Relation Contents in Multiple Networks

RONALD S. BURT AND THOMAS SCHØTT

Department of Sociology, Columbia University

Distinctions among kinds of relations (friendship, advice, intimacy, and so on) are typically ad hoc in empirical research. These ad hoc distinctions among relation contents can be expected to increase the likelihood of equivocal research conclusions. We develop three ideas indicating how standard, well-known, network models of relationship form can be used to clarify relationship content. (a) We begin with an idea for recovering the semantic context in which a relation content occurs. This context is cast as a network of tendencies for contents to be confused for one another and the form of this network—dissected with network models of relation form—holds insights into the ways in which relation contents are understood in a study population. (b) The network concept of structural equivalence is used to define content domains composed of specific relation contents that are substitutable for one another in described relationships. (c) The network concept of network prominence is used to define the ambiguity of contents in described relationships. The proposed perspective is analogous to a linguistic componential analysis of relation content. © 1985 Academic Press, Inc.

THE PROBLEM

The content of relationships is a problem for network analysis. The problem is nicely illustrated in the distinction between naturally occurring

287

relations and analytical relations—the first being the relations in which
people are actually involved, the second being the recreation of relations
for a network analysis. For example, when you go to a colleague for
advice, that interaction occurs in the context of other activities for which
you have sought her out; lunch, cocktails, dinner, committee work,
colloquies, leads to new acquaintances, and so on. Your naturally occurring
relation to this person, your relation to her as it exists in fact, is a bundle
of different interaction elements. Similarly, your naturally occurring re-
lations to other people are bundles of specific interactions, some consisting
of many elements, others containing very few.

    With the notable exception of ethnographers, network analysts rarely
capture the complexity of naturally occurring relations. Their concern
is less the complexity of the typical relationship between a pair of individuals
than it is the complexity of the structure of relations among many individuals
as a system. The relations described are analytical constructs—a relation's
form being its intensity or strength as a tendency to occur and its content
being its substance as a reason for occurring. The form of a friendship
relation, for example, would refer to the intensity or strength of the
relation while its content, its substantive meaning, would be friendship.
Network models of social structure typically describe the form of relations
while taking the content of those relations as a given, an item exogenous
to the model. The most general of these models purport to describe
formal structure in multiple networks among individuals in a system
where each network consists of relations having the same content. The
questions of why certain networks are to be distinguished in a system,
how individuals within the system interpret their relations, and how they
distinguish different kinds of interaction are assumed to be resolved a
priori.

    Unhappily, these unasked questions are quite unresolved; not only in
general, but in the particular. When someone poses a sociometric question
to you asking for the names of people to whom you go for such and
such, you must disentangle the welter of interactions in your naturally
occurring relations and classify some as such and such before you can
answer the question. If you are asked to name your best friends, for
example, you must decide which of your interactions indicate friendship.
If you are asked to name the people with whom you engage in leisure
activities, you must decide which of your interactions indicate leisure.
If you are asked to name the people who most influence your thinking
with their personal comments, you must decide which of your interactions
reflect influence. Obviously, people can differ in their interpretation of
specific interactions as manifestations of general kinds of relations; some
viewing as intimate, for example, what others view as no more than
friendly. More obviously, people in different social situations or from

different subcultures can differ in their interpretation of specific interactions as indicators of more general relation content.

The distinctions between relation contents needed to formulate sociometric questions before collecting network data are thus unsettlingly ad hoc. The sociometric questions finally selected for a study can be no more than a compromise between the practical impossibility of gathering data on all kinds of relations in which respondents might be involved and the other extreme of initial hunches as to the correct identification of some minimal number of the most significant kinds of relations in a study population.

The data definition problem creates problems for data analysis. Measurement error is an obvious problem. Ad hoc definitions of relation content increase the likelihood of random errors in identifying kinds of interaction in a relationship. If the meaning of friendship is unclear, for example, then respondents will be inconsistent when asked to identify those of their relationships which involve "friendship." As in any data analysis, random errors in network data can be expected to attenuate standardized effects, amplify standard errors, and so suppress evidence of true effects. But the problem is especially troublesome here because network indices are less often used as dependent variables than they are used to predict other variables. This means that random error in network data attenuates standardized and unstandardized effects for network indices. Network analyses based on ad hoc content definitions can be expected to produce equivocal research findings and spurious evidence of insignificant effects.

Beyond random measurement error, there are validity problems. Ad hoc content definitions increase the likelihood of misinterpreted relations. The kind of relation solicited in an ad hoc sociometric question is likely to be understood by a social scientist in a way distinct from its understanding in a study population. There is the related problem of erroneous inferences from comparative research. Even if identical sociometric questions are posed to individuals in two study populations, there is no guarantee that the questions have identical interpretations in the two populations. What the heroin addict understands to be friendship probably differs from the suburban housewife's friendship. More generally, ad hoc content distinctions make it difficult to analyze the coordination of different kinds of relations. But such analysis is integral, almost by definition, to a description of social structure in a multiple network system. Consider the concept of multiplexity. A naturally occurring relation is multiplex to the extent that multiple contents occur together in the relationship. For example, you have a multiplex relation to your colleague as described in this article's first paragraph. You have social, economic, and collegial interactions with the person. The relationship would be uniplex if you had only one kind of interaction.

But who is to say where one kind of relation stops and another begins? When does a colleague relation become a friendship relation? One observer might decide that each of the above interactions has a unique relation content—social, economic, and collegial—whereupon the relationship would be multiplex. Another observer might only distinguish two contents, kinship and nonkinship, whereupon the described relationship would be uniplex—it would consist of multiple examples of nonkinship interactions. Without evidence of the content distinctions recognized in a study population, these alternative distinctions between contents are ad hoc, raising an important analytical question: When is a uniplex relationship mistakenly treated as if it were a multiplex relationship merely because a network analyst has considered various aspects of a single relation content to be different contents? Consequential as a clear understanding of relation content is for describing social structure with network models ranging in sophistication from ego-network multiplexity to multiple network role structures, very little is known about it. Research inferences are correspondingly equivocal.

Fortunately, well-known network models of relation form have the potential to inform studies of relationship content. We explore one strategy, an analogy with linguistic analysis, by which this potential can be exploited. We begin with an idea for recovering the semantic context in which a relation content occurs. This context is cast as a network of tendencies for contents to be confused for one another and the form of this network— dissected with network models of relation form—holds insights into the ways in which relation contents are understood in a study population. Our discussion is methodological. Readers interested in data analysis are referred to a study of substitutability and ambiguity in relationships among several hundred Northern Californians during the late 1970s (Burt and Minor, 1983, Chap. 2).

