DIMACS Technical Report 99-35
June 1999

The Role Assignment Model Nearly Fits Most Social Networks

by

Aleksandar Pekeč
The Fuqua School of Business, Duke University
Durham, NC 27708-0120

Fred S. Roberts[1]
Department of Mathematics and DIMACS
Rutgers University, New Brunswick, NJ 08903

[1]Fred Roberts thanks the National Science Foundation for its support under grants SES-9211492 and NSF-SBR97-09134 to Rutgers University

DIMACS is a partnership of Rutgers University, Princeton University, AT&T Labs-Research, Bell Labs, Telcordia Technologies (formerly Bellcore) and NEC Research Institute.

DIMACS is an NSF Science and Technology Center, funded under contract STC–91–19999; and also receives support from the New Jersey Commission on Science and Technology.

ABSTRACT

Role assignments, introduced by Everett and Borgatti [6], who called them role colorings, formalize the idea, arising in the theory of social networks, that individuals of the same social role will relate in the same way to individuals playing counterpart roles. If G is a graph, a k-*role assignment* is a surjective function mapping each vertex into a positive integer $1, 2, \ldots, k$, so that if x and y have the same role, then the sets of roles assigned to their neighbors are the same. We show that all graphs G having no astronomical discrepancies between the minimum and the maximum degree have a k-role assignment. Furthermore, we introduce and study a natural measure expressing how close an onto map $f : V(G) \to \{1, \ldots, k\}$ is to being a k-role assignment of a graph $G = (V, E)$, and show that almost all graphs nearly have a k-role assignment.

1 Introduction

Role assignments, otherwise known as role colorings, provide a graph theoretical model of social role. This paper shows that there are large classes of social networks that can be captured by the role assignment model. Also, a measure of how close this model is to fitting a given social network is introduced, and it is shown that the model nearly fits all but finitely many social networks.

The idea behind the role assignment model is that individuals with the same social role will relate in the same way to individuals playing counterpart roles. This idea was formulated by White and Reitz [15] using the notion of regular equivalence. Regular equivalence, in turn, was motivated by the approach to blockmodeling of social networks introduced by Lorrain and White [9], Arabie, Boorman, and Levitt [2], and Sailer [14], and is aimed in part at providing homomorphic reductions of social networks.

Everett and Borgatti [6] formalized the regular equivalence concept of White and Reitz using a graph-theoretical model, as follows. Let $G = (V, E)$ be a graph with vertices representing individuals and edges representing relationships. An $f : V(G) \to \{1, \dots, k\}$ is a **k-role assignment** of G if f is surjective and if

$$f(v) = f(w) \Rightarrow f(N(v)) = f(N(w)) \tag{1}$$

holds for every $v, w \in V(G)$. Here $N(v)$ stands for the neighborhood of $v \in G$, that is $N(v) := \{u \in V : \{u, v\} \in E(G)\}$, and the notation $f(S)$ is used to denote $\{f(x) : x \in S\}$. Work on role assignments is summarized in the papers by Everett and Borgatti [7] and Roberts and Sheng [12].

In order to express how close a surjective map $f : V(G) \to \{1, \dots k\}$ is to a k-role assignment we count the number of pairs of vertices from G for which (1) holds. Let

$$M(v, w) := \begin{cases} 1 & f(N(v)) = f(N(w)) \\ 0 & \text{otherwise} \end{cases}$$

Then, for a given role $i \in \{1, \dots k\}$, the fraction of pairs of vertices from $V_i := \{v \in V : f(v) = i\}$ for which (1) holds is

$$M_i := \frac{1}{\binom{|V_i|}{2}} \sum_{\{v,w\} \subseteq V_i} M(v, w). \tag{2}$$

One way to measure how close f is to being a k-role assignment is to consider

$$M(f) := \min\{M_1, M_2, \dots, M_k\}.$$

Taking the maximum over all possible $M(f)$ gives a measure of how close G is to being **k-role assignable** (i.e., having a k-role assignment). In other words,

$$\alpha_k(G) := \max\{M(f) : f : V(G) \to \{1, \dots, k\}, f \text{ onto}\}$$

is trivially within $[0, 1]$, and $\alpha_k(G) = 1$ if and only if G is k-role assignable. Note that $\alpha_k(G)$ is not even defined for $k > n$. The problem of determining the exact value of $\alpha_k(G)$ (where k is fixed and G is input data) is NP-hard since even the special case of determining whether $\alpha_k(G) = 1$ is NP-hard (The latter was shown in [13].)

A measure similar to $\alpha_k(G)$ arises in the theory of blockmodeling of a digraph $D = (V, A)$. Here, for a given partition of the vertices $V = V_1 \cup V_2 \cup \ldots \cup V_k$, we measure the fraction M_{ij} of all pairs (x, y) from $V_i \times V_j$ for which there is an arc from x to y in D. We then seek a partition so that the measure M_{ij} is either very small or very large for every i and j. Early algorithms for finding such a partition are due to Breiger, Boorman, and Arabie [4] and Arabie, Boorman, and Levitt [2]. See [10]. Alternative possible definitions of $\alpha_k(G)$ are discussed in Section 6.

This paper is organized as follows. In Section 2, we find constructive bounds for $\alpha_k(G)$ for $k = 2$. We then show in Section 3 that most "large" graphs have $\alpha_k(G) = 1$, i.e., have k-role assignments. This allows us to respond to a question of Everett and Borgatti [8], who ask how common are graphs that have only trivial role assignments, the so-called role primitive graphs. In Section 4, we combine the results of Section 2 and Section 3 to find better bounds for $\alpha_2(G)$ and show that all but finitely many graphs are "almost" 2-role assignable. We extend this analysis in Section 5 and show that all but finitely many graphs are "almost" k-role assignable for arbitrary k. We close the paper with a brief discussion of our results and with some remarks on possible measures of approximate k-role assignability and on possible alternative graph theoretical models of social roles.

Throughout, we use the graph-theoretical terminology of [11]. In particular, the following standard graph notation will also be used throughout: the number of vertices in the graph is $n = |V(G)|$, the degree of a vertex is $d(v) = |N(v)|$, the minimum degree in the graph is $\delta = \min\{d(v) : v \in V\}$, the maximum degree in the graph is $\Delta = \max\{d(v) : v \in V\}$. A bit of non-standard notation will also be used. The neighbors of a set of vertices S are denoted by

$$N(S) := \bigcup_{v \in S} N(v) \setminus S,$$

and the degree of S by $d(S) := |N(S)|$.

