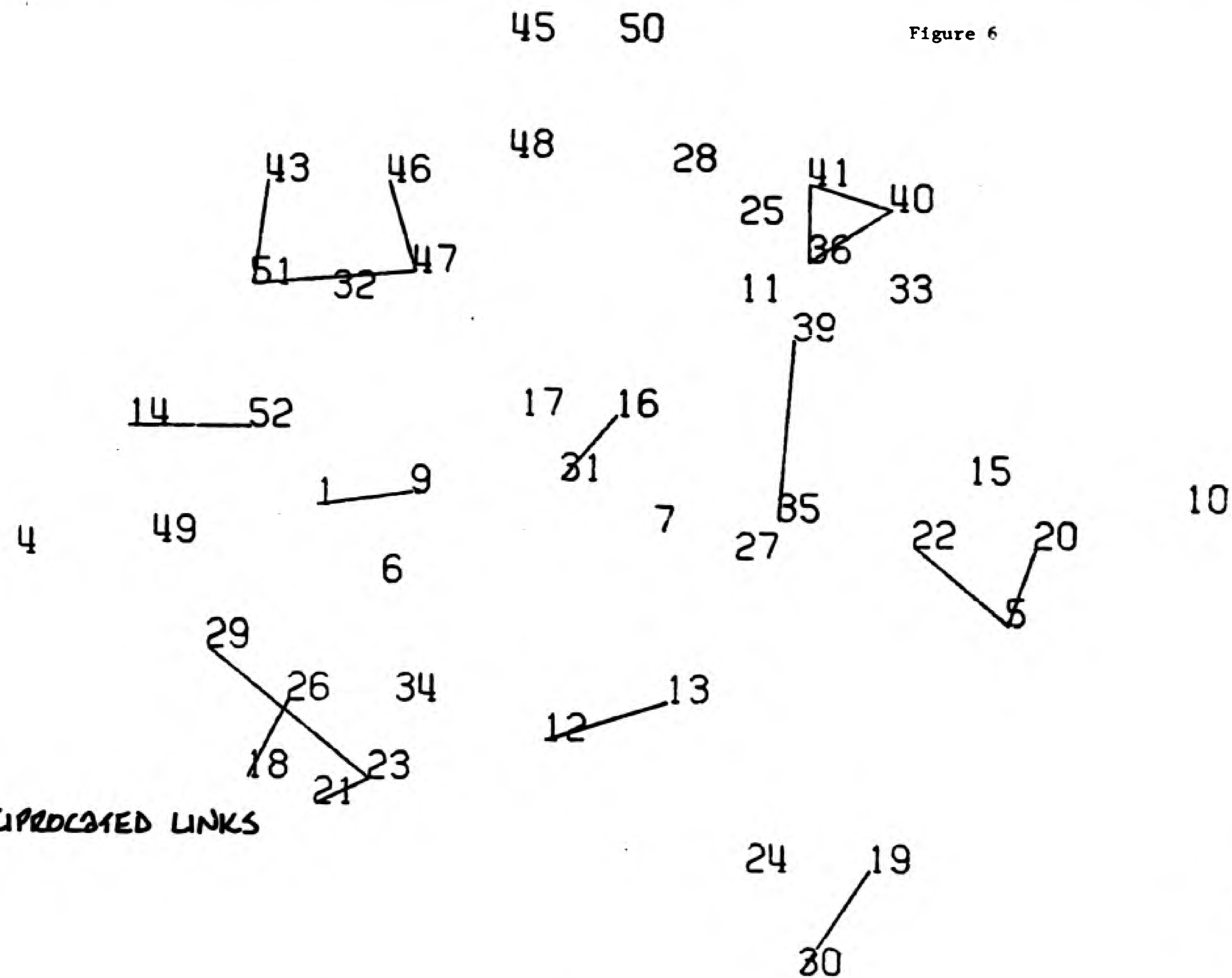
Figure 5-18 Links between persons in C-Cottage, given limited choice sociogram for reciprocated and unreciprocated choices.

Figure 5



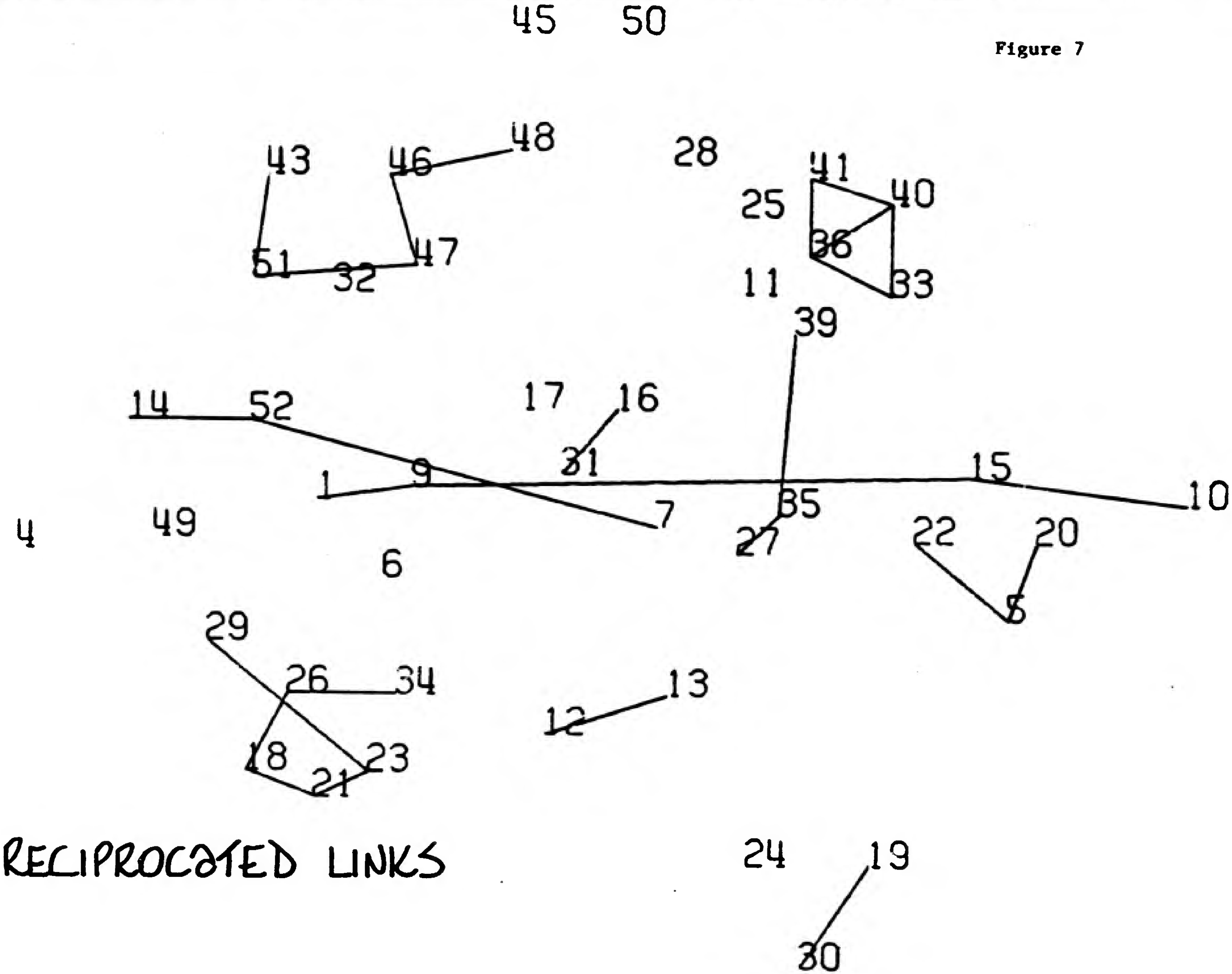
45
50
43
46
48
28
41
25
40
36
47
51
32
11
33
39
14
52
17
16
31
15
10
1
9
35
7
27
22
20
4
49
6
5
29
26
34
13
12
18
23
21
24
19
30
RECIPROCATED LINK

