Social Networks 11 (1989) 65–88
North-Holland

THE CLASS OF ALL REGULAR EQUIVALENCES: ALGEBRAIC STRUCTURE AND COMPUTATION †

Stephen P. BORGATTI *
University of California, Irvine

Martin G. EVERETT **
Thames Polytechnic

In this paper, we explore the structure of the set of all regular equivalences (White and Reitz 1983), proving that it forms a lattice, and suggest a general approach to computing certain elements of the lattice. The resulting algorithm represents a useful complement to the White and Reitz algorithm, which can only find the maximal regular equivalence of a graph. Using this algorithm, it is possible to compute several well-known equivalences, such as structural equivalence (Lorrain and White 1971), automorphic equivalence (Everett and Borgatti 1988), and Winship–Pattison equivalence (Winship and Mandel 1983). In addition, any number of other useful equivalences may be generated, providing suitable mathematical descriptions of them are available.

Introduction

Drawing on the work of social theorists such as Linton (1936), Nadel (1957) and Merton (1959), Lorrain and White (1971) introduced structural equivalence as a mathematical representation of the social role concept. They defined structural equivalence as follows:

> Objects a, b of a category C are *structurally equivalent* if, for any morphism M and any object x of C, aMx if and only if bMx, and xMa if and only if xMb. In other words, a is structurally equivalent

† This work arose from valuable discussions with D.R. White.
* School of Social Science, University of California, Irvine CA 92717, U.S.A.
** School of Mathematics, Statistics and Computing, Thames Polytechnic, Wellington Street, London SE18 6PF, U.K.

> to b if a relates to every object x of C in exactly the same ways as b does. From the point of view of the logic of the structure, then, a and b are absolutely equivalent, they are substitutable.
>
> (Lorrain and White 1971: 81)

In short, two actors are identically positioned if they are equally related to every other in the network. However, as Sailer (1978) notes, the difficulty with using structural equivalence as a formalization of the role concept is that for two actors to play the same role, the model requires that they know the same people. Thus two doctors, despite similar patterns of interaction with patients, nurses, suppliers, and other doctors, are not necessarily seen as playing the same role.

Several authors, including Sailer (1978), Winship and Mandel (1983), White and Reitz (1983), Everett (1985), and Breiger and Pattison (1986), have introduced models that more closely capture the notion of role. Of these, perhaps the most general is the White and Reitz formulation.

White and Reitz define *regular equivalence* as follows:

> If $G(V, R)$ is a graph (directed or undirected) with vertex set V, edge-set R, and no isolated vertices [1], and $\equiv$ is an equivalence relation on V then $\equiv$ is a regular equivalence if and only if for all a, b, c in V, $a \equiv b$ implies both
>
> (i) aRc implies there exists $d \in V$ such that bRd and $d \equiv c$;
>
> (ii) cRa implies there exists $d \in V$ such that dRb and $d \equiv c$.

Thus, two actors are regularly equivalent if they are equally related to equivalent others. Whereas in structural equivalence the focus is on the pattern of relations between individuals, in regular equivalence the focus is on the pattern of relations between positions or classes. Two doctors are equivalent because they have the same relationships with patients, nurses, and suppliers, not individuals.

An aspect of White and Reitz's definition of regular equivalence that has been largely overlooked is that it defines not one but a collection of regular equivalences, one of which is the equivalence computed by the

[1] For clarity of exposition, we assume G contains no isolates. In addition, we assume there is only one relation. However, all our results extend without difficulty to graphs with isolates and multiple relations.

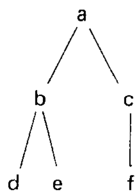

Fig. 1.

REGE algorithm (White 1984; White and Reitz 1985), and another is a form of structural equivalence. For example, for the graph in Figure 1, the following partitions are all consistent with regular equivalence:

(1) (a, b, c, d, e, f)
(2) $(a)(b, c)(d, e, f)$
(3) $(a)(b)(c)(d, e)(f)$
(4) $(a)(b)(c)(d)(e)(f)$

Partition #1, the maximal regular equivalence, is the one found by REGE. Partition #3 gives the maximal structural equivalence. It might be argued that neither the maximal regular equivalence nor the the maximal structural equivalence does a very good job of capturing the structure found in Figure 1. Rather, partition #2 would seem to be the best one. However, there are other regular partitions of this graph in addition to the ones shown, any one of which might be considered the "best" model. In the following section we describe the structure of the set of all regular equivalences of a graph.

The class of regular equivalences

Let $G(V, R)$ be a graph. A partition π of G is a subset of the power set of V, $P(V)$, which satisfies the following properties:

(i) $\emptyset$ is not an element of π
(ii) For all $v \in V$ there exists exactly one $A \in \pi$ with $v \in A$

If π_0 and π_1 are partitions of G, we will write $\pi_0 \leq \pi_1$ if for every $B \in \pi_0$, there exists $C \in \pi_1$ such that B is a subset of C.

As we have already indicated, any regular equivalence is a partition, which we call a *regular partition*.

The set of all partitions $E(G)$ of a graph form a partially ordered set. That is, the relation $\leq$ is reflexive, anti-symmetric, and transitive. In fact, $E(G)$ forms a lattice. That is, given any two elements π_0, $\pi_1 \in E(G)$ we can form the least upper bound $\pi_0 \vee \pi_1$ and the greatest lower bound $\pi_0 \wedge \pi_1$ in $E(G)$, where $\vee$ and $\wedge$ are the *join* and *meet* operators respectively.