2 Constructive bounds for $\alpha_2(G)$

A first step towards a general method for obtaining lower bounds on $\alpha_k(G)$ is to study 2-role assignments and $\alpha_2(G)$.

Lemma 2.1 *Let $v \in V(G)$. Suppose that either $\delta(G) \neq 1$ or that $d(v) \neq n - 1$. Then*

$$\alpha_2(G) \geq \frac{\binom{d(v)}{2} + \binom{n-1-d(v)}{2}}{\binom{n-1}{2}}. \tag{3}$$

Proof: If G is not connected, choose a connected component $C \subset V(G)$ with the minimum number of vertices and set $f(w) = 1$ for every $w \in C$ and $f(w) := 2$ for every $w \in V(G) \setminus C$. Note that f is a 2-role assignment, so $\alpha_2(G) = 1$ and (3) trivially holds.

Let G be a connected graph with $\delta(G) = 1$. Let u be such that $d(u) = 1 = \delta(G)$ (note that u is adjacent to a vertex of degree at least 2). Set $f(u) = 1$, and $f(w) = 2$ for any other $w \in V(G)$. Clearly, $M_1 = 1$ and

$$M_2 = \frac{\binom{n-2}{2}}{\binom{n-1}{2}} = \frac{n-3}{n-1}$$

since $M(x, y) = 1$ for every x, y among the $n - 2$ vertices that are not adjacent to u and $M(x, y) = 0$ whenever x or y is the vertex adjacent to u (note that the formula doesn't hold and $M_2 = 1$ in the totally uninteresting case of $n = 2$). So, we have $\alpha_2(G) \geq M(f) = (n - 3)/(n - 1)$. Note that $d(v) \neq 0$ since G is connected, and since $\delta(G) = 1$ we cannot have $d(v) = n - 1$ by assumption. The righthand side of (3), thought of as a function of $d(v)$, attains its maximum at the endpoints of the interval of allowable values, i.e., when $d(v) = n - 2$ or $d(v) = 1$. It is easy to check, for those endpoints, that the function value is $(n - 3)/(n - 1)$.

Finally assume that G is connected and that $\delta(G) \geq 2$. Set $f(v) = 1$ and $f(w) = 2$ for $w \in V(G)$, $w \neq v$. Obviously, $M_1 = 1$. Since every vertex has at least two neighbors, for every $w \neq v$,

$$f(N(w)) = \begin{cases} \{1, 2\} & v \in N(w) \\ \{2\} & v \notin N(w) \end{cases}$$

So, $M_2 = (\binom{d(v)}{2} + \binom{n-1-d(v)}{2})/\binom{n-1}{2}$. Therefore, $\alpha_2(G) \geq M(f) \geq M_2$, i.e., (3) holds. **Q.E.D.**

It is clear from (3) that the method from Lemma 2.1 produces lower bounds on $\alpha_2(G)$ that are higher for graphs with vertices of degree further from $(n - 1)/2$.

Corollary 2.2 *For all graphs G,*

$$\alpha_2(G) \geq \frac{1}{2} - \begin{cases} \frac{1}{2n-2} & n = 0 \ mod\, 2 \\ \frac{1}{2n-4} & n = 1 \ mod\, 4 \\ \frac{n-5}{2n^2-6n+4} & n = 3 \ mod\, 4 \end{cases}$$

Proof: The result is trivial for $G = K_2$. If $G \neq K_2$ there is a v satisfying Lemma 2.1. To finish the proof we find a lower bound for the righthand side of (3). Observe that the minimum over integral d's of the quadratic function $f(d) = \binom{d}{2} + \binom{n-1-d}{2}$ is attained at $d = n/2$ ($d = (n - 2)/2$) for n even, at $d = (n - 1)/2$ for n odd. This completes the proof in the case of $n = 0 \bmod 2$ and in the case $n = 1 \bmod 4$.

If $n = 3 \bmod 4$, there exists no $(n - 1)/2$-regular graph on n vertices (since both n and $(n - 1)/2$ are odd and every graph with odd $|V(G)|$ must have at least one vertex of even degree). Hence, in that case, there must exist a vertex v with $d(v) \neq (n-1)/2$. Then $f(d(v))$ is at least $f((n - 3)/2) = f((n + 1)/2)$. **Q.E.D.**

The main idea of Lemma 2.1 was to label only one vertex with 1 and label all the other vertices with 2. In this way it was ensured that $M_1 = 1$. One possible generalization of this approach would be to consider all possible surjective $f : V(G) \to \{1, 2\}$ such that $M_1 = 1$. In what follows, $G(S)$ denotes the subgraph of G generated by the set of vertices $S \subseteq V$.

Lemma 2.3 *Let $S \subset V(G)$ such that $\delta(G(S)) \geq 1$ or $\Delta(G(S)) = 0$, and such that $N(v) \cap S \neq N(v)$ for every $v \in V$. Then*

$$\alpha_2(G) \geq \frac{\binom{d(S)}{2} + \binom{n-|S|-d(S)}{2}}{\binom{n-|S|}{2}}. \tag{4}$$

Proof: Let $f(v) = 1$ for $v \in S$, and let $f(w) = 2$ for $w \in V(G) \setminus S$. Note that, for $v \in S$, $2 \in f(N(v))$ since every $v \in S$ is adjacent to at least one $w \notin S$ (this is what $N(v) \cap S \neq N(v)$

says). Also note that $\delta(G(S)) \geq 1$ implies that $\{1,2\} = f(N(v))$ for every $v \in S$; similarly $\Delta(G(S)) = 0$ implies $\{2\} = f(N(v))$ for every $v \in S$. In either case, $M_1 = 1$.

It remains to show that M_2 is equal to the righthand side of (4). Since, for any $v \in V$, $N(v) \cap S \neq N(v)$, then for any $w \notin S$, $f(N(w))$ can be either $\{1,2\}$ or $\{2\}$. Since there are exactly $d(S)$ vertices from $V(G) \setminus S$ that are adjacent to some vertex $v \in S$, we conclude that M_2 is equal to the righthand side of (4). **Q.E.D.**

As with Lemma 2.1, the righthand side of (4) is minimized when $d(S)$ is an integer closest to $\frac{n-|S|}{2}$. Thus, one obtains a bound on $\alpha_2(G)$ that is slightly less than $\frac{1}{2}$. This means that generalizing to sets S from single vertices would not bring much of an improvement in the bounds stated in Corollary 2.2. However, it is possible that, for a particular graph, Lemma 2.3 would give a better bound than Lemma 2.1.