Figure 6



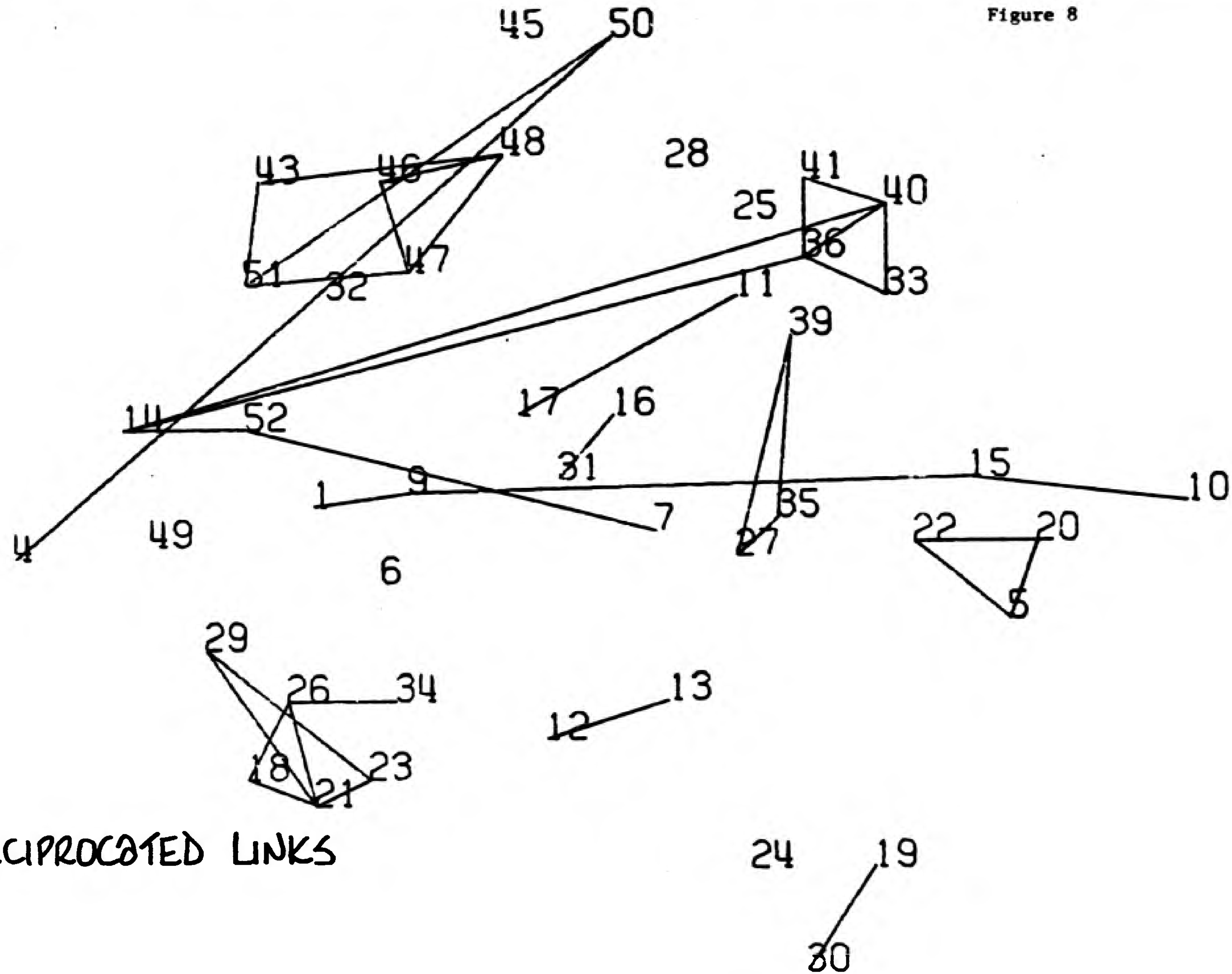
TWO RECIPROCATED LINKS

Figure 7



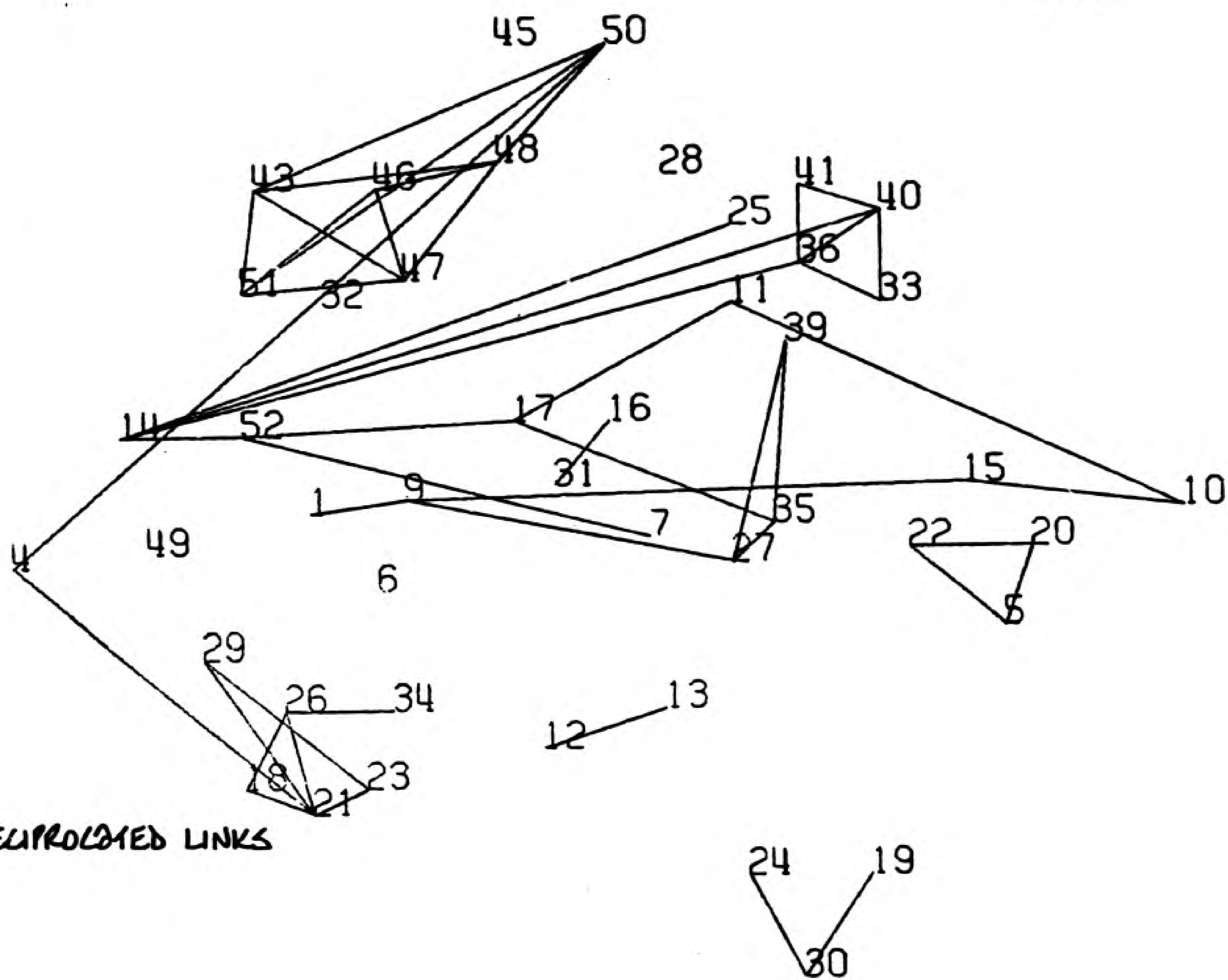