Let $R(G)$ be the set of all regular partitions of a graph G. Since $R(G)$ is a subset of $E(G)$, it follows that $R(G)$ is a partially ordered set.

We are now in a position to prove that the collection of regular equivalences $R(G)$ form a lattice.

Lemma. In a partially ordered set $(X, \leq)$ if $\vee H$ exists for any proper subset H of X, then $(X, \leq)$ is a lattice.

The result is well known and can be found in standard algebra texts. However, the case when $H = \emptyset$ requires special mention. Any element of X is then vacuously an upper bound, and so the existence of $\vee \emptyset$ is equivalent to saying X has a minimum element. (This element is often called the zero element; $0 \in X$ is a zero if and only if for all $x \in X$, $0 \leq x$.)

Theorem 1. The set $(R(G), \leq)$ is a lattice.

Proof. The partition $w(G)$ which places each vertex in its own equivalence class (the *equality* partition) is easily seen to be regular for any graph. Obviously $w(G)$ is a zero element and hence by the lemma we need only consider the least upper bound for non-empty collections of regular equivalences.

Let $\equiv_i$ be a family of regular equivalences indexed by I $(\neq \emptyset)$. Define $\equiv$ as follows: $a \equiv b$ if and only if there exists a sequence $Z_0, Z_1, \ldots, Z_n$ with $a = Z_0$ and $b = Z_n$ such that for all j in the range $1 \leq j \leq n$, there exists $i_j \in I$ with the property $Z_{j-1} \equiv_{i_j} Z_j$. We shall prove that $\equiv$ is equal to $\vee$ $(\equiv_i, i \in I)$. It is well-known that $\equiv$ is the join for the family of partitions, but we shall include the proof of this result for completeness.

We first show that $\equiv$ is an equivalence relation:

(*Reflexivity*) $a \equiv a$ since a, a is the required sequence.
(*Symmetricity*) If $a \equiv b$ then there exists $Z_0, Z_1, \ldots, Z_n$ with the required property. The sequence $Z_n, Z_{n-1}, \ldots, Z_0$ shows that $b \equiv a$.
(*Transitivity*) Suppose $a \equiv b$ and $b \equiv c$, putting together the two sequences corresponding to the two equivalences guarantees that $a \equiv c$.

Hence $\equiv$ is an equivalence relation.

If $a \equiv_i b$ then $a \equiv b$ since the sequence a, b guarantees this. That is, $\equiv_i \leq \equiv$ so that $\equiv$ is an upper bound. Next we show that it is a least upper bound.

Suppose Φ is another upper bound. If $a \equiv b$ then there exists $Z_0, Z_1, \ldots, Z_n$ with $Z_{j-1} \equiv_{i_j} Z_j$, but since Φ is an upper bound (i.e. $\equiv \cdots_i \leq \Phi$) then $Z_{j-1} \Phi Z_j$. But Φ is an equivalence relation and therefore by transitivity $Z_0 \Phi Z_n$. That is, $a \Phi b$ so that $\equiv \leq \Phi$.

Finally, we prove that $\equiv$ is regular. Suppose $a \equiv b$ and aRc. Since $a \equiv b$ there exists a sequence $a, Z_1, Z_2, \ldots, Z_{n-1}, b$ where $Z_i \equiv_{i_1} a$. But since $\equiv_{i_1}$ is regular and aRc then there exists d_1 such that $Z_1 R d_1$ and $d_1 \equiv_{i_1} c$. In addition $Z_2 \equiv_{i_2} Z_1$ and since $\equiv_{i_2}$ is regular and $Z_1 R d_1$ then there exists d_2 such that $Z_2 R d_2$ and $d_2 \equiv_{i_2} d_1$. Continuing in this manner we produce a sequence $c, d_1, d_2 \ldots$ until finally we have $b \equiv_{i_n} Z_{n-1}$, again by regularity there will exist a d_n such that bRd_n and $d_n \equiv_{i_n} d_{n-1}$. The construction of the sequence $c, d_1, d_2, \ldots, d_n$ guarantees that $c \equiv d_n$. Consequently we have that if $a \equiv b$ and aRc then d_n is such that bRd_n and $c \equiv d_n$, which is precisely the first condition of regularity. The case cRa is similar, and hence $\equiv$ is regular. #

The theorem provides an alternative non-constructive proof of the existence of a maximal regular equivalence. The existence of such an equivalence was proved in a constructive way by White and Reitz (1983). It should further be noted that $R(G)$ is *not* a sublattice of $E(G)$, since although the joins are constructed in the same way, the meets are not. In $E(G)$ arbitrary meets are constructed as follows: if $\equiv_i$ are a collection of partition equivalences, then

$$a \wedge \equiv_i b \text{ iff } a \equiv_i b, \qquad \text{for all } i.$$

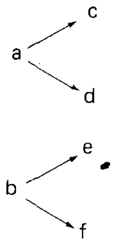

Fig. 2.

The following example shows that such a construction does not work for regular equivalences. In Figure 2 the partitions $\{\{a, b\}, (c, e\}, \{d, f\}\}$ and $\{\{a, b\}, \{c, f\}, \{d, e\}\}$ are both regular. However the meet equivalence $\{a, b\}, \{c\}, \{d\}, \{e\}, \{f\}\}$ is not regular since although $a \equiv b$ and aRc there is no $d \equiv c$ such that bRd.