A natural question to ask is when are the worst-case bounds from Corollary 2.2 tight? For example, it is not hard to see that $\alpha_2(C_5) = 1/3$. In order to see this, observe that it suffices to study only $f : V(C_5) \to \{1,2\}$ with $1 \leq |f^{-1}(1)| \leq 2$. If $|f^{-1}(1)| = 1$ then f is just as described in the proof of Lemma 2.1, so $M(f) = 1/3$. If $|f^{-1}(1)| = 2$ and the two vertices labeled by 1 are adjacent, then f is just as described in the proof of Lemma 2.3 and again $M(f) = 1/3$. Finally, if $|f^{-1}(1)| = 2$ and the two vertices labeled by 1 are non-adjacent, it is easy to see that the $M_1 = 1$ and $M_2 = 1/3$, so again $M(f) = 1/3$. Hence, $\alpha_2(G) = 1/3$. (In fact, we have shown that for every 2-assignment of C_5, $M(f) = 1/3$.) Note that C_5 is the only graph with $\alpha_2(G) \leq 1/3$. This follows trivially for $n \leq 4$. For $n \geq 5$, it follows directly from Corollary 2.2 and the fact that C_5 is the only 2-regular connected graph on 5 vertices.

Are there infinitely many n for which there exists a graph on n vertices with $\alpha_2(G) \leq \frac{1}{2}$ (i.e., such that the bounds from Corollary 2.2 are asymptotically tight)? The results from the next section will provide a negative answer to this question.

3 When is a graph k-role assignable?

We first show that if $\delta(G) > c_k \log n$ (where c_k is some constant that depends only on k), then there must exist $f : V(G) \to \{1, \ldots, k\}$ such that neighbors of any vertex are labeled with all k labels, i.e., such that $f(N(v)) = \{1, \ldots, k\}$ for every $v \in V(G)$. Note that such f is a k-role assignment since f is surjective and (1) holds trivially. Unlike the proofs in the previous section, such f will be shown to exist but will not be constructed.

Theorem 3.1 *For any integer $k > 1$ and any graph G,*

$$\delta(G) > \frac{\log(kn)}{\log(\frac{k}{k-1})} \tag{5}$$

implies that $\alpha_k(G) = 1$.

Proof: The proof is based on a standard probabilistic method argument.[1]

Define $f : V(G) \to \{1, \ldots, k\}$ as follows: for every $v \in V(G)$ choose $f(v)$ uniformly at random from $\{1, \ldots, k\}$.

Now, for any $v \in V$,

$$P(f(N(v)) \neq \{1, 2, \ldots, k\}) \leq k \left(\frac{k-1}{k}\right)^{d(v)} \tag{6}$$

[1]For more on the probabilistic method, see [1].

because all $f(w)$ are chosen independently and for a fixed value i, the probability that i is not used for any of the vertices adjacent to v is $(\frac{k-1}{k})^{|N(v)|} = (\frac{k-1}{k})^{d(v)}$, and there are k choices for i.

Define a random variable X_v such that $X_v = 1$ if $f(N(v)) \neq \{1, 2, \ldots, k\}$ and $X_v = 0$ if $f(N(v)) = \{1, 2, \ldots, k\}$. Let

$$Z := \sum_{v \in V} X_v.$$

Note that Z is integer valued and that $Z = 0$ implies that $f(N(v)) = \{1, 2, \ldots, k\}$ for every $v \in V$, i.e., in this case f is a k-role assignment.

Therefore, whenever $\mathtt{E}[Z] < 1$, there must exist at least one f such that $Z = 0$ (otherwise $\mathtt{E}[Z]$ would be at least one). Note that

$$\mathtt{E}[Z] = \sum_{v \in V} \mathtt{E}[X_v] = \sum_{v \in V} P(f(N(v)) \neq \{1, 2, \ldots, k\}) \leq \sum_{v \in V} k \left(\frac{k-1}{k}\right)^{d(v)}.$$

So, G is k-role assignable whenever

$$\sum_{v \in V} k \left(\frac{k-1}{k}\right)^{d(v)} < 1. \tag{7}$$

Since the largest summand in the sum above is when $d(v) = \delta(G)$,

$$\sum_{v \in V} k \left(\frac{k-1}{k}\right)^{d(v)} \leq nk \left(\frac{k-1}{k}\right)^{\delta(G)}$$

and G is k-role assignable whenever

$$nk \left(\frac{k-1}{k}\right)^{\delta(G)} < 1. \tag{8}$$

The last inequality is equivalent to (5). **Q.E.D.**

Corollary 3.2 *For any integer $k \geq 1$ there exists a constant c_k such that for any graph G*

$$\delta(G) > c_k \log n \Rightarrow \alpha_k(G) = 1.$$

Proof: Note that $k \leq n$ and, hence, (5) holds whenever

$$\delta(G) > \frac{\log(n^2)}{\log(\frac{k}{k-1})}.$$

Take $c_k = 2/\log(\frac{k}{k-1})$. **Q.E.D.**

Remark 3.3 Note that a k-role assignment f was only shown to exist whenever (7) holds. However, such f can be found in polynomial time by applying a basic derandomization technique: the method of conditional probabilities. Simply define $f(v_1) = 1$ and then, with $f(v_1), \ldots, f(v_{i-1})$ already defined, choose any value $f(v_i)$ such that

$$\mathtt{E}[Z|f(v_1), \ldots, f(v_{i-1}), f(v_i)] \leq E[Z].$$

To see that there is such a value note that $\mathtt{E}[Z|f(v_1)] = \mathtt{E}[Z]$ and that

$$\mathtt{E}[Z|f(v_1), \ldots, f(v_{i-1})] = \sum_{j=1}^{k} \mathtt{E}[Z|f(v_1), \ldots, f(v_{i-1}), f(v_i) = j] P(f(v_i) = j),$$

so there exists at least one j such that

$$\mathtt{E}[Z|f(v_1), \ldots, f(v_{i-1}), f(v_i) = j] \leq \mathtt{E}[Z|f(v_1), \ldots, f(v_{i-1})].$$

Therefore, by induction, it is possible to define f for which $Z \leq \mathtt{E}[Z]$ as long as it is possible to compute $\mathtt{E}[Z|f(v_1), \ldots, f(v_{i-1}), f(v_i)]$ for all possible choices of $f(v_i)$. But this can be done in a way analogous to the computation of the lefthand side of (6) and this can obviously be done in polynomial time. We will have to do at most k such calculations when determining $f(v_i)$ and there are n values of f that need to be determined. Hence, f can be found in at most kn rounds of computing a value similar to the lefthand side of (6).