3 RECIPROCATED LINKS

Figure 8



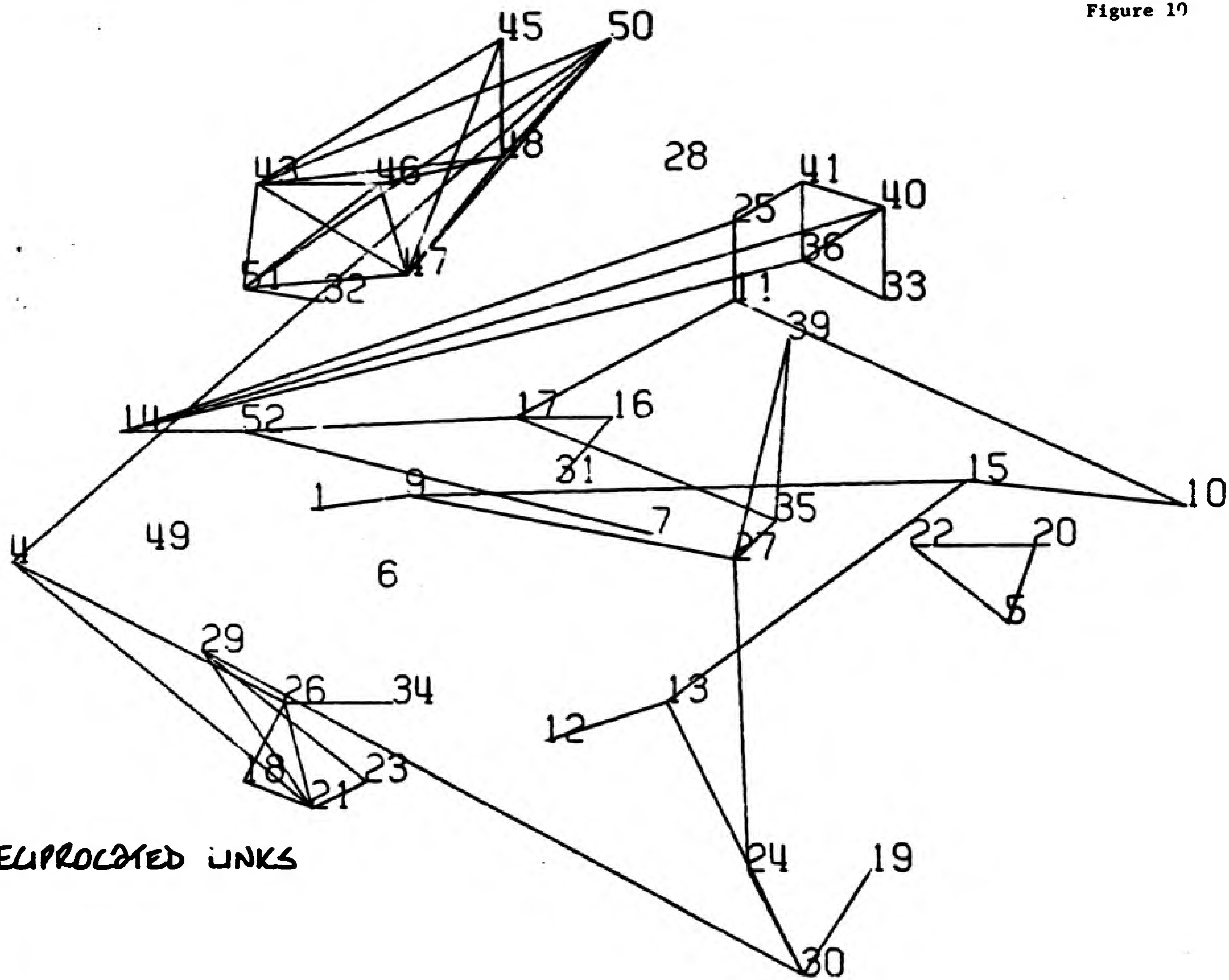
4 RECIPROCATED LINKS

Figure 9