The REGE algorithm provided by White and Reitz (White 1984; White and Reitz (1985) is based upon their constructive proof of the existence of a maximal regular equivalence and therefore only finds this element of the lattice. In many cases, this will be the complete partition $i(G)$ which groups all vertices into a single class. The following theorem characterizes these cases.

Theorem 2. The complete partition $i(G)$ is regular (and obviously maximal) if and only if G contains no sources or sinks. (Sources are vertices with zero indegree and sinks are vertices with zero outdegree.)

Proof. Suppose G contains no sources or sinks and let a and b be any two vertices of G. Hence there exists c such that aRc and d such that bRd. Further, in $i(G)$ all vertices are equivalent so $c \equiv d$. The relation cRa is similar and hence $i(G)$ is regular.

Conversely, suppose $i(G)$ is a regular partition. Suppose a is a source; since G contains no isolates there exists b such that aRb. Under $i(G)$ $b \equiv a$ and therefore there exists d such that dRa and $a \equiv d$, contradicting the fact that a is a source. The case where a is a sink is similar and the result follows. #

Corollary. In any undirected graph, $i(G)$ is the maximal regular equivalence.

These results are well-known to practitioners, but the proof has not previously appeared in the literature.

As Faust (1988) has pointed out in passing, the workings of REGE may be understood in terms of sinks and sources. A discrete version of the algorithm is as follows:

Step 1: Divide all vertices into three *classes*: sources, sinks, and others. If no sinks or sources, exit (there is just one equivalence class).

Step 2: Cluster together vertices whose set of alters have the same combination of classes. Let these new clusters become the classes. Repeat Step 2 until no change in class assignments.

Step 3: Print equivalence classes.

Non-trivial regular equivalences

We call a regular partition *non-trivial* if it is not $i(G)$, the complete partition, nor $w(G)$, the equality partition.

A graph is k-partite if the vertex set can be partitioned into k disjoint sets $V_1 \dots V_k$ such that any edge only connects vertices in V_i to V_j where $i \neq j$. A 2-partite graph is known as bipartite. It is well-known that a graph is bipartite if and only if it contains no odd semi-cycles. We shall say that a k-partite graph is a *semi-complete k-partite* graph if whenever there exists an edge connecting a vertex in V_i to a vertex in V_j then for every $x \in V_i$ there exists $y \in V_j$ such that (x, y) is an edge and for every $p \in V_j$ there exists $q \in V_i$ such that (q, p) is an edge. An example of a semi-complete 3-partite graph is shown in Figure 3.

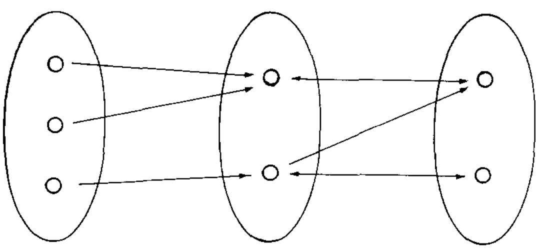

Fig. 3

Theorem 3. If G is a semi-complete k-partite graph then the vertex partition $V_1 \dots V_k$ is regular.

Proof. The result follows directly from applying the definition of regular equivalence to the vertex partition $V_1 \dots V_k$. #

Corollary. In an undirected bipartite graph the vertex partition V_1, V_2 is regular.

The corollary tells us that non-trivial regular partitions exist for undirected trees or any undirected graph which does not contain odd cycles. It should be noted that there exist graphs (both directed and undirected) for which no non-trivial regular partition exists.

The need for alternative regular equivalences

In this section we argue that there are two important situations where it is desirable to compute other elements of the lattice besides the maximal regular equivalence computed by REGE.

One such situation is evident in the results of the foregoing section: the circumstance where the maximal equivalence is trivial. An example is provided by the directed graph in Figure 4.

Clearly, the graph is highly structured. For example, the regular partition

$$\{\{a, b\}, \{c, d\}, \{e, f\}\}$$

captures that structure quite nicely. But since the graph contains no sinks or sources, the maximal regular partition is

$$\{\{a, b, c, d, e, f\}\}$$

which completely misses the structure.

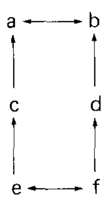

Fig. 4.

The second situation occurs when it is desired that the obtained equivalence or partition possess certain properties in addition to the basic requirements of regular equivalence.

Consider for example the case of "true substitutability". Under regular equivalence, there is no reason why a doctor with 2 nurses and 5 patients can not be equivalent to a doctor with 3 nurses and 11 patients. What determines the doctor role is that each player have the same relations with nurses and patients, regardless of how many such alters they might have. But two such doctors are not *substitutable*; they do not, for example, require the same amount of office space, the same number of parking spaces, and so on. This strict notion of substitutability is captured precisely by *automorphic equivalence*, as Winship (1974), Everett (1985) and Everett and Borgatti (1988) have shown. Automorphic equivalence is a regular equivalence, but a particularly strict form of it which will rarely coincide with the maximal element of the lattice.

It is worth noting that automorphic equivalence has another property in addition to substitutability that is not shared by all regular equivalences. This is the property that two automorphically equivalent nodes of G are also automorphically equivalent in the complement graph G' (where two nodes are adjacent if and only if they are not adjacent in G). In graphs where the coding of information as arcs or not-arcs is arbitrary, this is a critical property. For example, if the vertices of a graph represent people and the underlying relation is their similarity or difference in age, then we would expect to get the same blockings whether we record an arc for "same age" or for "different age"! However, invariance under such simple coding transformations is not a property of regular partitions in general, as the graph in Figure 5 demonstrates.