Remark 3.4 As shown in the proof of Theorem 3.1, G has a k-role assignment f whenever (7) holds. However, there is no information about sizes of sets of vertices getting the same label, i.e., on $|f^{-1}(i)|$ for $i = 1, \ldots, k$. Still, we may assume that $\lfloor \frac{n}{k} \rfloor \leq |f^{-1}(i)| \leq \lceil \frac{n}{k} \rceil$ for $i = 1, \ldots, k$ (that is, f defines an equipartition of $V(G)$).

To formally prove this, let us denote by $\mathtt{R}$ the model under which f is defined by choosing $f(v)$ independently and uniformly at random from $\{1, \ldots, k\}$ and by $\mathtt{S}$ the model in which we define f by choosing a random k-equipartition of V, that is $V_1 \cup V_2 \cup \ldots \cup V_k$ with $\lfloor \frac{n}{k} \rfloor \leq |V_i| \leq \lceil \frac{n}{k} \rceil$, and then define $f(v) = i$ if and only if $v \in V_i$. Let $P_{\mathtt{R}}(X)$ denote the probability of event X under model $\mathtt{R}$ and similarly for model $\mathtt{S}$. Let

$$F_{(s_1, s_2, \ldots, s_k)} = \{f : |f^{-1}(i)| = s_i, i = 1, \ldots, k\}.$$

For a given $v \in V(G)$, consider the event

$$A = \{f : f(N(v)) \neq \{1, \ldots, k\}\}.$$

We will show that $P_{\mathtt{S}}(A) \leq P_{\mathtt{R}}(A)$, so that (6) holds in model $\mathtt{S}$ also.

First observe that

$$P_{\mathtt{R}}(A) = \sum P_{\mathtt{R}}(A|F_{(s_1, \ldots, s_k)}) P_{\mathtt{R}}(F_{(s_1, \ldots, s_k)}) \tag{9}$$

where the summation goes over all possible nonnegative integers $s_1, \ldots, s_k$ such that $s_1 + \ldots + s_k = n$. Note that Equation (9) also holds with $P_{\mathtt{S}}$ replacing $P_{\mathtt{R}}$ throughout. It is easy to see that

$$P_{\mathtt{R}}(A|F_{(s_1, \ldots, s_k)}) = \sum_{i=1}^{k} \frac{\binom{n-s_i}{d(v)}}{\binom{n}{d(v)}}$$

is minimized for equipartitions, i.e., when $max_{i,j}|s_i - s_j| \leq 1$. It is also straightforward that for all equipartitions $(s_1, \ldots, s_k)$ and $(t_1, \ldots, t_k)$,

$$P_{\mathtt{R}}(A|F_{(s_1, \ldots, s_k)}) = P_{\mathtt{S}}(A|F_{(s_1, \ldots, s_k)}).$$

Moreover,

$$P_{\mathtt{R}}(A|F_{(s_1,\dots,s_k)}) = P_{\mathtt{R}}(A|F_{(t_1,\dots,t_k)})$$

and

$$P_{\mathtt{R}}(F_{(s_1,\dots,s_k)}) = P_{\mathtt{R}}(F_{(t_1,\dots,t_k)})$$

and similarly for $P_{\mathtt{S}}$. Note also that $P_{\mathtt{S}}(F_{(s_1,\dots,s_k)}) = 0$ if $(s_1, \dots, s_k)$ is not an equipartition. Thus, if we fix one particular equipartition $(u_1, \dots, u_k)$, (9) implies that

$$P_{\mathtt{R}}(A) \geq P_{\mathtt{R}}(A|F_{(u_1,\dots,u_k)}) \sum P_{\mathtt{R}}(F_{(s_1,\dots,s_k)}) = P_{\mathtt{R}}(A|F_{(u_1,\dots,u_k)})$$

and, similarly, (9) with $P_{\mathtt{S}}$ implies that

$$P_{\mathtt{S}}(A) = P_{\mathtt{S}}(A|F_{(u_1,\dots,u_k)}).$$

It follows that $P_{\mathtt{S}}(A) \leq P_{\mathtt{R}}(A)$. Thus, (6) holds in model $\mathtt{S}$ also, which means that the rest of the proof of Theorem 3.1 also holds.

Remark 3.5 A graph of at least three vertices is called **role primitive** if the only role assignments are the **trivial assignments** that give each vertex the same role or each vertex a different role. Everett and Borgatti [6] showed that there exist role primitive graphs. In [8], Everett and Borgatti asked how common are role primitive graphs. In fact, on several occasions while giving lectures on the topic at various meetings, Everett has conjectured that, asymptotically, almost all graphs are role primitive [5]. The proof of Theorem 3.1 shows that large classes of graphs have a role assignment in which the neighborhood of each vertex has all roles in $\{1, \dots, k\}$. Thus, asymptotically, almost all graphs are not role primitive. [2]

As will be shown in the next section, Theorem 3.1 shows that $\alpha_k(G) = 1$ whenever G doesn't have vertices of a very small degree. The next result, whose proof is similar in flavor and again based on a more sophisticated probabilistic argument, the Lovász Local Lemma, shows that $\alpha_k(G) = 1$ whenever the graph is close to being regular. In this context, close to regular will mean that $\delta(G) \geq c \log \Delta(G)$.

Before stating and proving the theorem, let us recall a uniform version of the Lovász Local Lemma.

Theorem 3.6 (Lovász Local Lemma[1]) *Let A_i, $i \in I$, be events in a discrete probability space. If there exists M such that for every $i \in I$*

$$|\{A_j : A_j \textit{ is not independent of } A_i\}| \leq M,$$

and if there exists $p > 0$ such that $P(A_i) \leq p$ for every $i \in I$, then

$$ep(M+1) \leq 1 \Rightarrow P(\bigcap_{i \in I} A_i^c) > 0.$$

Theorem 3.7 *For every integer $k > 1$ there exists a constant c_k such that*

$$\delta(G) > c_k \log(\Delta(G)) \Rightarrow \alpha_k(G) = 1.$$

[2] This assumes some standard measure on the set of all graphs, e.g., any measure arising from standard random graph models (cf. [1])

Proof: As in the proof of Theorem 3.1, let f be defined by choosing all $f(v)$ independently and uniformly at random from $\{1, \ldots, k\}$. For $v \in V(G)$, let A_v be the event that $f(N(v)) \neq \{1, \ldots, k\}$. As shown in the proof of Theorem 3.1,

$$P(A_v) \leq k \left(\frac{k-1}{k}\right)^{d(v)} \leq k \left(\frac{k-1}{k}\right)^{\delta(G)}.$$