COINCIDENCE RELATIONS

One key to the meaning of relations lies in the variable tendencies of kinds of interaction to be perceived in the same naturally occurring relations. Confronted with a sociometric question, a respondent must sort through his relationships and identify those containing one or more of the specific contents solicited by the question. By confronting the person with repeated questions asking him to identify specific contents in his relationships, we can see variable tendencies for the same relationships to be identified in response to different questions. In other words, we can see variable tendencies for specific relation contents to be confused for one another in naturally occurring relations.

More specifically, suppose that a set of questions are put to a respondent, eliciting a pool of N naturally occurring relationships and K arbitrarily distinguished relation contents. The N relationships could be with people,

groups, or formal organizations. The K contents could be kinds of interactions (e.g., friendship, advice, socializing) or attributes believed to flavor the meaning of relationships (e.g., the race, age, or occupation of people named as the object of relationships). Assume that the N relationships are elicited to represent a reasonable population of the respondent's relationships. Data such as these are available in diverse research designs. The respondent could be an individual in the closed system typically studied with sociometric data (e.g., Moreno, 1960) or a survey respondent in an area probability sample (e.g., Fischer, 1982a) or a subject in a name generating experiment (e.g., Killworth, Bernard, and McCarty, 1984). Note that data on relations to the respondent from others are not necessary here; only the respondent's perceptions of his relationships are needed. Moreover, the respondent could be evaluating personal relationships, relationships in which a corporation is involved, or relationships in which some informal social group is involved.

These data define a (K,K) symmetric matrix of frequency data where element n_{ij} is the number of relationships in which the respondent identified both content i and content j. For example, if i refers to "best friend" and j refers to "elderly," then n_{ij} would be the number of best friends perceived to be elderly, n_{ii} would be the number of people perceived to be best friends, and n_{jj} would be the number of people perceived to be elderly. These frequency data define two classes of variables that we shall discuss as elements in a network of coincidence relations.

Let c_{ii} be the probability that the respondent perceives content i in any one of his N naturally occurring relations. For the purposes here, c_{ii} can be computed as the following ratio of frequencies (see Note 1):

$$c_{ii} = n_{ii}/N,\tag{1}$$

where N is the number of different people the respondent named and n_{ii} is the number of those relationships reported to contain content i. Obviously, the accuracy of c_{ii} as the probability of content i depends on the representativeness of the N elicited relationships. Thus our assumption is that the N relationships are elicited so as to represent a reasonable population of the respondent's relationships. The n_{ii} relationships form a subset of all N relationships, so c_{ii} will vary from a minimum of 0 (no relationship contains content i) up to a maximum of 1 (each relationship contains content i). If content i were friendship, for example, a c_{ii} equal to .67 would indicate that two of any three of the respondent's N relationships could be expected to involve friendship.

Let c_{ij} be the conditional probability that the respondent perceives content j in a naturally occurring relationship he perceives to contain content i. In other words, c_{ij} is the probability of content j being perceived in a relationship known to contain content i. For the purposes here, c_{ij}

can be computed as the following ratio of frequencies:

$$c_{ij} = n_{ij}/n_{ii},$$
(2)

ignoring the easily resolved problem of n_{ii} equal to zero (Note 2) and the less easily resolved problem of missing data (Note 3). The n_{ij} relationships perceived to simultaneously contain contents i and j form a subset of all n_{ii} relationships containing content i, so c_{ij} will vary from a minimum of 0 (no relationships contain both contents) up to a maximum of 1 (content j is perceived in every relationship containing content i). Note that n_{ii} need not equal n_{jj}, so c_{ij} need not equal c_{ji} even though n_{ij} equals n_{ji}. Note also that the coincidence relation c_{ij} is analogous to the familiar regression coefficient predicting the strength of relation j from the strength of relation i (e.g., frequent contact leads to friendship, see Fischer, 1982b; Laumann et al., 1974; Wellman, 1979; for illustrative analyses). This correspondence is not exact, however; a regression coefficient represents a change in probability while a coincidence relation is a conditional probability (Note 4).

COINCIDENCE RELATIONS AS SEMANTIC DATA

We propose that people make distinctions among relation contents in so far as they are able to refer to different people, different relationships, with the contents. Distinct relationships are necessary for cognitive distinctions between relation contents. In the absence of any understanding of a content, some sense of its meaning can be obtained by observing the manner in which the content appears in relationships with other contents that are understood. In the same way that the meaning of a word can be derived in part from the structure of the words combined in sentences containing the word, the meaning of a content can be derived in part from the structure of the contents combined in the relationships in which the content is perceived. Thus, a coincidence relation is a semantic datum. It defines the extent to which one kind of relation, one content, is prominent in the interpretation of another content.

More specifically, the meaning of relation content i is given by the elements in the ith row and column of the (K,K) asymmetric matrix of probabilities defined by Eqs. (1) and (2). The diagonal element c_{ii} measures the extent to which content i is perceived in all relationships. Ceteris paribus, the more often that a respondent perceives content i in his interactions with different persons (i.e., the higher that c_{ii} is), the less clearly content i is defined in terms of a specific kind of relationship. The off-diagonal element c_{ij} measures the extent to which content j is perceived in any relation containing content i. Ceteris paribus, the more often that the respondent perceives content j in any relationship containing content i (i.e., the higher that c_{ij} is), the less likely he will be to think

about content j as something distinct from content i—the more likely he will be to confuse content j for content i.

More importantly, the structure of coincidence relations can be studied with familiar models of network form to reveal insights into the meaning of relationships. To begin, it is possible to identify content domains as general classes of relation content defining kinds of relationships in a study population.

SUBSTITUTABILITY WITHIN CONTENT DOMAINS

When two contents i and j have identical patterns of coincidence relations with other contents, they derive identical meaning from other contents and contribute identical meaning to other contents. To the extent that content meaning is reflected in these interdependences among contents, contents i and j are semantically equivalent elements in relationships, or, more simply, they are substitutable in the sense that they refer to the same kinds of relationships; not the same relationships per se, but the same kinds of relationships.

The Form of Substitutability

More specifically, contents i and j are substitutable to the extent that (a) they have equal tendencies to be perceived in relationships (i.e., $c_{ii} = c_{jj}$), (b) they have equal tendencies to be confused for one another (i.e., $c_{ij} = c_{ji}$), and (c) they have equal tendencies to be confused with any third content (i.e., $c_{ik} = c_{jk}$ and $c_{ki} = c_{kj}$ for all contents k, $i \neq k \neq j$). In the absence of missing data, the first two conditions are determined by the extent to which the two contents occur with equal frequency (Note 5). The third condition is determined, in part, by the extent to which the two contents occur with equal frequency in any relationship containing any third content (Note 6).

Maximizing the analogy with network models of relationship form, it is convenient to adopt a spatial representation of substitutability. To the extent that contents i and j meet the above three conditions for substitutability, the following euclidean distance will equal zero:

$$d_{ij} = \{(c_{ii} - c_{jj})^2 + (c_{ij} - c_{ji})^2 + \sum_k [(c_{ik} - c_{jk})^2 + (c_{ki} - c_{kj})^2]\}^{1/2}, \quad (3)$$

where summation is across all K contents excluding the two being assessed for their substitutability ($i \neq k \neq j$). This equation defines the distance between each pair, i and j, of K contents within a $2 + 2 (K - 2)$ dimensional semantic space.