Of course, there are cases where such invariance is not of interest. For example, in a primate study we might record an arc if a certain relation is observed between a pair of monkeys during the test period, and not otherwise. Thus while the presence of an arc has a definite meaning, the absence can mean either that a monkey does not have that relation with another, or that we have simply not observed it. The ones and zeros in the adjacency matrix are not measured in the same way, and in fact should not be treated as such.

Another property that one might wish a regular equivalence to preserve is centrality (Freeman 1979). In an advice network, for example, it is clear that highly central actors play substantially different

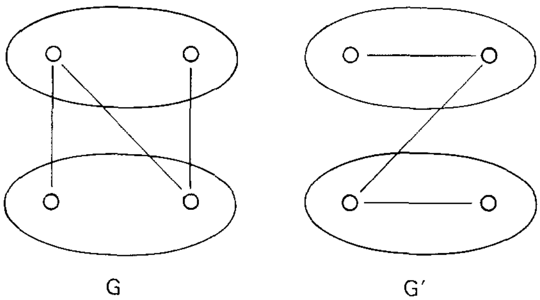

Fig. 5.

roles than do peripheral ones. Yet regular equivalence will not necessarily observe this distinction. For example, for the graph in Figure 6, the following betweenness centralities are found:

Vertex	Betweenness
a:	0.33
b:	6.00
c:	0.33
d:	6.00
e:	20.33
f:	20.33
g:	6.00
h:	0.33
i:	6.00
j:	0.33

Yet the following partitions are both regular:

$\{\{a, b, c, d, e, f, g, h, i, h, j\}\}$

$\{\{a, c, h, j, e, f\}\{b, d, g, i\}\}.$

Obviously, both partitions confound highly central points with maxi-

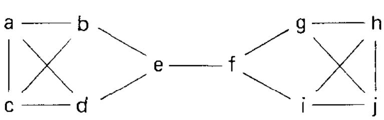

Fig. 6.

mally peripheral points. In such a case we might prefer a less maximal but centrality-preserving equivalence such as:

$$\{\{a, c, h, j\}\{e, f\}\{b, d, g, i\}\}.$$

Clearly, what is needed are computational tools for constructing other regular equivalences besides the maximal regular equivalences. In the next section we describe such a procedure called REGE/A.

Basic algorithm

Consider the following simplified operating description of the standard REGE algorithm:

Standard REGE:
For all $i, j, k, m \in V$, i and j are equivalent at iteration $t+1$ if for all k related to i there exists an m related to j such that
(i) i is to k as j is to m
(ii) $E_t(k) = E_t(m)$.

The phrase "i is to k as j is to m" means that the ties linking i with k are the same as those linking j with m. In a network with just one relation, there are three ways an actor may be linked to another: either $i \rightarrow k$, $i \leftarrow k$, or $i \leftrightarrow k$. Thus we can define a function $R^*(i, k)$:

$$R^* = \begin{cases} 1 & \text{if } i \rightarrow k \text{ only} \\ 2 & \text{if } j \leftarrow k \text{ only} \\ 3 & \text{if } i \leftrightarrow k \\ 0 & \text{otherwise} \end{cases}$$

The description of standard REGE may then be rewritten as follows:

Standard REGE:
For all $i, j, k, m \in V$, i and j are equivalent at iteration $t+1$ if for all k related to i there exists an m related to j such that
(i) $R^*(i, k) = R^*(j, m)$
(ii) $E_t(k) = E_t(m)$.

The condition "$E_t(k) = E_t(m)$" means that k and m were put in the same class on the previous iteration. An actor's class may be viewed as a categorical attribute of the actor, like gender or place of birth, that the procedure must preserve. By "preserve" we mean that two actors cannot be equivalent if their alters do not share this attribute. The only difference between the class attribute and one such as birthplace is that its value may change from iteration to iteration.

White and Reitz describe regular equivalence as the situation that occurs when a set of actors are equally related to equivalent others. This is easily verified using the notation above. Consider again the graph in Figure 6. A regular partition is:

$\{\{a, c, h, j\}, \{b, d, g, i\}, \{e, f\}\}$

In Table 1, column 2, we give the list of alters for each actor for this graph. In column 3, we replace the actor's identifying letter with the equivalence class in which they are placed by the above partition. It is quickly seen that while actors in a given equivalence class, say $\{a, c, h, j\}$, have different sets of alters, these alters have the same attribute values $E(k)$ (in this case they are all "1"s and "2"s). In short, the friends of equivalent actors belong to the same clubs.

Note that for an undirected graph there is only one way a point may be related to another; for such graphs the check in condition (i) is unnecessary. This reduces the algorithm to checking simply that the set of alters of one actor contains the same set of attributes as the set of alters of the other actor, where the categorical attribute is merely the equivalence class found in the previous iteration.

Table 1
List of alters and equivalence classes for Figure 6

Actor	Alters	$E(k)$
a	$c\ b\ d$	1 2 2
b	$a\ c\ e$	1 1 3
c	$a\ b\ d$	1 2 2
d	$a\ c\ e$	1 1 3
e	$b\ d\ f$	2 2 3
f	$g\ i\ f$	2 2 3
g	$h\ j\ f$	1 1 3
h	$j\ g\ i$	1 2 2
i	$h\ j\ f$	1 1 3
j	$h\ g\ i$	1 2 2

We now give a preliminary description of the REGE/A procedure:

REGE/A, version 1:
For all i, j, k, $m \in V$, i and j are equivalent at iteration $t+1$ it for all k related to i there exists an m such that
(i) $R^*(i, k) = R^*(j, m)$
(ii) $E_t(k) = E_t(m)$
(iii) $A_h(i) = A_h(j)$ for all $h \in H$.