Also note that A_v and A_w are not independent if and only if $N(v) \cap N(w) \neq \emptyset$. Hence, A_v with $M = \Delta(G)^2$ and $p = k(\frac{k-1}{k})^{\delta(G)}$ satisfies the conditions of the Lovász local lemma. Therefore,

$$ek \left(\frac{k-1}{k}\right)^{\delta(G)} (\Delta(G)^2 + 1) \leq 1 \Rightarrow P(\bigcap_{v \in V(G)} A_v^c) > 0. \tag{10}$$

If the righthand side of (10) holds, there exists at least one f such that $f(N(v)) = \{1, \ldots k\}$ for every $v \in V(G)$, that is, there exists at least one k-role assignment. In order to finish the proof it remains to show that the lefthand side of (10) is equivalent to

$$\delta(G) \geq \frac{\log(ek(\Delta(G)^2 + 1))}{\log(\frac{k}{k-1})}$$

which is a matter of simple algebraic manipulation. It is straightforward to see that there exists c_k such that $c_k \log(\Delta(G))$ is greater than the righthand side of the above inequality. **Q.E.D.**

Even though the proof of Theorem 3.7 is nonconstructive, there exist methods to derandomize the proof and thereby construct a k-role assignment. The derandomization method of the Lovasz Local Lemma is described in [3].

Note that the use of the local lemma doesn't improve the general bound from Theorem 3.1 since $\Delta(G)$ can be as large as $n-1$. Still, as already mentioned, the bound on $\delta(G)$ in terms of $\Delta(G)$ could be useful in the case of "almost" regular graphs; for example, if $\delta(G) \leq \log(n)$, Theorem 3.1 doesn't guarantee existence of a k-role assignment while Theorem 3.7 does, provided that $\delta(G) \geq c_k \log(\Delta(G))$.

It should also be noted that the use of the probabilistic method in the study of k-role assignability is not limited to the results presented in this section. For example, instead of the uniform version, we could have used the non-uniform version of the Lovasz Local Lemma[1] to obtain a result that would be slightly more general and more technical[3] than that of Theorem 3.7. Similarly, the application of more involved techniques of the probabilistic method would likely widen the class of graphs that are provably k-role assignable. However, our goal here was not to dig deeply into technicalities, but to use the probabilistic method as a tool in order to show that almost all social networks can be captured by the role assignment model (cf. Remark 3.5 and its footnote).

[3]The condition for k-role assignability would be more general and would involve degrees of all the vertices in the graph, and not only $\delta(G)$ and $\Delta(G)$.

4 Better bounds for $\alpha_2(G)$

Combining the results from Theorem 3.1 and Lemma 2.1 leads to the following theorem.

Theorem 4.1 *Let G be a graph on at least three vertices. Then*

$$\alpha_2(G) > 1 - \frac{2\lceil \log_2 n \rceil}{n-2} \tag{11}$$

Proof: First note that, for $n \leq 10$, the righthand side of (11) is nonpositive, hence the theorem holds trivially. So, we may assume that $n > 10$.

Setting $k = 2$ in (5) yields that $\alpha_2(G) = 1$ whenever $\delta(G) > 1 + \log_2 n$. If $\delta(G) = 1 + \log_2 n$ then in the proof of Theorem 3.1, $\mathtt{E}[Z] \leq 1$. We can still conclude that there exists f with $Z = 0$ (otherwise $Z = 1$ for every f which is not true since, for example, $f \equiv 1$ implies $Z = n$). Hence, $\alpha_2(G) = 1$.

If $\delta(G) < 1 + \log_2 n$, that is, if $\delta(G) \leq \lceil \log_2 n \rceil$, we apply Lemma 2.1 with v such that $d(v) = \delta(G)$, so

$$\alpha_2(G) \geq \frac{\binom{\delta(G)}{2} + \binom{n-1-\delta(G)}{2}}{\binom{n-1}{2}}.$$

Since the function $f(t) = \binom{t}{2} + \binom{n-1-t}{2}$ is a quadratic function attaining minimum at $t = (n-1)/2$ and since $\delta(G) \leq \lceil \log_2 n \rceil \leq (n-1)/2$ for $n > 10$, it follows that $f(\delta(G)) \geq f(\lceil \log_2 n \rceil)$. Therefore,

$$\alpha_2(G) \geq \frac{\binom{\lceil \log_2 n \rceil}{2} + \binom{n-1-\lceil \log_2 n \rceil}{2}}{\binom{n-1}{2}} \tag{12}$$

The righthand side can be bounded below as follows:

$$\begin{aligned}
\frac{\binom{\lceil \log_2 n \rceil}{2} + \binom{n-1-\lceil \log_2 n \rceil}{2}}{\binom{n-1}{2}} &= \frac{\lceil \log_2 n \rceil(\lceil \log_2 n \rceil - 1) + ((n-1) - \lceil \log_2 n \rceil)((n-2) - \lceil \log_2 n \rceil)}{(n-1)(n-2)} \\
&= 1 - \frac{\lceil \log_2 n \rceil}{n-1} - \frac{\lceil \log_2 n \rceil}{n-2} + \frac{2\lceil \log_2 n \rceil^2 - \lceil \log_2 n \rceil}{(n-1)(n-2)} \\
&\geq 1 - \frac{\lceil \log_2 n \rceil}{n-1} - \frac{\lceil \log_2 n \rceil}{n-2} \\
&\geq 1 - \frac{2\lceil \log_2 n \rceil}{n-2}
\end{aligned}$$

and this concludes the proof. **Q.E.D.**

Theorem 4.1 shows that for any $\epsilon > 0$, all but finitely many graphs have $\alpha_2(G) > 1 - \epsilon$, that is, all but finitely many graphs are almost 2-role assignable (according to α_2 as a measure of closeness), no matter how picky one wants to be about "almost". Of course, this is because for every $\epsilon > 0$, there exists n_ϵ such that, for every $n > n_\epsilon$, $2\lceil \log_2 n \rceil/(n-2) < \epsilon$. So for every G on more than n_ϵ vertices, $\alpha_2(G) > 1 - \epsilon$ (by Theorem 4.1), and there are only finitely many graphs on at most n_ϵ vertices.