Two contents separated by zero distance in this space are completely substitutable in describing relationships. The three substitutability conditions correspond to the terms in Eq. (3). (a) When contents i and j have the same probability of appearing in one of the respondent's re-

lationships, c_{ii} equals c_{jj} so the first term in Eq. (3) equals zero. (b) When the contents are equivalently confused for one another, c_{ij} equals c_{ji} so the second term in Eq. (3) equals zero. (c) When the contents are equivalently confused with every other content under consideration, the third, bracketed, term in Eq. (3) equals zero; c_{ik} equals c_{jk} and c_{ki} equals c_{kj} for every other content k.

More generally, the domain containing contents i and j contains all contents substitutable for i and j. So defined, a content domain is a maximal complete set of substitutable contents—it contains all and only those contents that are substitutable for one another. Content domains are distinct points in the semantic space defined by Eq. (3), each representing a unique mixture of contents in described relationships, a unique meaning that cannot be substituted for the meaning of other content domains. Underlying the K arbitrarily distinguished contents are one or more domains of relation content, each domain containing one or more substitutable contents.

Detecting Content Domains

In order to detect these content domains one could use Eq. (3) to compute a (K,K) distance matrix and search for values of d_{ij} equal to zero. Each set of contents $(i, j, ..., k)$ separated by zero distance $(d_{ij} = ... = d_{ik} = d_{jk} = 0)$ would constitute a content domain.

In practice, however, this strategy is inadequate. Diverse research errors such as poorly constructed questionnaires, respondent fatigue, coding errors, and so on are likely to distort the estimation of coincidence relations. Moreover, colloquial descriptions of relationships are inconsistent over time. Words used to describe a content in today's relationships are likely to be replaced with other words at other times—not to mention actual changes in the bundle of contents constituting a relationship. For reasons of research errors, living language variations in the use of words, or the shifting composition of relationships positive distances between substitutable contents are to be expected.

Instead of restricting content domains to be points in a semantic space defined by the distances in Eq. (3), therefore, domains should be defined as discernible areas, fields, in the space. So defined, a content domain consists of all contents i and j that are separated by negligible distance

$$d_{ij} \leq \mathbf{d}, \tag{4}$$

where $\mathbf{d}$ is a criterion of negligible distance. If $\mathbf{d}$ equals 0, then the domain consists of completely substitutable contents. The higher $\mathbf{d}$ is, the more equivocally domains are defined in terms of specific kinds of relations. With respect to a semantic space, high values of $\mathbf{d}$ would allow domains to include large areas and so an increasing variety of contents distributed in the space.

Domains of substitutable contents can thus be detected by the same methods used to detect statuses of structurally equivalent actors. We have merely adopted the formal concept of structural equivalence to define substitutability. Contents i and j are substitutable to the extent that they are structurally equivalent in a matrix of coincidence relations. That is, substitutable contents are syntactically equivalent in respondent descriptions of relationships. When **d** in Eq. (4) is zero, contents i and j are equivalent under a strong criterion and a positive **d** would define contents equivalent under a weak criterion. The cluster and factor analysis procedures used to locate actors structurally equivalent in social networks can be applied in the same way to locate substitutable contents in a matrix of coincidence relations (e.g., see Burt, 1982:Chaps. 2,3; Burt and Minor, 1983:Chaps. 13,14 for review and numerical illustration) (Note 7).

Coincidence Relations among Content Domains

Let **k** equal the number of different content domains detected in a coincidence matrix. Where c_{ij} is a coincidence relation between contents i and j, let c_{IJ} be a coincidence relation between the Ith and Jth domains of relation content. If each content on which data are obtained is discovered to be nonsubstitutable for every other, then each would define its own content domain. In other words, K would equal **k** and each c_{ij} would correspond to a c_{IJ}. But if two or more of the K initial contents are substitutable, i.e., if K is greater than **k**, then coincidence relations between some content domains I and J would be a reduction of multiple coincidence relations between specific contents observed within each domain. Reduction of the c_{ij} to the c_{IJ} is straightforward and should be completed before formal network models are used to interpret contents (Note 8).

Where c_{ii} is the probability with which content i occurs in a respondent's relationships, c_{II} is the probability with which interactions indicating the Ith domain of content occur in his relationships. The following ratio of frequencies defines the latter probability (Note 9):

$$c_{II} = n_{II}/N, \tag{5}$$

where N is the number of different relationships named by the respondent (cf. Eq. (1)), and n_{II} is the number of those relationships that contain any one of the substitutable contents within domain I. If data were obtained on two contents i and j within domain I, then n_{II} would equal $n_{ii} + n_{jj} - n_{ij}$, where n_{ii} n_{jj}, and n_{ij}, respectively, would be the number of the respondent's relationships that contain the ith, the jth, and both contents. If "discussing personal matters" and "socializing" were the contents observed within a friendship content domain I, then the probability of friendship being a content in a respondent's relationships (c_{II}) would

equal the number of people named by the respondent as individuals with whom he socializes or discusses personal problems (n_{II}) divided by the total number of people he named (N). If data were obtained on three contents i, j, and k within domain I, then n_{II} would equal $n_{ii} + n_{jj} + n_{kk} - n_{ij} - n_{ik} - n_{jk} + n_{ijk}$, where n_{ijk} would be the number of the respondent's relationships that contain all three contents i, j, and k. More generally, c_{II} is defined by the union of contents within domain I (Note 10). Of course, if data were obtained on only one content i within domain I then n_{II} would equal n_{ii}, so c_{II} would equal c_{ii} in Eq. (1).

Interdomain coincidence relations are also defined by the union of specific contents. Corresponding to c_{ij} in Eq. (2), let c_{IJ} be the conditional probability of an interaction with domain J content occurring in a relationship containing an interaction of domain I content. The following ratio of frequencies defines the interdomain coincidence relation (Note 11).

$$c_{IJ} = n_{IJ}/n_{II}, \tag{6}$$

where n_{II} is given in Eq. (5) and n_{IJ} is the number of relationships in which any content domain I interactions occur with any content domain J interactions. For example, suppose that data are obtained on two contents—discussing personal problems and socializing—within a friendship content domain, domain I, and data are obtained on two contents—job-related advice-seeking and supervision—within a work content domain, domain J. The denominator in Eq. (6), n_{II}, would be the number of different people named by the respondent as individuals with whom he discusses personal problems or socializes. The numerator in Eq. (6), n_{IJ}, would be the number of people with whom he has any of four combinations of interaction: discussing personal problems and seeking advice about his job, discussing personal problems and work supervision, socializing and seeking advice about his job, or socializing and work supervision. The ratio of n_{IJ} to n_{II}—that is to say, c_{IJ} in Eq. (6)—would then equal the probability of a work content appearing in a relationship containing a friendship content (Note 12). Naturally, if data were only obtained on a single content i within domain I and a single content j in domain J, then n_{IJ} would equal n_{ij}, so c_{IJ} would equal c_{ij} in Eq. (2).