The "H" in condition (iii) refers to a set of general point attributes, such as centrality or degree. $A_h(i)$ is the value held by actor i on attribute h. It is important to note that condition (iii) implies

$$A_h(k) = A_h(m) \qquad \text{for all } h \in H$$

on the second iteration. This says that i and j are equivalent if and only if they are equally related to the same kind of alters, where "same kind" is defined as having equal values on any number of user-specified attributes. This is because on the second iteration condition (ii) would require that a pair of possible equivalent actors have equivalent alters, that is alters who on the previous iteration were equal on all attributes.

The importance of this is that to compute the maximal regular equivalence that simultaneously preserved some attribute, one could not simply run standard REGE and subdivide the resulting blocks by the desired attribute. This fails because the resulting partitions need not be regularly equivalent. The undirected graph in Figure 7 is an obvious example. Let centrality be our only point-attribute. The sets of actors with equal centrality are:

Group 1: (1)
Group 2: (2)
Group 3: (5)
Group 4: (3 4 6 7)

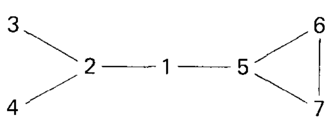

Fig. 7.

Subdividing REGE's partition (the complete partition) by the sets of equal centrality yields those same sets, but this partition is not a regular equivalence since not every actor in group 4 is connected to a member of group 3, as required by the definition.

It should be clear that there is no fundamental difference between the attribute in condition (ii), and an attribute of condition (iii). In fact, in a later section we will suggest that a useful choice of attributes is an equivalence classification generated by a competing measure of role.

It should also be clear that by describing the attributes in condition (iii) in the form that we have, we force any computer program implementing REGE/A to implicitly compute a boolean similarity matrix whose cells contain TRUE if the pair of actors in question have the same values for all attributes, and FALSE otherwise. For computational convenience, we could directly enter a similarity matrix S instead of a two-mode rectangular attribute matrix. In our previous notation,

$$S(i, j) \text{ is TRUE if } A_h(i) = A_h(j) \qquad \text{for all } h \in H.$$

Further, we could enter more than one of these, requiring that alters be similar on each similarity matrix. We rewrite the algorithm again as follows:

REGE/A, version 1b:
For all i, j, k, $m \in V$, i and j are equivalent at iteration $t+1$ if for all k related to i there exists an m such that
(i) $R^*(i, k) = R^*(j, m)$
(ii) $E_t(k) = E_t(m)$
(iii) $S_q(i, j) = \text{TRUE} \qquad \text{for all } q \in Q$,

where Q is a set of actor-by-actor similarity matrices. We will refer to any $S(i, j)$ in Q as a "point-similarity matrix".

Obviously, we can use REGE/A to find the maximal regular equivalence that preserves any point-attribute we choose. Examples of point-attributes might be network-theoretic, such as centrality and degree, or background information, such as age or occupation.

This puts an interesting light on Doreian's (1987) method of using REGE to generate a non-trivial equivalence on symmetric matrices. Essentially, he proposes splitting a symmetric relation into two asymmetric ones, and submitting these to REGE. The splitting criterion is simply that if CENTRALITY(i) > CENTRALITY(j) then RELATION1(i, j) = 1, else RELATION2(i, j) = 1. Doreian shows that this splitting operation preserves regular equivalence, and the result is the maximal regular equivalence that preserves centrality. Often, this will be a non-trivial equivalence (although it is possible for a graph with a high degree of structure to still only contain trivial partitions). However, using REGE/A, we are able to input the original relation unmodified, using point-centrality as our attribute. This also must result in the maximal regular equivalence that preserves centrality, and has the advantage of keeping attributes and relations apart.

An interesting attribute to take is "actor's name" or "actor's id number". This yields a partition where two actors are equivalent if and only if they are connected to all the same people: in short, it computes a form of structural equivalence.

An example of a more complex set of attributes is found in the computation of a very important member of the lattice: automorphic equivalence. Maximal sets of automorphically equivalent vertices are called orbits. While no polynomial-time algorithm is known to solve the automorphism problem exactly, Everett and Borgatti (1988) describe an approximate method based upon comparing successively wide neighborhoods of actors on several "difficult" attributes. The method is guaranteed to partition vertices in such a way that the orbit partition must either be equal to the algorithm's partition or be a refinement of it. However, when it does fail to produce orbits, the partition obtained will not have any specifiable properties; it may not, for example, necessarily conform to regular equivalence.

However, by taking the partition generated by the orbit algorithm as the attribute in REGE/A, we can use REGE/A to find the maximal regular equivalence preserving that structure. This ensures that we work from a model with known and desirable properties.

Similarly, the result of any equivalence-finding algorithm can be "regularized" simply by submitting to REGE/A both the data matrix (as the primary relation) and the current equivalence matrix (as the point-similarity matrix).