Another important consequence of Theorem 4.1 is that the problem of determining $\alpha_2(G)$, has a PTAS (polynomial time approximation scheme). Recall that an NP-hard optimization problem (e.g., the problem of determining $\alpha_2(G)$) has a polynomial time approximation scheme if for every $\epsilon > 0$, there exists a polynomial time algorithm that finds a feasible value that is within ϵ of the optimum. Hence, there is a PTAS for the problem of finding $\alpha_2(G)$ because, for a given $\epsilon > 0$, we can find the exact value of $\alpha_2(G)$ for finitely many graphs G on $1, 2, \ldots$ or n_ϵ vertices (by simply checking all possible 2-assignments) and for all that we only need constant time (the constant is enormous, though). If $|V(G)| > n_\epsilon$, then we can find $\delta(G)$ in time linear in $|E(G)|$ and knowing $\delta(G)$, using Theorem 4.1, we either conclude $\alpha_2(G) = 1$ or use Lemma 2.1 to construct f (using v with $d(v) = \delta(G)$) that is within $2\lceil \log_2 n \rceil/(n-2)$ of $\alpha_2(G)$. Note that f from Lemma 2.1 can be found in polynomial time in a trivial manner, while the f from Theorem 3.1 can be found in not so trivial manner but still in polynomial time (Remark 3.3). Therefore, not only can we find in polynomial time an approximation of $\alpha_2(G)$ that is within ϵ, but we can also construct in polynomial time a k-assignment f such that the approximation in question is $M(f)$.

Even though $\alpha_2(G) > 1 - \epsilon$ for all but finitely many graphs, it might be worth finding $\alpha_2(G)$ for small G. We have already shown in Section 2 that $\alpha_2(C_5) = 1/3$ and C_5 is the unique graph with minimum possible $\alpha_2(G)$. For example, Theorem 4.1 (with bound from (12) rather than the one from the statement of the theorem) also gives that $\alpha_2(G) > \frac{1}{2}$ for all G on 13 or more vertices. This is readily checked for $n = 13, 14, 15$. The rest follows by observing that, for $p = 4, 5, \ldots$, the righthand side of (12) is greater than 1/2 whenever $n = 2^p$, and is an increasing function in n within the interval $(2^p, 2^{p+1}]$ (note that $\lceil \log_2 n \rceil$ is a constant there).

5 Almost all graphs are nearly k-role assignable

In order to prove an analogue of Theorem 4.1 in the case of k-role assignments, an analogue of Lemma 2.1 is needed.

Lemma 5.1 *Let $S \subset V(G)$, $|S| = k - 1$, such that S contains all isolated vertices of G (if any) or such that all vertices of S are isolated. Then*

$$\alpha_k(G) \geq \frac{\binom{n+1-k-d(S)}{2}}{\binom{n+1-k}{2}}.$$

Proof: If all vertices are isolated, the lemma holds trivially since $\alpha_k(G) = 1$.

Let $S = \{v_1, \ldots, v_{k-1}\}$. Define $f(v_i) = i$ for every $v_i \in S$ and $f(w) = k$ for every $w \in V(G) \setminus S$. Note that $M_1 = \ldots = M_{k-1} = 1$. Also note that $M_k \geq \frac{\binom{n-d(S)-|S|}{2}}{\binom{n-|S|}{2}}$ because there are $n - d(S) - |S|$ vertices which have no neighbors in S. So for any such vertex w, $f(N(w)) = \{k\}$. (If some $w \in V(G) \setminus S$ is isolated, then all vertices in S must be isolated. Then we redefine $f(w) = k - 1$ for every isolated $w \in V(G) \setminus S$ and this makes f a k-role assignment because not all vertices are isolated.) **Q.E.D.**

In order to show that, for every $\epsilon > 0$, $\alpha_k(G) > 1 - \epsilon$ for all but finitely many graphs, Lemma 5.1 and Theorem 3.1 are combined as follows. First calculate $\delta(G)$. If (5) holds, then $\alpha_k(G) = 1$. Otherwise set $f(v_1) = 1$ where $d(v_1) = \delta(G)$ and consider $k - 1$ assignments on $G - v_1$. If (5) holds for $\delta(G - v_1)$ (and with $k - 1$ instead of k), then there exists a $k-1$-role assignment of $G - v_1$ where the number of vertices of each label is (almost) the same. Then, by Remark 3.4, even if all neighbors of v_1 get the same label, $\alpha_k(G)$ is still going to be large (where we consider a k-assignment f obtained by merging $f(v_1) = 1$ and the $(k-1)$-role assignment of $G - v_1$) If, however, $\delta(G - v_1)$ violates (5), then we set $f(v_2) = 2$ where v_2 is the vertex of a small degree in $G - v$ and repeat the procedure by considering $(k-2)$-assignments of $G - v_1 - v_2, \ldots$ If this procedure has to be repeated until $G - v_1 - \ldots - v_{k-1}$ (which means that $d(v_i)$ are all small), then Lemma 5.1 with $S = \{v_1, \ldots, k-1\}$ would give a satisfactory bound.

Theorem 5.2 *For any positive integer k there exists a positive constant C_k such that for every G on at least k vertices,*

$$\alpha_k(G) \geq 1 - \frac{C_k \log n}{n}.$$

Proof: Let $S = \{v_1, v_2, \ldots, v_{k-1}\}$ be a set of vertices of G with the $k-1$ smallest degrees, i.e., $\delta(G) = d(v_1) \leq d(v_2) \leq \ldots \leq d(v_{k-1}) \leq d(w)$ for any $w \notin S$.

First suppose that

$$\sum_{i=1}^{k-1} (d(v_i) - i + 1) \leq \sum_{i=1}^{k-1} \frac{\log((k+1-i)n)}{\log(\frac{k+1-i}{k-i})}. \tag{13}$$

In this case, by (13),

$$d(S) \leq \sum_{i=1}^{k-1} d(v_i) \leq \frac{(k-2)(k-1)}{2} + (k-1)\frac{\log(kn)}{\log(\frac{k}{k-1})}$$

In other words, there exists a constant D_k such that $d(S) \leq D_k \log n$. Now, by Lemma 5.1,

$$\alpha_k(G) \geq \frac{\binom{n+1-k-d(S)}{2}}{\binom{n+1-k}{2}} \geq \frac{\binom{n+1-k-\lfloor D_k \log n \rfloor}{2}}{\binom{n+1-k}{2}}$$

whenever $n \geq k - 1 + 2\lfloor D_k \log n \rfloor$ since $f(t) = \binom{n+1-k-t}{2}$ is a quadratic function in t with minimum at $(n+1-k)/2 \geq \lfloor D_k \log n \rfloor \geq d(S)$. So, for $n \geq k - 1 + 2\lfloor D_k \log n \rfloor$,

$$\alpha_k(G) \geq 1 - \frac{\lfloor D_k \log n \rfloor}{(n-k+1)} - \frac{\lfloor D_k \log n \rfloor}{(n-k)}.$$

Therefore, there exists a constant A_k such that

$$\alpha_k(G) \geq 1 - \frac{A_k \log n}{n}$$

for every G satisfying (13). (We can always choose A_k large enough to get $1 - \frac{A_k \log n}{n} \leq 0$ for every $n < k - 1 + 2\lfloor D_k \log n \rfloor$.)