Subculture Content Domains

The ideas proposed for studying the way in which an individual describes his relationships can be used to study the ways in which members of different subcultures describe their relationships. Chicano relationships can be compared to those of blacks. The content of a failure's relationships can be compared to the content of a successful person's relationships. The Catholic's view of relationships can be compared to the Protestant's. In fact, relation contents can be a guide to detecting subcultures—two

subcultures are distinct to the extent that distinct coincidence matrices
are observed in each subculture.

The meanings of interaction within a subculture can be inferred from
the structure of coincidence relations typical of the subculture. Data on
respondent m, one of M persons sampled from a subculture, will be
indexed with a subscript m. Respondent m's tendency to perceive content
j in a relationship containing content i, $c_{ij(m)}$, is defined in Eq. (2). The
subcultural tendency to see content j in a relationship containing content
i, c_{ij}, is then the expected value of respondent specific tendencies (Note
13):

$$c_{ij} = \sum_m c_{ij(m)}/M.$$ (7)

At a higher level of abstraction, respondent m's tendency to perceive
the Jth domain of content in a relationship containing the Ith, $c_{IJ(m)}$, is
defined in Eq. (6). The subcultural tendency for interactions of content
J to be perceived in relationships containing domain content I is then
the expected value of these respondent specific tendencies:

$$c_{IJ} = \sum_m c_{IJ(m)}/M,$$ (8)

where content domains I and J are defined across individual respondents
within the subculture.

The distance data needed to detect substitutable contents within sub-
culture content domains can be obtained in several ways. Two are il-
lustrative. The simpler is to compute distances from average coincidence
relations. Subculture distances between contents can be computed from
Eq. (3) using the average coincidence relations defined in Eq. (7). Al-
ternatively, subculture distances can be computed directly from each
respondent's coincidence relations. The distance between contents i and
j within the semantic space created by respondent m's description of his
relationships, $d_{ij(m)}$, is defined in Eq. (3). The square root of the sum
across respondents of these distances squared is the euclidean distance
between contents i and j across all respondents sampled from the subculture:

$$d_{ij} = \left[\sum_m d_{ij(m)}^2 \right]^{1/2}.$$ (9)

This equation defines the distance between each pair, i and j, of K
contents within an $M[2 + 2(K - 2)]$ dimensional semantic space. Distances
computed from the mean coincidence relations for a subculture define
the distance between each pair, i and j, of K contents within a $2 + 2(K -
2)$ dimensional semantic space. Dimensions increase by a factor of M in
Eq. (9) to register differences in the way that people in different subcultures
describe their relationships. These differences can be striking and are
ignored by distances based on coincidence relations averaged across

individuals (Note 14). The distance in Eq. (9) is the better definition of
content substitutability whenever individuals within a subculture are noted
to differ in their descriptions of relationships (See, however, Note 13).

Once defined, subculture distances could be analyzed to detect sub-
stitutable contents within subculture content domains in the same way
that respondent specific distances would be analyzed to detect respondent
specific content domains. Relations of each domain content define a
distinct network among people within the subculture, the total number
of distinct networks in the subculture equaling the number of content
domains its members distinguish (Note 15). Illustrative analysis of content
domains and meanings across social groups is provided elsewhere (Burt
and Minor, 1983:49–66).

CONTENT AMBIGUITY

We have argued that the meaning of a relation content can be derived
in part from the pattern of coincidence relations linking the content to
others. Having distinguished contents with nonsubstitutable meanings,
we turn to the problem of describing those meanings.

One aspect of meaning indicated by a pattern of coincidence relations
is content ambiguity. A content used to describe multiple kinds of re-
lationships is ambiguous in describing any one kind of relationship. Two
relationships are of different kinds to the extent that they contain different
contents. Thus a content is ambiguous when it occurs in many relationships
with many different contents. Ambiguous contents serve the important
function of communicating general qualities of relationship, indicating
whether a relationship is generally good or generally bad, for example;
but their presence in many different kinds of relationships makes it
difficult to assign them any meaning independent of other contents. The
way in which a content occurs with other contents can be studied to
learn the extent to which it is ambiguous in descriptions of relationships.
In a sentence, a content is ambiguous to the extent that it appears in
diverse relationships containing ambiguous contents. This ostensibly cir-
cular idea is greatly clarified with a little algebra.

The Form of Ambiguity

Ambiguity has absolute and relative qualities. Let u denote the ambiguity
of the most ambiguous of k content domains. We will use this maximally
ambiguous content as a numeraire to express the relative ambiguity of
other contents. Let g_I be the ratio of content domain I ambiguity to
numeraire ambiguity. Content I is less ambiguous than the numeraire by
definition, so g_I is a fraction. Specifically, it is the fraction by which
numeraire ambiguity would be multiplied in order to express the ambiguity
of interactions with domain I content. In other words, the absolute am-

biguity of content I is the following product:

$$g_I u, \tag{10}$$

and the relative ambiguity of content I is the ratio of absolute ambiguity over numeraire ambiguity;

$$g_I = (g_I u)/u. \tag{11}$$

Content ambiguity is generated by the ambiguity of the contents with which it appears in relationships. The conditional probability of interactions with domain I content appearing in a relationship containing content J is defined in Eq. (6) as c_{JI}. Content I is therefore ambiguous to the extent that high values of c_{JI} occur with high values of g_J. Summing the product of these two terms across all k content domains J yields a fractional measure of content I ambiguity;

$$\sum_J c_{JI} g_J. \tag{12}$$

Equations (10) and (12) both define measures of content I ambiguity, Eq. (10) with respect to the ambiguity of other contents and Eq. (12) with respect to the circumstances generating ambiguity. Bringing the two measures together yields the following equation defining the form of content ambiguity:

$$g_I u = \sum_J c_{JI} g_J. \tag{13}$$

Numeraire ambiguity is defined in Eq. (13) as follows: $u = c + \sum_J c_J g_J$, where the summation across all k contents J excludes the numeraire (recall that the relative ambiguity of the numeraire, Eq. (11), equals 1). This means that numeraire ambiguity is high (i.e., u is much greater than zero) to the extent that numeraire content interactions appear in many relationships (i.e., the numeraire's diagonal coincidence element, c, is much greater than 0) and appear in relationships containing other interactions of especially ambiguous content (i.e., c_J, the coincidence relation from content domain J to the numeraire domain, is much greater than 0 at the same time that domain J ambiguity, g_B, is much greater than 0). More generally, the ambiguity of content domain I is defined as follows:

$$g_I = (c_{II} g_I + \sum_J c_{JI} g_J)/u,$$

where the summation is across all k domain contents J, excluding I. In other words, the ambiguity of interactions with domain I content will be low (i.e., g_I close to 0) to the extent that they very rarely appear (i.e., c_{II} close to 0) and rarely appear in relationships containing other contents, especially when those other contents are ambiguous (i.e., c_{JI} close to 0

for all contents J for which g_J/u is large) (Note 16). In short, Eq. (13) is the mathematical expression of an earlier verbal statement: A content is ambiguous to the extent that it appears in diverse relationships containing ambiguous contents.