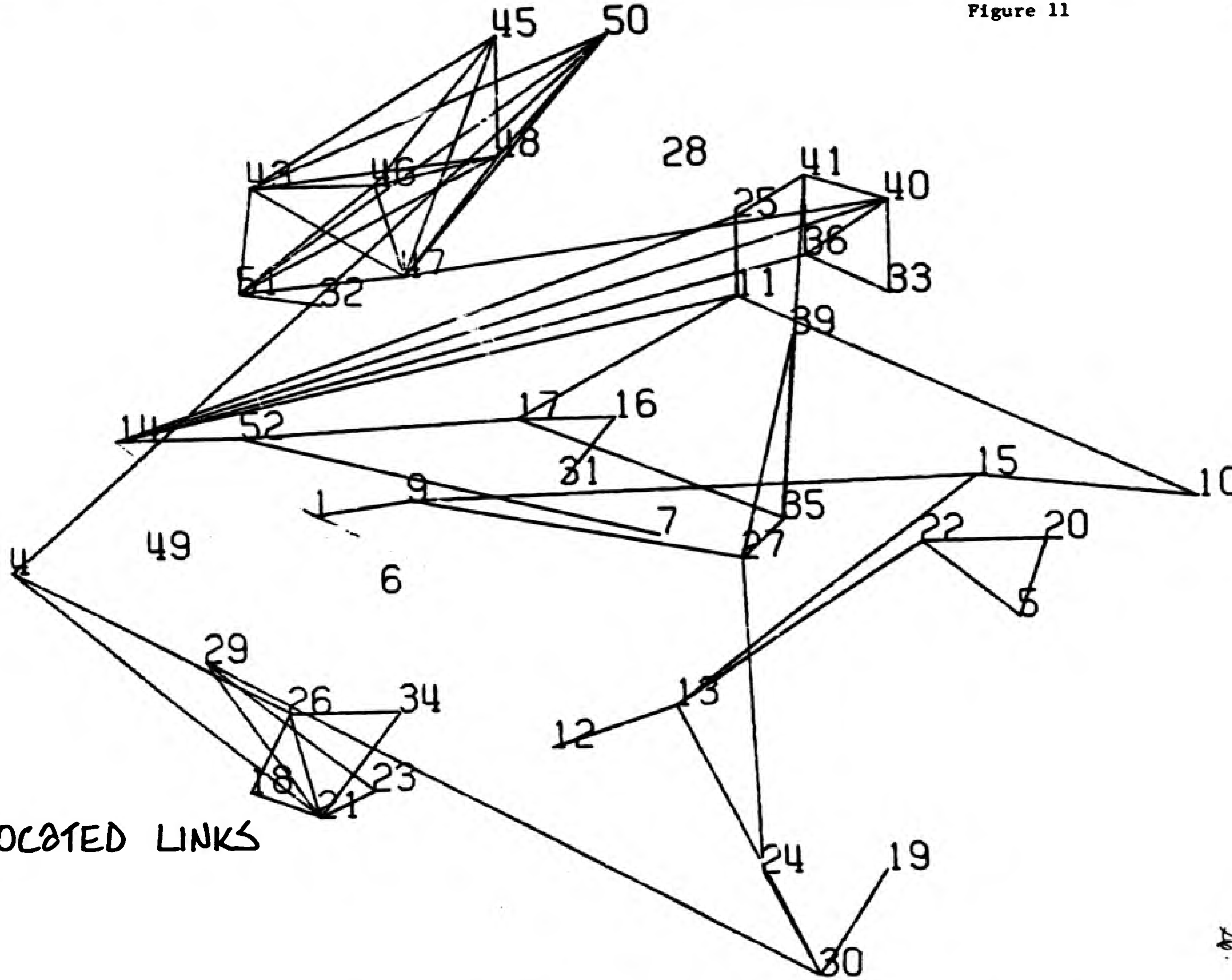
5 RECIPROCATED LINKS

Figure 19



RECIPROCATED LINKS

Figure 11



7 RECIPROCATED LINKS

Figure 12