Next suppose that (13) doesn't hold, that is,

$$\sum_{i=1}^{k-1} (d(v_i) - i + 1) > \sum_{i=1}^{k-1} \frac{\log((k+1-i)n)}{\log(\frac{k+1-i}{k-i})}.$$

Let j be the smallest number among $1, 2, \ldots, k-1$ for which

$$d(v_j) - j + 1 > \frac{\log((k+1-j)n)}{\log(\frac{k+1-j}{k-j})}$$

Such j must exist; otherwise (13) would hold. If $j = 1$, then $\delta(G)$ satisfies (5) and $\alpha_k(G) = 1$ by Theorem 3.1. So, we may assume $j \geq 2$.

Let $G_j := G - v_1 - v_2 - \ldots - v_{j-1}$. Note that $\delta(G_j) \geq d(v_j) - j + 1$ since $d_{G_j}(w) = d(w) - |N(w) \cap \{v_1, v_2 \ldots v_{j-1}\}| \geq d(w) - j + 1$ and since $v_1, v_2, \ldots v_{j-1} \notin V(G_j)$. Hence, (5) holds for G_j with $k+1-j$ instead of k, so G_j has a $(k+1-j)$-role assignment such that $f(N(v)) = \{j, j+1, \ldots, k\}$ for all $v \in G_j$. Moreover, by Remark 3.4, f can be chosen so that $\lfloor \frac{n-j+1}{k+1-j} \rfloor \leq |f^{-1}(i)| \leq \lceil \frac{n-j+1}{k+1-j} \rceil$, for every $i = j, j+1, \ldots, k$. This f can be extended to a k-assignment of G by setting $f(v_i) = i$, for $i = 1, \ldots, j-1$. Note that $M_1 = M_2 = \ldots = M_{j-1} = 1$. It remains to calculate M_i for $j \leq i \leq k$. For every $v \in V(G_j)$, $f(N(v)) = \{j, j+1, \ldots k\}$ if and only if $\{v_1, \ldots, v_{j-1}\} \cap N(v) = \emptyset$. Note that $d(\{v_1, \ldots, v_{j-1}\}) \leq d(v_1) + \ldots + d(v_{j-1})$, so there can be at most $d(v_1) + \ldots + d(v_{j-1})$ vertices v such that $f(v) = i$ and such that $\{v_1, \ldots, v_{j-1}\} \cap N(v) \neq \emptyset$. By the choice of j,

$$\sum_{t=1}^{j-1} d(v_t) \leq \sum_{t=1}^{j-1} \frac{\log((k+1-t)n)}{\log(\frac{k+1-t}{k-t})} + \frac{(j-1)(j-2)}{2}$$

$$\leq \sum_{t=1}^{k-1} \frac{\log((k+1-t)n)}{\log(\frac{k+1-t}{k-t})} + \frac{(k-1)(k-2)}{2} \leq D_k \log n.$$

(The last inequality was shown earlier in the proof.) Hence,

$$M_i \geq \max\{\frac{\binom{\lceil \frac{n-j+1}{k+1-j} \rceil - \lfloor D_k \log n \rfloor}{2}}{\binom{\lceil \frac{n-j+1}{k+1-j} \rceil}{2}}, \frac{\binom{\lfloor \frac{n-j+1}{k+1-j} \rfloor - \lfloor D_k \log n \rfloor}{2}}{\binom{\lfloor \frac{n-j+1}{k+1-j} \rfloor}{2}}\}$$

So, a simple algebraic manipulation similar to the one preceding the definition of A_k shows that there exists a constant B_k such that

$$\alpha_k(G) \geq 1 - \frac{B_k \log n}{\lfloor \frac{n-j+1}{k+1-j} \rfloor}.$$

Hence, there exists a constant $C_k \geq A_k$ such that

$$\alpha_k(G) \geq 1 - \frac{C_k \log n}{n}.$$

Q.E.D.

Theorem 5.2 shows that all but finitely many graphs are "almost" k-role assignable and this is true no matter how picky we are about "almost".

Note that the same argument that was used to show that there exists a PTAS for the problem of finding $\alpha_2(G)$ also works (now with Theorem 5.2) for $\alpha_k(G)$. In other words, there is a PTAS for the problem of finding $\alpha_k(G)$.

6 Concluding Remarks

We have shown that the role assignment model fits any social network in which the difference in the number of relationships is not too large for any two individuals in the network (Theorem 3.7). For example, social roles in "democratic" social networks, the networks in which all individuals are involved in roughly the same number of relationships, can be captured by the k-role assignment model. A slightly weaker result (Theorem 3.1) shows that social roles in all social networks in which each individual is involved in a significant number of relationships can also be captured by the role assignment model.

As already mentioned in Section 1, the role assignment model has been studied extensively and it is known that it is hard to determine the existence of and to construct a k-role assignment describing social roles in a given network. In view of these difficulties, one could be led to believe that for most social networks only trivial role assignments (cf. Remark 3.5) are possible, i.e., that the role assignment model cannot effectively reduce the size of the data set (homomorphically "reduced" social network) needed to analyze social roles in a network. We have shown that this belief is dramatically incorrect and that, in fact, large social networks are very likely to have a non-trivial role assignment. Hence, most of the difficulties associated with determining a possible role assignment are encountered when dealing with small networks. One might argue that finding a non-trivial role assignment in a small social network is not so crucial given that the data set is manageable in this case.

Rarely does a mathematical model fit the given data perfectly and one is often satisfied when the number of inconsistencies or errors in the model output is negligible. We have shown (Theorem 5.2) that the role assignment model has exactly this property. Our result is based on the measure of closeness of fit of the model that is motivated by the blockmodeling literature and introduced in Section 1. A natural question to ask is whether the results of this paper would hold with a different measure of the closeness of fit of the model. There are several other possibilities for defining $\alpha_k(G)$. For example, one might argue that M_i (cf. (2)) does not necessarily give the most natural count of the proportion of pairs of individuals that are assigned role i by the model and that satisfy (1). [4] Alternatively, one could redefine M_i to be $\binom{s}{2}/\binom{|V_i|}{2}$ where s is the largest number of vertices of V_i such that they have the same set of labels in their neighborhoods. Another way of redefining M_i would be to set it to $s/|V_i|$. Further definitions could be obtained by changing $M(f)$ to be, say, the arithmetic mean of the M_i's, that is $M(f) := (M_1 + M_2 + \dots M_k)/k$ or any other average (min is only one possible choice for averages). Yet another way of defining a measure of closeness different than α_k would be not to normalize M_i's for all i's separately but to normalize after aggregating. The reason for doing this would be to penalize for model errors regardless of the size of the sets $|V_i|$. It is important to note that Theorem 5.2 (and its proof) would still hold if α_k is replaced by any of these proposed alternative measures of closeness of fit. Thus, our result does not depend on the particular choice of the measure of closeness.