Once again clarifying the tie to formal network models, Eq. (13) corresponds to the eigenvector models used to define prestige and centrality in social networks (e.g., see Burt, 1982:35–37; Burt and Minor, 1983:Chap. 10, for review and illustration). In the same sense that a central, prestigious person is prominent in a network of relations among other individuals, an ambiguous content is prominent in a network of coincidence relations among other contents. That is to say, a content is ambiguous to the extent that it occupies a prominent semantic position in descriptions of relationships. While eigenvectors provide the most elegant model of network prominence, their use is sensitive to the structure of the network from which scores are derived.

Before closing with guidelines for computing ambiguity scores, we wish to emphasize how limited our exploration of formal models has been. Like the prominence model used here, there are a great variety of network models that describe observed forms of relations in comparison to theoretically significant idealized forms. Examples are multiplexity, range, structural autonomy, power, and so on; not to mention the diversity of alternative network centrality and prestige models. Each of these formal models offers insights into a content's meaning when used to study the form of the content's coincidence relations.

Obtaining Ambiguity Scores

Matrix notation is convenient here. The k equations represented by Eq. (13), one for each content domain I, can be expressed as the following matrix equation;

$$uG = GC,$$

where C is the (k,k) matrix of coincidence relations and G is a row vector containing the k content ambiguities defined in Eq. (11), i.e., $G = (g_1, g_2, \ldots, g_k)$. The equation can be rewritten to the following;

$$0 = G(C - uI), \tag{14}$$

which is the characteristic equation for the matrix C where I is a (k,k) identity matrix, u is the maximum eigenvalue for C, and G is its corresponding left-hand eigenvector. The maximum eigenvalue is appropriate here because it equals numeraire ambiguity, the maximum possible. Note that domain-level coincidence relations can be defined for individuals (Eqs. (5), (6)) or subcultures (Eq. (8)), so content ambiguity can be analyzed as an individual or subcultural phenomenon. Equation (14)

provides an unequivocal definition of content ambiguity as long as a satisfactory, unique, solution exists (Note 17).

If the coincidence relations *cannot*, by some reordering of rows and columns in C, be reduced to the form

$$C = \begin{bmatrix} A & 0 \\ 0 & B \end{bmatrix},$$

with A and B square, then Eq. (14) has a dominant, positive eigenvalue u (meaning that u is greater than $|v|$ for any other eigenvalue v for the matrix C) and a unique corresponding left-hand eigenvector G composed of positive elements (e.g., see Gantmacher, 1954:65). This eigenvector is a satisfactory solution yielding positive content ambiguities unique to a factor, here chosen to be the dominant eigenvalue u used to denote numeraire ambiguity.

If the coincidence relations *can* be reduced to the above form of two (or more) square matrices A and B, then there are two (or more) independent subsystems of domain contents in the described relationships. These subsystems should be analyzed separately. For contents in subsystem A, ambiguities can be obtained from the characteristic equation $0 = G_A(A - u_A I)$, and ambiguities for the contents in subsystem B can be obtained from the characteristic equation $G_B(B - u_B I) = 0$. Ambiguities within each subsystem would be obtained from Eq. (11) using the appropriate maximum subsystem eigenvalue, u_A or u_B. The relative magnitudes of u_A and u_B would indicate the relative ambiguity of contents on average in the two subsystems, the larger indicating the more ambiguous.

Separating the content subsystems preserves a complete description of content ambiguities. Simultaneous analysis of all k contents in a reducible C matrix would result in all contents in the less ambiguous subsystem receiving ambiguity scores of zero. If u_A were greater than u_B, for example, the ambiguities in Eq. (14) for subsystem B contents would be 0. Clearly, this is less informative than the proposed procedure.

CLOSING

This article was motivated by the typically ad hoc distinctions made among kinds of relationships in empirical research. These ad hoc distinctions among relation contents increase the likelihood of equivocal research conclusions. We developed three ideas—coincidence matrix, content substitutability, and content ambiguity—to illustrate a strategy by which standard, well-known, network models of relationship form could be used to improve the treatment of relationship content in empirical research.

A matrix of coincidence relations was used to represent the semantic context in which relation contents are interpreted within a study population. Element c_{ii} is the probability of content i being perceived in a relationship and element c_{ij} is the conditional probability of content j being perceived

in a relationship containing content *i*. We proposed that people make distinctions among relation contents in so far as they are able to refer to different people, different relationships, with the contents. Distinct relationships are necessary for cognitive distinctions between relation contents. In the absence of any understanding of a content, some sense of its meaning can be obtained by observing the manner in which the content appears in relationships with other contents that are understood. In the same way that the meaning of a word can be derived in part from the structure of the words combined in sentences containing the word, the meaning of a content can be derived in part from the structure of the contents combined in the relationships in which the content is perceived. Thus, a coincidence relation is a semantic datum. If defines the extent to which one kind of relation, one content, is prominent in the interpretation of another content.

The structure of coincidence relations can be studied with well-known models of network form to reveal insights into content meaning. We used the concept of structural equivalence to define the extent to which two contents are substitutable for one another in the sense of having the same meaning in described relationships. Methods used to identify statuses of structurally equivalent actors in social networks can be applied in the same way to a matrix of coincidence relations in order to identify content domains, classes of substitutable relation contents that define kinds of relationships in a study population. We used the concept of network prominence to define the ambiguity with which contents are perceived in described relationships. Methods used to obtain eigenvectors of network prestige can be applied to a coincidence matrix in order to estimate content ambiguities. The proposed concepts of content substitutability and ambiguity can greatly simplify the empirical task of distinguishing kinds of relationships in a study population (Burt and Minor, 1983: Chap. 2).

It was initially hoped that established methods of linguistic analysis could be drawn upon. To the extent that the proposed analytical strategy has a linguistics analogy, it seems most analogous to componential analysis. We interpret contents by the pattern of propositions (coincidence relations) linking them to component contents. Kempson (1977:18–20, 86–102) provides a textbook treatment of linquistic componential analysis. Romney and D'Andrade's (1964) empirical study of sememes combining to define the meaning of kinship relations more clearly illustrates the analogy between componential analysis and the network strategy proposed here.