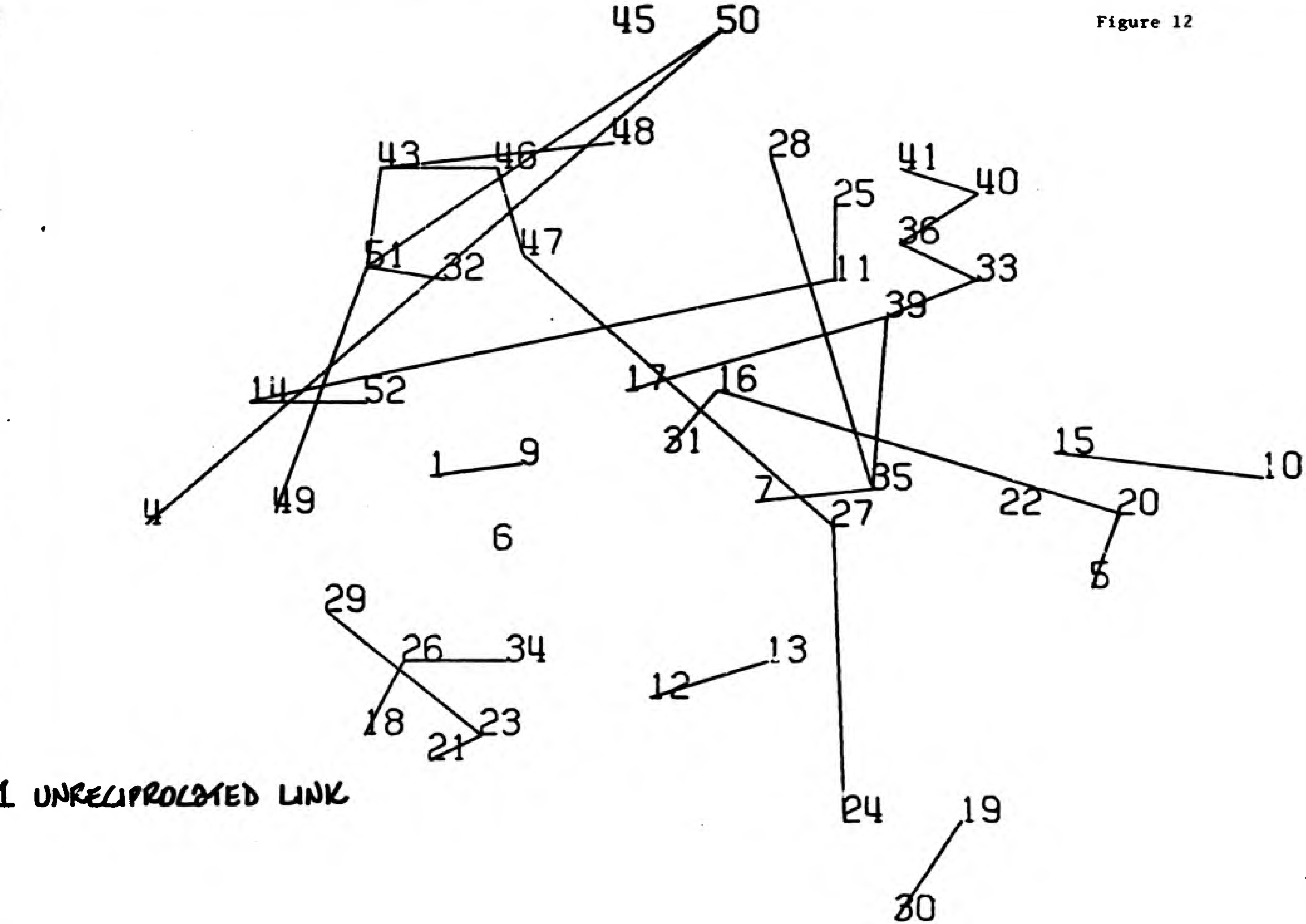
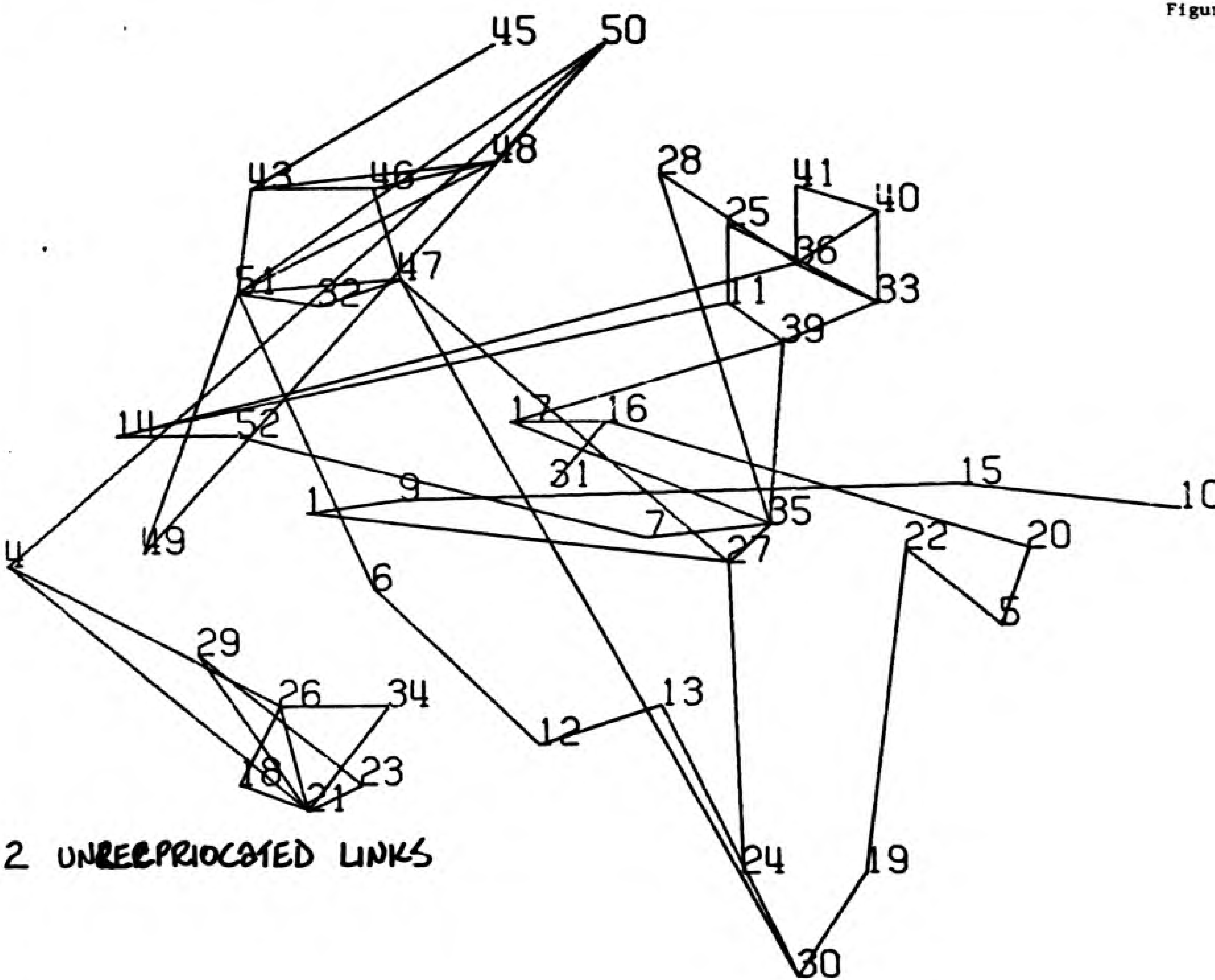
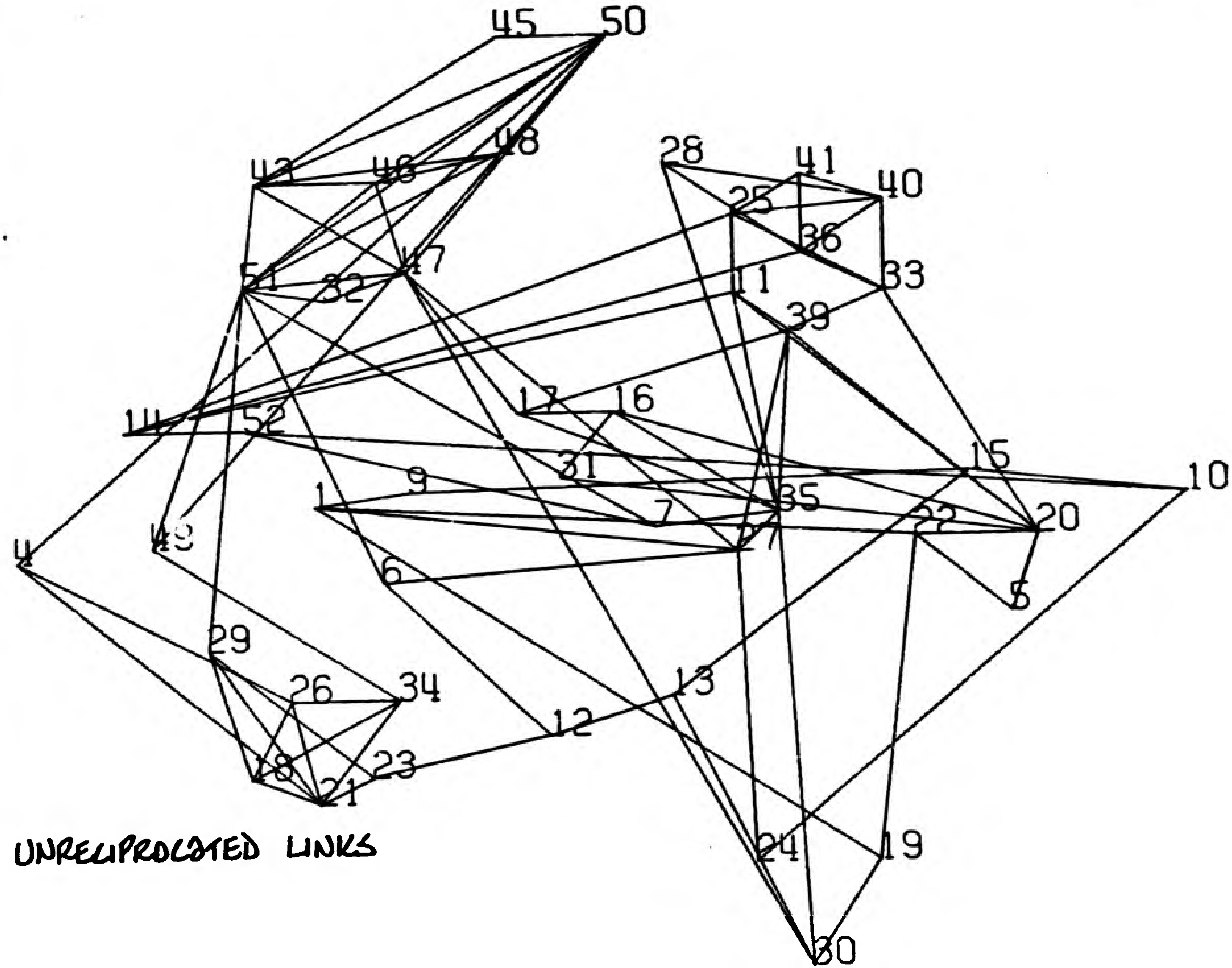