An interesting application of REGE/A is to test an a priori partition. For example, suppose a biologist records interactions between members of a small colony of organisms. He builds up a model of how many types of members there are and which individual represents what type. For example, he might hypothesize four types: workers, guards, nurses and queens. The question is, does this partition fit our rigorously defined notion of role, namely regular equivalence? The test is easily performed by entering the hypothesized partition as the point-similarity attribute (simply a nominal code indicating which partition each actor is in). The result will be the maximal regular equivalence consistent with that partition. If that equivalence is identical to the hypothesized partition, then the latter is regular and can legitimately be regarded as a role-model. If it is not, in which case the output equivalence is a refinement of the hypothesis, then a new model is called for.

The latter case is useful in its own right. Rather than test an hypothesis, one may be interested simply in finding the maximal regular equivalence that keeps two actors (or two sets of actors) apart. For example, one might wonder whether there exists any regular equivalence such that actor x occupies a different role from actors y and z (who may or may not be playing the same role). This is entered simply by an attribute with the following values:

Actor	Attribute
x	1
y	2
z	2

REGE/A will split y from z if it needs to, but cannot put x and y or x and z together. Obviously, much more complicated role patterns may be hypothesized and tested in this manner.

Extensions

The standard REGE procedure seeks pairs of actors that are alike in having the same kind of alters (a single point-attribute, namely previous-iteration equivalence class) and the same kind of relationship to them (a set of dyad-attributes, namely the set R). Up to this point, the only difference between standard REGE and REGE/A has been that the point-based attributes were extended beyond the previous-iteration equivalence class to include any point-based attribute(s) the analyst might require. No similar extension of the dyadic attributes was made. We now present such an extension:

REGE/A, version 2:
For all i, j, k, $m \in V$, i and j are equivalent at iteration $t+1$ if for all k related to i there exists an m such that
(i) $R^*(i, k) = R^*(j, m)$
(ii) $E_t(k) = E_t(m)$
(iii) $S_q(i, j) = \text{TRUE}$ for all $q \in Q$
(iv) $D_p(i, k) = D_p(j, m)$ for all $p \in P$.

The "P" in condition (iv) refers to a set of general dyadic attributes, such as partial dependencies (Freeman 1979) or geodesic distance. $D_p(i, k)$ is the value assigned to the relationship between actors i and k on the pth dyadic attribute. In short, this says that i and j are equivalent if and only if they have the same kinds of relationships with the same kinds of alters, where both "kinds of relationships" and "kinds of alters" mean equality on specified sets of attributes.

Of course, it is not absolutely necessary that conditions (i) and (ii) be kept separate from (iv) and (iii) respectively. Any dyadic attribute included in condition (iv) could just as well appear as a primary relation in condition (i). However, for conceptual clarity we keep them apart.

The modified REGE/A algorithm may be used to find blockings of actors that are not necessarily regular equivalences. A case in point is the role-equivalence proposed by Winship (1974), Mandel (1978), Mandel and Winship (1979), and Winship and Mandel (1983), and referred to by Mandel as "Winship–Pattison role equivalence". The following definition is taken from Winship and Mandel:

> *Definition*. Two individuals are *role-equivalent* if their role sets contain the same role relations. That is, for every role relation associated with each individual, there is at least one (and perhaps more) identical role relation associated with the other individual. (1983: 324)

As in other work by the "semi-group school" of role modeling (Pattison 1982; Wu 1983; Breiger and Pattison 1986), the approach is based on the semigroup of relations arising from composition of the measured

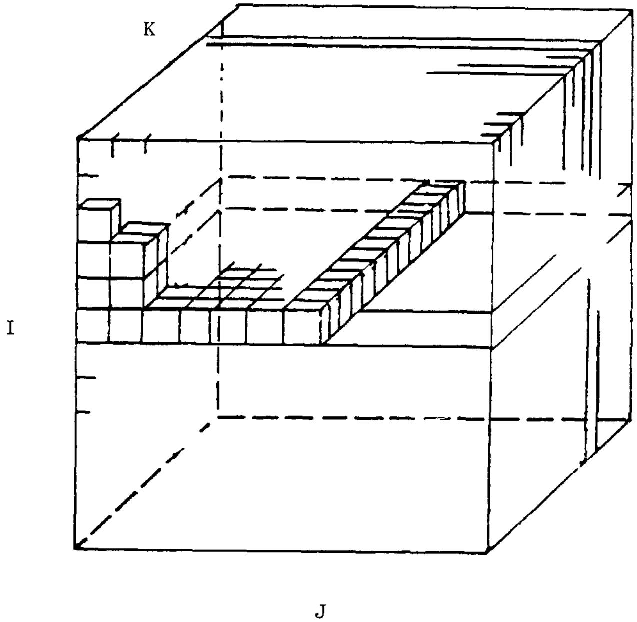

Fig. 8.

relations. The term "role relation" in their terminology is a vector of ones and zero indicating whether an actor has a particular compound relation with a given other. In a network of n actors, each actor has n role relations (including the one with himself). An actor's "role set" (sometimes called "role plane") is merely his collection of n role relations (technically, it is the set of non-identical role relations, but the distinction is irrelevant in this context); hence, it is an $n \times n$ matrix. Since each actor has a role set, there are n of these matrices stacked one on top of the other. In Figure 8, the role sets are represented by horizontal slices or planes through the "relation box". Vertical slices are adjacency matrices for each relation in the semigroup.

Let us label the vertical dimension of the box with "I", the horizontal dimension with "J" and the depth dimension with "K". Slices parallel to the KJ face of the box are role sets. Each column J of the ith KJ slice is a vector of ones and zeros representing the role relation of the ith actor with the jth actor.