The analogy must be qualified in two important ways. First, the component propositions we use to interpret a relation are continuous and probabilistic rather than the categorical and deterministic propositions typical of componential analysis. Second, and more importantly, we offer no distinction between primitive and derived contents. There is no hierarchy of contents in which some are primitive terms, obtaining their meaning

from universals beyond the contents under study. Each content is endogenous in our analysis, deriving meaning from, and contributing meaning to, each other content under study.

By these differences, our strategy for analyzing relation contents has a definite limit on the profundity of insight it can provide at the same time that it provides better tools for empirical research not aspiring beyond that limit. The limit is hierarchy. The lack of logical propositions between primitive and derived contents means that ours is not a strategy for recovering the hierarchy of elements that generates content meaning. One could never fully understand relation contents with the proposed strategy. At the same time, the strategy provides better tools than established linguistics for empirical research on relation contents. We are relieved of sorting contents into primitive and derived terms, a time-consuming task heavily dependent on personal judgments by the individual performing the task. Moreover, our emphasis on continuous probabilistic propositions frees us from the ultimately unreliable task of sorting contents into one meaning or another. Rather than having to code contents as ambiguous when they have more than one independent meaning, for example (cf. Kempson, 1977, on ambiguity), we measure the degree to which any one content is more ambiguous than other contents. In sum, ours is a strategy for using network models and empirical data to recover the meaning of relation contents in a study population so that networks can be properly specified.

NOTES

[1] There are two special cases that deserve mention here. First, we assume that data on several relationships have been obtained. The smaller that N is, the less variable coincidence relations can be. For an N of 2, coincidence relations have three possible values (0.0, 0.5, 1.0). For an N of 1, coincidence relations have two possible values (0.0 and 1.0). For the respondent acknowledging no relationships, $N = 0$, coincidence relations could be set equal to 0 by definition. No knowledge is gained from such computation, but isolates can occur in a large sample of respondents so there is value in anticipating such cases. The second special case to be acknowledged is missing data. Suppose that N people were named by a respondent, but only N^*_i of those people have nonmissing data on sociometric question i. In other words, N^*_i, not N, is the upper limit for n_{ii}. In keeping with the probability interpretation of c_{ii}, the following ratio of frequencies with missing data gives the number of times that content i was perceived in any relationship in which it could have been perceived: $c_{ii} = n_{ii}/(N^*_i + a)$, where a is the constant in Note 2 (needed when $N^*_i = 0$).

[2] All contents will not occur in every respondent's relations. In other words, values of n_{ii} equal to 0 are to be expected. In computer programs generating coincidence relations, Eq. (2) should include a constant, $a = 1/10^r$, in the denominator; $c_{ij} = n_{ij}/(n_{ii} + a)$ to eliminate the definitional problem of dividing by zero, where the exponent r is one larger than the number of decimal places to which coincidence relations will be rounded. For example, $1/10^4 = 0.0001$ would be an appropriate value for a if coincidence relations were to be rounded to three decimal places.

[3] Of alternative, reasonable, computations allowing for missing data, we have yet to see the following produce odd results (such as probabilities greater than one): $c_{ij} = n_{ij}/(N^*_{ij} + a)$,

where a is the constant in Note 2 and N^*_{ij} is the number of relationships containing content i and nonmissing data on content j. In other words, N^*_{ij} is the upper limit for n_{ij} so the ratio of n_{ij} to $(N^*_{ij} + a)$ is the number of times that content j was perceived in a relationship containing content i over the number of times that content j could have been perceived in a content i relationship.

[4] We wish to make the analogy between coincidence relations and regression coefficients explicit because the latter are so readily available. Let each of the N persons named by a respondent be a unit of analysis. In other words, let each observed naturally occurring relation be a unit of analysis. Let x_{jn} be a binary variable equal to 1 if content j appears in the nth relationship and let it equal 0 if such interaction does not appear in the relationship. Let x_{in} be a similarly defined variable for content i. Ordinary least squares estimates of parameters in the following equation could be computed:

$$x_{jn} = b + b_{ji}x_{in} + e_n,$$

where b, b_{ji}, and e_n, respectively, are regression intercept, coefficient, and residual in predicting the appearance of content j from the appearance of content i in relationship n. The intercept b is the probability that content j appears in a relationship in which content i is absent. The coefficient b_{ji} is the change in that probability associated with the appearance of content i in the relationship. The coefficient is defined as follows:

$$b_{ji} = \left[\sum_n (x_{jn} - \bar{x}_j)(x_{in} - \bar{x}_i) \right] \Big/ \left[\sum_n (x_{in} - \bar{x}_i)^2 \right],$$

where summation is across N relationships, $\bar{x}_j$ is the mean of x_{jn} across all N relationships, and $\bar{x}_i$ is similarly the mean of x_{in}. Rewriting the definition in terms of sociometric citation data yields the following:

$$b_{ji} = [n_{ij} - n_{ii}n_{jj}/N - n_{ii}n_{ij}/N + n_{ii}n_{jj}/N]/[n_{ii} - 2n^2_{ii}/N + n^2_{ii}/N]$$
$$= [n_{ij} - n_{ii}n_{jj}/N]/[n_{ii} - n^2_{ii}/N],$$

where $\bar{x}_i = n_{ii}/N$ and $\bar{x}_j = n_{jj}/N$. This can be rewritten to express the difference between the regression coefficient and the coincidence relation:

$$b_{ji} = n_{ij}/n_{ii} - (n_{jj}/N - b_{ji}n_{ii}/N)$$
$$= c_{ij} - (\bar{x}_j - b_{ji}\bar{x}_i)$$
$$= c_{ij} - b,$$

so c_{ij} equals b plus b_{ji}. In other words, the regression coefficient b_{ji} is the conditional probability of j given i, minus the probability of j occurring in the absence of i. Equivalently, the coincidence relation c_{ij} is the zero-order regression coefficient predicting the appearance of j from the occurrence of i, plus the mean tendency for j to appear in the absence of i.

[5] The diagonal elements c_{ii} and c_{jj} equal n_{ii}/N and n_{jj}/N, respectively, and so are equal—ignoring the possibility of missing data—when i and j occur with equal frequency; $n_{ii} = n_{jj}$. The off-diagonal elements c_{ij} and c_{ji} equal n_{ij}/n_{ii} and n_{ji}/n_{jj}, respectively, and so are equal—again ignoring the possibility of missing data—when i and j occur with equal frequency since n_{ij} by definition equals n_{ji}. This second criterion merits a brief note. At first glance, it seems reasonable to require substitutable contents to be completely coincident with one another in the sense that c_{ji} and c_{ij} both equal their maximum value, 1. Not only does this trivialize the idea of content substitutability (cf. Note 7), it is very sensitive to questionnaire design. For example, a respondent could seek advice from superiors and subordinates where he works. This interaction would give his work relations an advice-seeking content. But suppose that two sociometric questions were used to solicit advice-seeking relations from the respondent, question i asking about superiors and question j asking about subordinates. A single person could not be named on both questions so c_{ij} and c_{ji} would equal

zero, implying that advice-seeking from superiors in no way means the same thing as advice-seeking from subordinates. This might be true, but in this instance the appearance of such a fact has been forced upon the data by question wording. The same problem would arise if interaction with living ancestors was solicited by a question separate from the question soliciting interaction with living descendants. More generally, the problem arises when sociometric questions solicit kinds of interaction with people who have particular attributes (e.g., superiors, subordinates, older than respondent, younger than respondent) or interaction in particular contexts (at home, at work, etc.). We are grateful to William Batchelder for calling attention to this issue during a colloquium discussion of content substitutability.