45
50
48
43
46
28
41
40
25
36
47
51
32
11
33
39
17
16
14
52
31
15
10
1
9
35
22
20
49
7
27
4
6
5
29
26
34
13
12
18
23
21
1 UNRECIPROCATED LINK
24
19
30

Figure 13

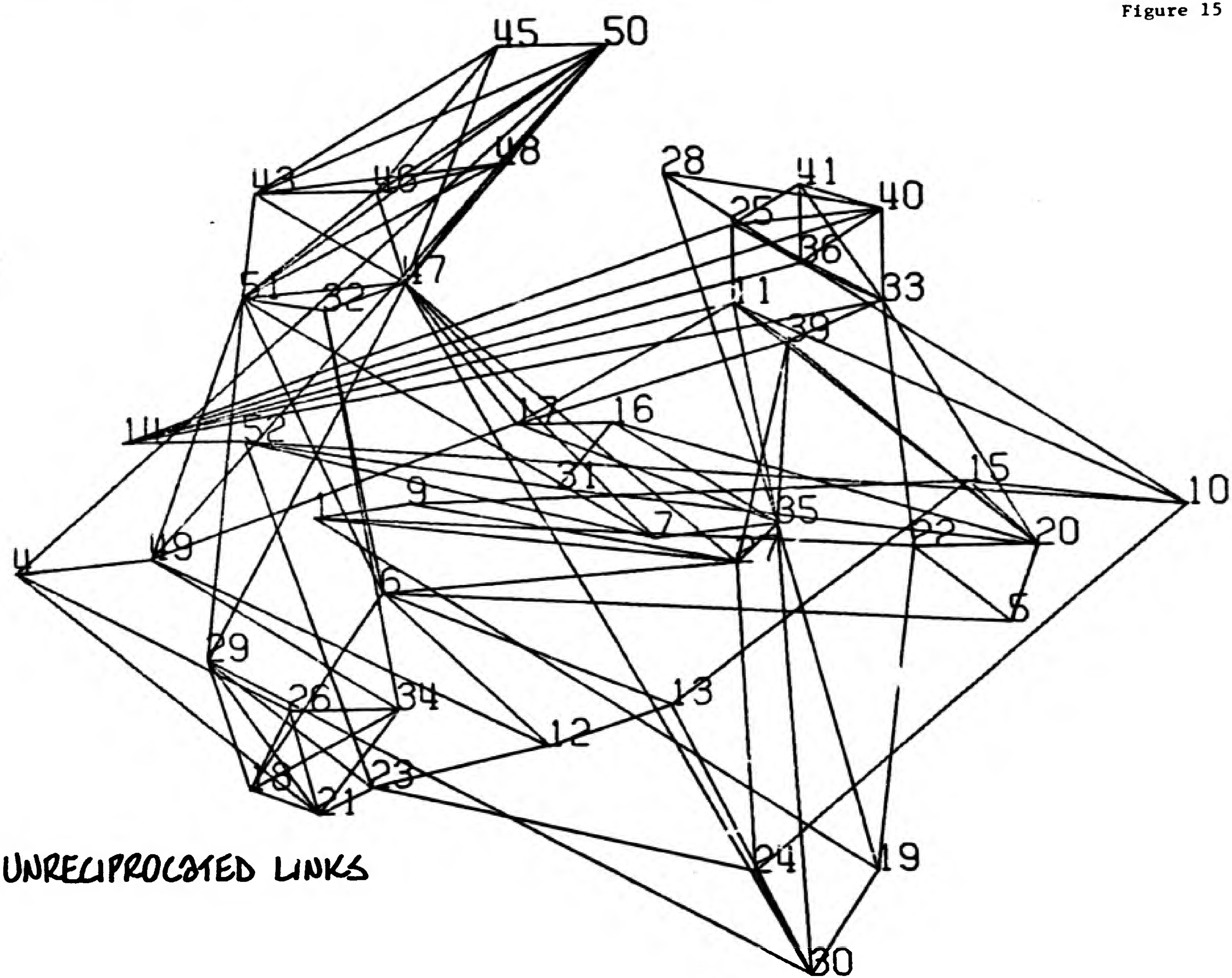


2 UNRECPRIOCATED LINKS



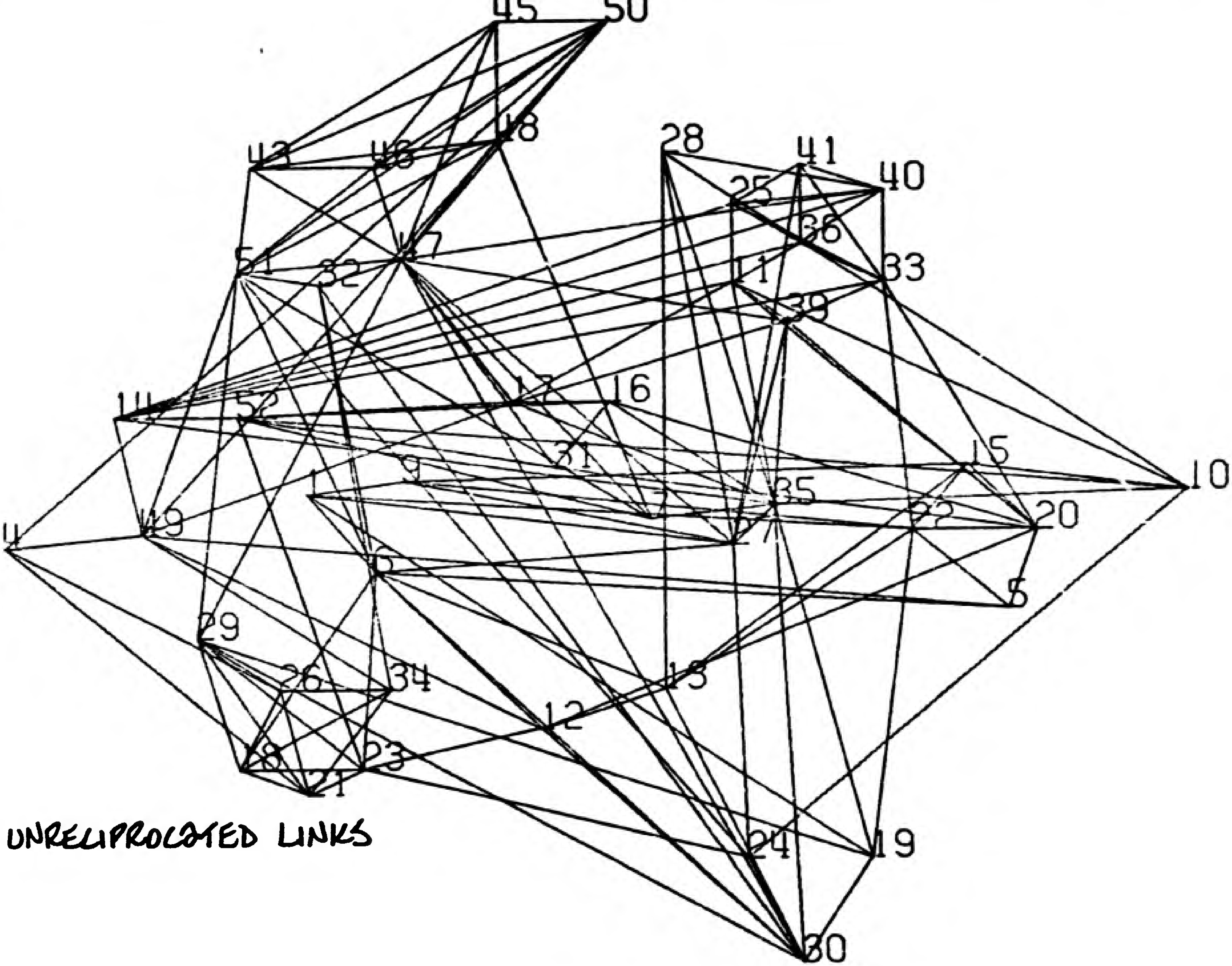
3 UNRECIPROCATED LINKS

Figure 15



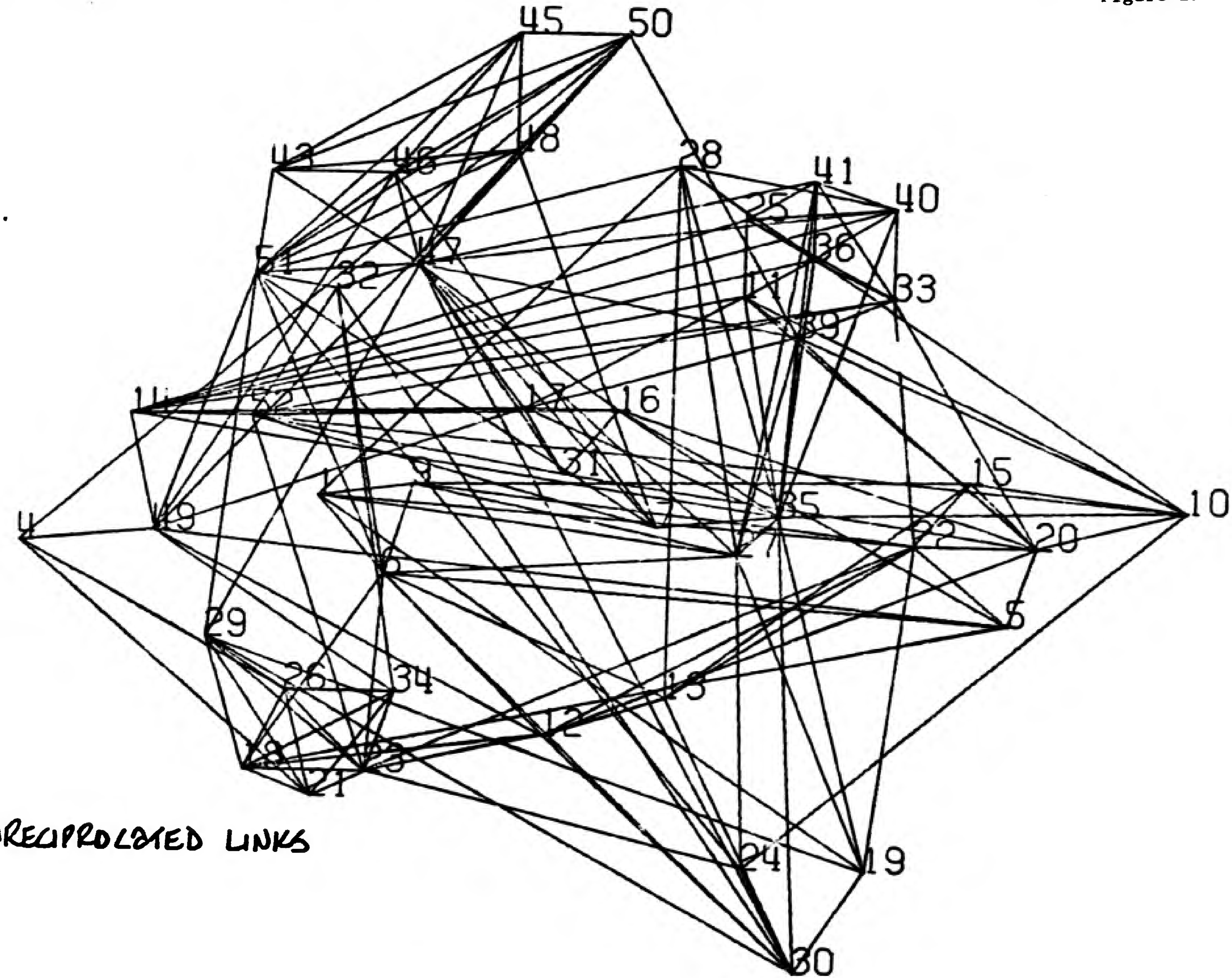