According to the definition, two actors are role equivalent if every distinct role relation that one has the other has also, but not necessarily

$$
\begin{array}{c|cccccccccc}
 & & & & & J & & & & & \\
 & a & a & b & b & c & a & b & c & c & a \\
 & a & c & c & a & c & a & a & a & a & a \\
 & b & c & b & c & b & a & b & c & a & b \\
 & a & c & c & a & c & a & a & a & a & a \\
I & b & c & b & c & b & a & b & c & a & b \\
 & b & c & c & b & c & b & b & b & b & b \\
 & d & d & d & d & d & d & d & d & d & d \\
 & a & c & c & c & c & c & c & c & c & c \\
 & b & a & b & a & b & a & b & b & a & b \\
 & b & c & b & c & b & a & b & c & a & b
\end{array}
$$

Fig. 9.

with the same alter, and not necessarily in the same quantity. Formally, we might put in this way:

For all i, j, k, $m \in V$, i and j are role-equivalent if for all k there exists an m such that

Role Relation(i, k) = Role Relation(j, m)

Consider now what happens if we assign a unique code, such as a letter or a color to every distinct role relation. In effect, this collapses the relation box across the K-dimension, as illustrated in Figure 9. The net result is an actor-by-actor matrix whose cells are categorical labels representing the set of relationships obtaining between each actor and every other. Obviously, this matrix can be entered as the dyadic attribute matrix in REGE/A, where the relation in condition (i) is the complete graph. Alternatively, the collapsed matrix may be entered directly in condition (i) as a valued (but categorical) relation. [2] Finally, we need not collapse the relation box at all, but just enter each IJ slice as an additional binary primary relation. The one thing we must do, however, to duplicate Winship and Mandel, is to employ just one iteration. While Winship and Mandel require that equivalent actors have the same relations with others, they do not require that those others be equivalent, which is the sole function of the REGE iterations.

Note, however, that by allowing multiple iterations we can "regularize" the Winship and Mandel equivalence. That is, we can find the

[2] Such a matrix cannot be processed by all REGE routines. One routine that will handle it is the GQ procedure in the AL computer package (Borgatti 1987).

maximal regular equivalence consistent with the Winship and Mandel equivalence.

Computer implementation

Earlier we cited conceptual clarity as the reason for keeping the four conditions of REGE/A separate, but there is another reason as well. When we apply concepts such as regular equivalence to empirically derived data we often run into a problem: no pair of actors are equivalent. A common approach to this problem, and the one taken in standard REGE programs, is to provide a measure of the extent of equivalence. While no-one would claim that the degrees of equivalence produced by REGE are "right" or even particularly interpretable, it does make sense to think twice before mixing relations and attributes.

At the very least we would like to control the relative weight of the attributes versus the relations. Further, the logic by which REGE goes about determining whether two actors are similarly related across multiple relations is not particularly appropriate to a mixture of relations and attributes. Briefly, REGE tries to take a kind of weighted percentage of the relations that are the same for i and j for a given alter. It actually fails but that is incidental here. The question is, should we include the attributes with the relations? What if a pair of points are similar on a couple of attributes but not on the true relation? The percentage of same "relations" would be high so the algorithm would find the pair highly equivalent. Yet we could question that result on the grounds that the relation is what the regular equivalence should be all about: the attributes should merely exist to break apart otherwise equivalent actors. Consequently, we have chosen to keep relations and attributes separate.

While we are considering degrees of equivalence, it should be noted that until now we have required that attributes match exactly. For example, conditions (iii) and (iv) of REGE/A read as follows:

(iii) $S_q(i, j) = \text{TRUE}$ for all $q \in Q$
(iv) $D_p(i, k) = D_p(j, m)$ for all $p \in P$

An obvious relaxation is to allow S_q to take on values between zero and one to indicate the degree of similarity. Similarly, we can evaluate the proportion of dyadic attributes D_p that match, rather than recording only whether they all match. These two adjustments have been implemented in our computer program (Borgatti 1987).

Discussion

We have argued that (a) a graph may contain a plurality of regular equivalences (forming a lattice), (b) situations exist where the maximal regular equivalence is not of interest, and (c) that REGE/A may be used to compute alternative regular equivalences. What we have not argued is why we restrict our search for alternative equivalences to the set of regular equivalences.

In the broadest possible manner, we can describe regular equivalence as the classification of actors by the kinds of relations they have with others. The first iteration of REGE, in fact, merely serves to sort the vertices into sinks, sources, and "repeaters". This description also fits other non-regular equivalences. The Winship and Mandel approach, for example, classifies actors by the sets of compound relations they have with others. If actor i has relations R, R^2, and R^4 with actor b, and R^2 and R^3 with actor c, and actor j has the same sets of relations ($\{R, R^2, R^4\}$ and $\{R^2, R^3\}$) with actors d and e, then i and j are equivalent. Similarly, Breiger and Pattison (1986) compute for each actor the semigroup of relations in which ego is involved, and then compare the semigroups.