[6] This statement refers to the two sums in Eq. (3). The second sum (differences between c_{ki} and c_{kj} for varying k) is zero when i and j occur with equal frequency in any relationship containing some other content k. Since c_{ki} and c_{kj} equal n_{ki}/n_{kk} and n_{kj}/n_{kk}, respectively, they will be equal when n_{ki} equals n_{kj}. The first sum (difference between c_{ik} and c_{jk}) does not reduce to a direct equivalence of frequencies. Since c_{ik} and c_{jk} equal n_{ik}/n_{ii} and n_{jk}/n_{jj}, respectively, they are only equal when the frequency with which i and k occur together—relative to the overall frequency of i—equals the frequency with which j and k occur together—relative to the overall frequency of j.

[7] Of the two concepts used to define network subgroups, cohesion and structural equivalence, there are several reasons for preferring structural equivalence (e.g., see Burt and Minor, 1983:Chap. 13), but the preference is particularly sharp with respect to content substitutability. Structural equivalence groups together individuals who have similar patterns of relations with all other individuals in a network. Cohesion groups together individuals who have strong relations with one another. With respect to a matrix of coincidence relations, two contents i and j would be grouped together as cohesive to the extent that they were complete coincident ($c_{ij} = c_{ji} = 1$). This criterion is only met when the two contents always occur with one another in the same relationships. Such a criterion would define a trivial form of content substitutability. It is more an indicator of poor questionnaire construction. The ith and jth sociometric questions refer to identical relationships and so neither offers a gain in information over the other (cf. Note 5). The structural equivalence criterion of substitutability is more suited to the purpose of studying relation content, with respect to coincidence relations at least. Two contents i and j need never appear in the same relationships in order to be structurally equivalent (i.e., c_{ij} and c_{ji} could equal 0). In order to be structurally equivalent, however, they would have to appear to the same extent with every other content. In other words, structurally equivalent contents describe the same kinds of relationships even if they do not describe the same specific relationships. Putting methodological advantages aside, in short, structural equivalence is conceptually better suited to detecting substitutable contents. Structural equivalence identifies contents that are semantically substitutable while cohesion would identify semantically redundant contents.

[8] The application of formal network models to the C matrix can be distorted by the presence of substitutable contents in the matrix (e.g., see Note 16).

[9] The issues raised in Note 1 with respect to computing c_{ii}, low N and missing data, apply in the same way to Eq. (5).

[10] Continuing the analogy with structural equivalence models, it might seem reasonable to compute the density of coincidence relations among substitutable contents. This is the usual procedure by which group relations are obtained from relations among structurally equivalent actors. Densities, or weighted averages, are not appropriate to the probabilistic meaning of c_{II} in Eq. (5). Routinely computed, an intradomain coincidence relation would be the average c_{ij} between all contents i and j substitutable within domain I;

$$\sum_i \sum_j c_{ij}/I^2,$$

where I equals the number of contents i and j within content domain I. As an average of probabilities (diagonal elements of C) and conditional probabilities (off-diagonal elements of C), this is a nonsensical computation (excluding the trivial case of a domain based on a single sociometric question, whereupon the average would equal c_{ii} in Eq. (1) and c_{II} in Eq. (5)). If only the off-diagonal elements were summed,

$$\sum_i \sum_j c_{ij}/I(I - 1), \; i \neq j,$$

it would not capture the probability of interactions with domain content I occurring in the respondent's relationships. It would instead measure their tendency to occur together. Nor would an average of diagonal elements, $\Sigma_i c_{ii}/I$, capture the probability of interactions with domain content I. It would instead measure the average probability of each content within the domain occurring in the respondent's relationships,

$$\sum_i c_{ii}/I = \sum_i (n_{ii}/I)/N = \bar{n}_{II}/N,$$

where $\bar{n}_{II}$ is the average number of different relationships of content I elicited from the respondent. For example, an $\bar{n}_{II}$ of 2 would mean that the respondent named two people on average in answering each of the sociometric questions indicating domain I content. This average is replaced in Eq. (5) with n_{II}—the number of different people with whom the respondent has domain I content relationships—in order to preserve the probabilistic meaning of c_{II}. As illustration, suppose that discussing personal problems and socializing were substitutable friendship contents for a respondent and four of his ten relationships contained friendship; two people were named as individuals with whom he socialized and two others were named as individuals with whom he discussed personal problems. The average number of people named on friendship contents is 2. The average diagonal coincidence element is 0.2 for the domain ($\bar{n}_{II}/N = 2/10 = 0.2$). But this is not the probability of friendship. Four of the respondent's ten relationships involve friendship. Friendship therefore has a 0.4 probability of occurring in a relationship and c_{II} in Eq. (5) equals 0.4 for this respondent.

[11] Problems created by missing data and n_{II} of zero can be handled in Eq. (6) in the same ways that they were handled with respect to Eq. (2). See Note 2 regarding zero n_{II} and Note 3 on missing data.

[12] This probability is not captured by the density of coincidence relations between contents within each domain for the same reasons that the density of diagonal coincidence matrix elements is an inappropriate measure of c_{II} (cf. Note 10). The point is worth demonstrating because densities are so often used to represent relations among structurally equivalent actors. The density of coincidence relations from content domain I to J is the average relation from any content in domain I to any content in domain J, $\Sigma_i \Sigma_j c_{ij}/IJ$, where i is one of the I contents with domain I content and j is one of the J contents with domain J content. Rewritten in terms of sociometric citations, this density is given as follows:

$$\sum_i \left(\sum_j n_{ij}/n_{ii} \right)/IJ = \sum_i \left[\sum_j (n_{ij}/J)/n_{ii} \right]/I = \sum_i (\bar{n}_{iJ}/n_{ii})/I,$$

where $\bar{n}_{iJ}$ is the mean number of persons cited by the respondent as the object of a relationship containing content i and any one of the J contents in content domain J. For example, if the respondent named five persons as the object of interaction i, and each was named as the object of a different one of five kinds of substitutable contents within domain J, then n_{ii} would equal 5 and $\bar{n}_{iJ}$ would equal 1 (an average of one person was named as the object of content i and any one of the contents in domain J; $\bar{n}_{iJ} = \Sigma_j n_{ij}/J = 5/5 = 1$). The simplest case is sufficient to illustrate our point here. Suppose that content domain I consists solely of context i. The coincidence relation from i to each of the J contents is $1/5$ so the density of coincidence relations from domain I to domain J is 0.2

(also given by $\bar{n}_{ij}/n_{ii} = 1/5 = 0.2$). This is not the probability of a content J interaction appearing in a relationship containing a content I interaction. Each of the five relationships containing content I interaction contains an interaction with content J. Content J interaction therefore has a 1.0 probability of occurring in a relationship containing content I interaction and 1.0 is the value for this respondent of c_{IJ} in Eq. (6).