It should also be noted that results presented in this paper readily generalize in the digraph case of the role assignment model, i.e., when the relationship among two individuals in a social network is not necessarily bidirectional (e.g., individual i can be affected by

[4] For example, suppose that $|V_i| = 1000$ with 990 of them with the set A being the set of the labels of their neighbors, and the remaining 10 with the set B being the set of the labels of their neighbors, $A \neq B$. Then $M_i = (\binom{990}{2} + \binom{10}{2})/\binom{1000}{2}$. One might argue that, given that the vast majority of individuals labeled i have A as the set of the labels of their neighbors, the neighborhood of all individuals labeled i should be A, and thus the vertices labeled i with B as the set of labels of their neighbors should be considered labeled in error. However, in computing the score for M_i, these 10 "error" individuals also contributed to M_i.

individual j's action, role, state,... and perceive this as a relationship with individual j, while at the same time individual j might not be affected at all by individual i and thus perceive no relationship with i).

Even though some of our proofs (Theorem 3.1 and Theorem 3.7) are nonconstructive, we have noted that there are ways to derandomize the argument there and thus construct a role-assignment for a particular social network. Moreover, we have addressed computational complexity issues related to finding such role assignments. In our view, because of the larger and larger sizes of data sets that are being analyzed, use of mathematical models in the social sciences requires at least awareness of the computational complexity of the model that is being used.

The peculiar thing about the methods and results in this paper is that we always find a k-role assignment with a special property that (almost) all vertices are adjacent to all k labels. Of course, it might be the case that for a given social network there exist other k-role assignments. However, our results do not guarantee the existence of such role assignments. It is conceivable that in some applications a role assignment of a social network such that all neighborhood sets have all k labels is not very informative. Thus, in such situations, one would be interested in finding other role assignments or possibly modifying the role assignment model in order to exclude these role assignments that played a crucial role in our analysis. If we don't allow these role assignments[5] then our earlier conclusion that large social networks are very likely to have a non-trivial role assignment is of course not necessarily correct.

We close the paper by proposing a modification of the role assignment model that we feel might be interesting to those applying the model, as well as to those interested in graph theoretical issues arising from models of social networks. The proposed modification of the role assignment model would limit the number of different roles in the neighborhood of an individual. This limit could be different for different individuals (or associated to roles assigned to them) or could be uniform for the network. In the latter case, we could define a **restricted s-size k-role assignment** of a graph G to be a k-role assignment in which $|f(N(v))| \leq s$ must hold for all $v \in V(G)$. In the case when $s = 1$, i.e., when everyone in the neighborhood of an individual should be assigned the same label, it is straightforward to see that G has a restricted 1-size k-role assignment if and only if $k \leq c(G) + b(G)$, where $c(G)$ is the number of connected components of G and $b(G)$ is the number of connected components of G that are bipartite but not isolated vertices. (In particular, a connected graph G on at least two vertices has a restricted 1-size 2-role assignment if and only if G is bipartite.) The cases with $s \geq 2$ do not seem to be as straightforward and could be an interesting topic for further research.

References

[1] N. Alon and J. Spencer *The Probabilistic Method*, Wiley, NY, 1992.

[2] P. Arabie, S. Boorman, and P. R. Levitt, Constructing blockmodels: How and why, *J. Math. Psychol.*, **17** (1978), 21–63.

[5] For example, that would be the case if the analysis were restricted to k-role assignments for which the opposite direction of the implication in (1) must also hold for every pair of vertices $v, w \in V(G)$: $[f(N(v)) = f(N(w))] \Rightarrow [f(v) = f(w)]$. Everett and Borgatti [8] call such assignments *perfect regular colorations*, while White and Reitz [15] say that such assignments define *structural homomorphisms*.

[3] J. Beck, An algorithmic approach to the Lovasz Local Lemma, *Random Structures & Algorithms*, **2** (1991), 343–365.

[4] R. L. Breiger, S. A. Boorman, and P. Arabie, An algorithm for clustering relational data with applications to social network analysis and comparison to multidimensional scaling, *J. Math. Psychol.*, **12** (1975), 328–383.

[5] M. G. Everett, *personal communication*, June 2, 1999.

[6] M. G. Everett and S. P. Borgatti, Role coloring a graph, *Math. Soc. Sci.* **21** (1991), 183–188.

[7] M. G. Everett and S. P. Borgatti, Regular equivalence: General theory, *J. of Math. Soc.*, **19** (1994), 29–52.

[8] M. G. Everett and S. P. Borgatti, The regular coloration of graphs, in C. Mitchell (ed.), *Applications of Combinatorial Mathematics*, The Institute of Mathematics and its Applications Conference Series No. 60, Clarendon Press, Oxford, 1997, 49–58.

[9] F. Lorrain and H. C. White, Structural equivalence of individuals in social networks, *J. Math. Sociology*, **1** (1971), 49–80.

[10] F. S. Roberts, Graph theory and the social sciences, in R. J. Wilson and L. W. Beineke (eds.), *Applications of Graph Theory*, Academic Press, New York, 1979, 255–291.

[11] F. S. Roberts, *Applied Combinatorics*, Prentice-Hall, Englewood Cliffs, NJ, 1984.

[12] F. S. Roberts and L. Sheng, Role assignments, in Y. Alavi, D.R. Lick, and A. Schwenk (eds.), *Combinatorics, Graph Theory, and Algorithms*, Vol. II, New Issues Press, Kalamazoo, MI, 1999, 729–745.

[13] F. S. Roberts and L. Sheng, How hard is it to determine if a graph has a 2-role assignment?, preprint, Department of Mathematics, Rutgers University, 1997.

[14] L. D. Sailer, Structural equivalence: meaning and definition, computation and application, *Social Networks*, **1** (1978), 73–90.

[15] D. R. White, & K. P. Reitz,, Graph and semigroup homomorphisms on networks of relations, *Social Networks*, **5** (1983), 193-234.