4 UNRECIPROCATED LINKS

Figure 16



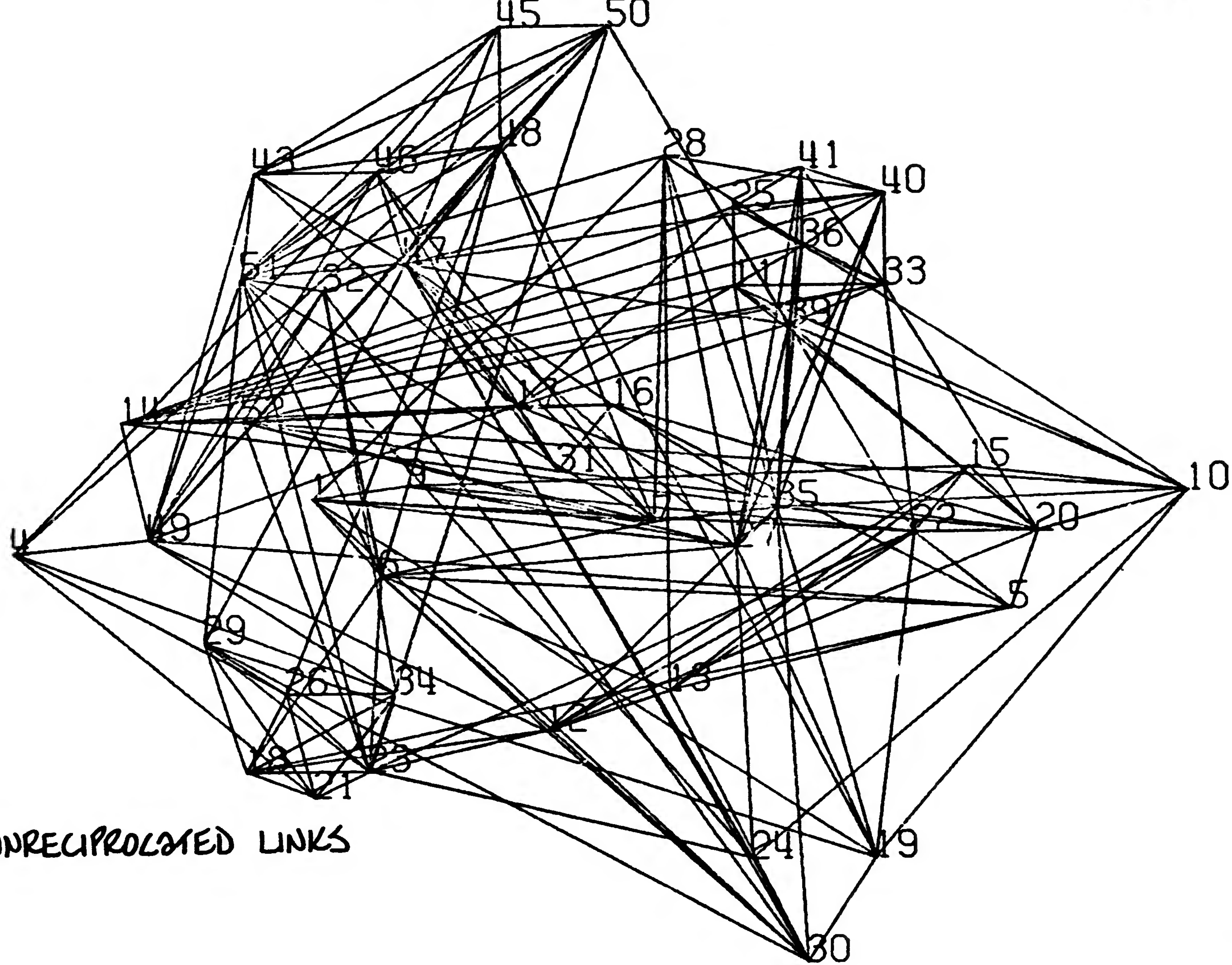
5 UNRECIPROCATED LINKS

Figure 17



UNRECIPROCATED LINKS

Figure 18



7 UNRECIPROCATED LINKS