[13] As a specification of subcultural tendency for one content j to be perceived in another content i, the expression in Eq. (7) is appropriate only when the respondents in the subculture are rather homogeneous in their coincidence relations $c_{ij(m)}$. If they are heterogeneous, however, it seems more appropriate to specify the subcultural tendency for content j to be confused for content i as the ratio of the total number of relations containing both j and i to the total number of relations containing i, for the respondents in the subculture

$$c_{ij} = \sum_m n_{ij(m)} / \sum_m n_{ii(m)}$$

summing over the respondents in the subpopulation; $n_{ii(m)}$ is respondent m's number of relations with content i, and $n_{ij(m)}$ is his number of relations with both contents i and j, cf. Eq. (2). Similar consideration applies to specification of subcultural coincidence relations among content domains, cf. Eq. (8). Furthermore, this specification of subcultural specific coincidence relations among contents can be inserted in Eq. (3) to compute subculture-specific distances between contents in a way different from the two ways discussed in the following text, cf. Eq. (9).

[14] This point is easily illustrated. Consider three contents and a subculture from which two respondents have been sampled. Suppose that contents i and j have equal probability of appearing in either respondent's relationships ($c_{ii} = c_{ij}$ and $c_{ji} = c_{ij}$). Hence, these elements have no effect on d_{ij}. Suppose further that the first respondent has values of 0.00 and 1.00 for two coincidence relations; c_{ik} and c_{jk}. In other words, he never perceives contents i and k in a relationship together but content k appears in every one of his relationships containing content j. Finally, suppose that the opposite is true of the second respondent; content k is perceived in every one of his relationships containing content i ($c_{ik} = 1.00$) while contents j and k never appear in the same relationships ($c_{jk} = 0.00$). With respect to their tendencies to be coincident with content k, contents i and j are clearly not substitutable in either respondent's described relationships. The distance in Eq. (9) would be positive, reflecting the nonsubstitutability. With respect to c_{ik} and c_{jk}, d_{ij} in Eq. (9) would equal the following: $d_{ij} = [(0.00 - 1.00)^2 + (1.00 - 0.00)^2]^{1/2}$ which is the square root of 2. In contrast, a distance based on mean coincidence relations would indicate complete substitutability. The mean value of c_{ik} and c_{jk} is one half, $(0.00 + 1.00)/2$ and $(1.00 + 0.00)/2$, respectively. The distance computed from these mean relations would then be zero; d_{ij} in Eq. (3) equals $[(0.5 - 0.5)^2]^{1/2}$. Across the many coincidence relations compared with computing distance, in short, subculture mean relations can obscure striking nonsubstitutabilities in the way individuals describe their relationships.

[15] This has an interesting implication: individuals within a subculture should have a content label for each of these networks. If a set of interaction activities seem to reflect a single content domain, then respondents in the subculture containing that domain should have some term they use to identify that content. For example, if shared leisure activities, intimacy, and economic exchanges are found to be substitutable kinds of interaction, then people in the subculture for which such interactions are substitutable should have a word, a role label, for relationships consisting of leisure–intimacy–economic interactions. This provides an interesting strategy for checking on proposed content domain models as well as a strategy for interpreting role relations unique to specific subcultures. We are grateful to Linton Freeman for calling attention to this issue during a colloquium discussion of content domains.

[16] Note that content ambiguity can be distorted by misspecified content domains. Suppose that the substitutable contents within some domain occur with one another in the same relationships. Coincidence relations among them are therefore high. Suppose further that

each content occurs with low probability and rarely occurs with other domain contents. The domain level coincidence relations for this content domain are therefore low; low c_{ii} because its constituent contents occur rarely and low c_{ji} because they rarely occur in conjunction with other contents. Given these low coincidence relations, domain ambiguity is low. But its ambiguity would be inflated by including each of its constituent contents in the C matrix from which ambiguity is derived. In such a misspecified matrix of domain level coincidence relations, there would be high off-diagonal elements among the constituent contents. While the true c_{ji} in Eq. (13) are low for this content domain, in other words, inclusion of substitutable contents would introduce high coincidence relations. This in turn would increase the ambiguity with which the content domain seemed to be used in describing relationships and would inflate numeraire ambiguity. The moral is that ambiguity should be derived from coincidence relations among nonsubstitutable contents, especially when substitutable contents are connected to one another by high coincidence relations.

[17] In the trivial case of some content domain never appearing, we assign it zero ambiguity. This will be consistent with the following where we, for the sake of simplifying the discussion, only consider contents I for which c_{ii} is positive.

REFERENCES

Burt, R. S. (1982). Toward a Structural Theory of Action: Network Models of Stratification, Perception, and Action, Academic Press, New York.

Burt, R. S., and Minor, M. J. (Eds.), (1983). Applied Network Analysis: A Methodological Introduction, Sage, Beverly Hills.

Fischer, C. S. (1982a). To Dwell among Friends: Personal Networks in Town and City, Univ. of Chicago Press, Chicago.

Fischer, C. S. (1982b). "What do we mean by 'friend'? an inductive study," Social Networks 3, 287–306.

Gantmacher, F. P. (1954). Applications of the Theory of Matrices (J. L. Brenner, D. W. Bushaw, and S. Evanusa, Trans. and rev., 1959), Interscience, New York.

Kempson, R. M. (1977). Semantic Theory, Cambridge Univ. Press, New York.

Killworth, P. D., Bernard, H. R., and McCarty, C. (1984). "Measuring patterns of acquaintanceship," Current Anthropology 25, 381–397.

Laumann, E. O., Verbrugge, L. M., and Pappi, F. U. (1974). "A causal modelling approach to the study of a community elite's influence structure." American Sociological Review 39, 162–174.

Moreno, J. L. (Ed.), (1960). The Sociometry Reader, Free Press, New York.

Romney, A. K., and D'Andrade, R. G. (1964). "Cognitive aspects of English kin terms," American Anthropologist 66, 146–170.

Wellman, B. (1979). "The community question: the intimate networks of East Yorkers," American Journal of Sociology 84, 1201–1231.