In short, many models of role equivalence can be characterized as classifying actors by the patterns of relations they have with others. Where these models principally differ is in the restrictions they place (or do not place) on who the actors are having these relations with. For example, Winship and Mandel place no restrictions. If actor i has role-relation vectors 101 and 219 with alters a and b and actor j has the same vectors with alters c and d, then i and j are equivalent regardless of the relationship between $\{a, b\}$ and $\{c, d\}$. Structural equivalence on the other hand places the ultimate restriction on the alter sets: actors i and j must have the same relations with the *same* alters. The maximal regular equivalence falls between Winship/Mandel and structural equivalences in restrictiveness: it requires that i and j have the same relations to the same *types* of alters. Completing the running example, i and j are equivalent if $R(i, a)$ matches $R(j, c)$, $R(i, b)$ matches $R(j, d)$, and $E(a) = E(c)$ and $E(b) = E(d)$. The minimal regular equivalence that does not require the same alters is automorphic equivalence.

The notion of social role as developed by Linton (1936), Nadel (1957), and Merton (1959) requires both the condition of similarity of relationships, and the restriction on what alters the relationships are with. Linton, for example, focused on reciprocal relations between groups of actors:

> ...the functioning of societies depends upon the presence of patterns for reciprocal behavior between ... groups of individuals. (1936: 113–114)

All of these theorists understand social structure as emerging from relations between classes of people rather than individuals. For example, Nadel, quoting Parsons in part, states:

> We arrive at the structure of a society though abstracting from the concrete population and its behavior the pattern or network (or 'system') of relationships obtaining 'between actors in their capacity of playing roles relative to one another'. (1957: 12)

Thus a model of role must require that actors have similar sets of relations with players of other roles, and this implies regular equivalence since in the homomorphic model representing the sets of regularly equivalent actors and the relationship between them, all the actors associated with one image will have the same relations with a member of a second image. This is not the case in the Winship and Mandel or Breiger and Pattison approach, and therefore we exclude "un-regularized" versions of such models from the list of candidates. And since we do not want to require that people must know all and only the same people to play the same role, we also exclude structural equivalence, except in the sense that any actors that happen to be structurally equivalent must certainly be playing the same role. This restricts the set of proper models of social role to the set of regular equivalences from automorphic equivalence up to the maximal element of the lattice.

References

Borgatti, Stephen P.

1987 *User's Manual for AL: Analytic Language*. Irvine: University of California.

1988 "A comment on Doreian's regular equivalence in symmetric structures." *Social Networks 10*: 265–272.

Breiger, R.L. and P.E. Pattison

1978 "The joint role structure of two community elites." *Sociological Methods and Research 7*: 213–226.

1986 "Cumulated social roles: The duality of persons and their algebras." *Social Networks 8*: 215–256.

Doreian, Patrick

1987 "Measuring regular equivalence in symmetric structures." *Social Networks 9*: 89–107.

Everett, Martin G.

1985 "Role similarity and complexity in social networks." *Social Networks 7*: 353–359.

Everett, Martin G. and Stephen P. Borgatti

1988 "Calculating role similarities: An algorithm that helps determine the orbits of a graph." *Social Networks 10*: 71–91.

Faust, Katherine

1988 "Comparison of methods for positional analysis: Structural and general equivalence." *Social Networks 10*: 313–342.

Freeman, L.C.

1979 "Centrality in social networks: I – Conceptual clarification." *Social Networks 1*: 215–239.

Linton, Ralph

1936 *The Study of Man*. New York: D. Appleton-Century.

Lorrain, F.P. and H.C. White

1971 "Structural equivalence of individuals in networks." *Journal of Mathematical Sociology 1* (January): 49–80.

Mandel, Michael J.

1978 *Roles and Networks: A Local Approach*. Senior honors thesis, Harvard University.

1983 "Local roles and social networks." *American Sociological Review 48*: 376–386.

Mandel, Michael J. and Christopher Winship

1979 "Roles, positions and networks." Presented American Sociological Association meetings, Boston.

Merton, R.K.

1959 *Social Theory and Social Structure*. 2nd ed. Glencoe, IL: Free Press.

Nadel, S.F.

1957 *The Theory of Social Structure*. London: Cohen & West.

1958 *The Foundations of Social Anthropology*. London: Cohen & West.

Pattison, Philippa E.

1977 "On analyzing a blockmodel". Unpublished manuscript.

1982 "The analysis of semiGroups of multirelational systems." *Journal of Mathematical Psychology 25*: 87–118.

Pattison, P.E. and Warren K. Bartlett

1982 "A factorization procedure for finite algebras." *Journal of Mathematical Psychology 25*: 51–81.

Sailer, Lee D.

1978 "Structural equivalence: Meaning and definition, computation and application." *Social Networks 1*: 73–90.

White, Douglas R.
1984 "REGE: A regular graph equivalence algorithm for computing role distances prior to blockmodeling." Unpublished manuscript, University of California, Irvine.
White, D.R. and K.P. Reitz
1983 "Graph and semigroup homomorphisms on networks of relations." *Social Networks 5*: 193–235.
1985 "Measuring role distance: Structural, regular and relational equivalence." Unpublished manuscript, Univ. of California, Irvine.
Winship, C.
1974 "Thoughts about roles and relations." Unpublished paper, Dept. of Sociology, Harvard University.
Winship, C. and M.J. Mandel
1983 "Roles and positions: A critique and extension of the blockmodeling approach," in: S. Leinhardt (ed.) *Sociological Methodology 1983–1984*. San Francisco: Jossey-Bass.
Wu, Lawrence L.
1983 "Local blockmodel algebras for analyzing social networks," in: Samuel Leinhardt (ed.) *Social Methodology 1983–1984*. San Francisco: Jossey-Bass.