ON THE THEORY OF LOCALLY AND GLOBALLY CONNECTED DESIGNS

BU-410-M

by

John Eccleston

Cornell University

Abstract

This study is concerned with the characterization and investigation of connected block designs. Every elementary treatment contrast is estimable if and only if the design is at least locally connected. Thus connectedness is an important and desirable property which every block design should enjoy. In particular, globally connected designs should yield better estimators (with respect to some optimality criterion, see conclusion) of all the elementary contrasts.

The definition of locally connected designs is the same as the connected designs of Chakrabarti [3] and Bose [1]. Several theorems which characterize locally connected designs, in terms of the incidence matrix N or some function of it, are given in section 3. A set of necessary and sufficient conditions for a design to be globally connected is given in section 4, and a new class of connected designs, pseudo-globally connected, is introduced and characterized in section 5.

Some invariance properties of both locally and globally connected designs are presented in section 6. In addition we have considered the proposition of combining connected designs so that the newly composed design has some connected nature.

There is a strong analogy between some graph theory concepts and experimental design theory. Several theorems and concepts from Harary [4] and Busacker and Saaty [2] yield graph theoretic analogies to several of the theorems in section 3 and further properties and notions applicable to experimental (block) design.

Paper No. BU-410-M in the Biometrics Unit Mimeograph Series, Department of Plant Breeding and Biometry, Cornell University, Ithaca, New York 14850.

ON THE THEORY OF LOCALLY AND GLOBALLY CONNECTED DESIGNS

BU-410-M by February, 1972

John Eccleston
Cornell University

1. Introduction

This study is concerned with the characterization and investigation of connected block designs. Every elementary treatment contrast is estimable if and only if the design is at least locally connected. Thus connectedness is an important and desirable property which every block design should enjoy. In particular, globally connected designs should yield better estimators (with respect to some optimality criterion, see conclusion) of all the elementary contrasts.

Our definition of locally connected designs is the same as the connected designs of Chakrabarti [3] and Bose [1]. Several theorems which characterize locally connected designs are given in section 3. All of these theorems involve the incidence matrix N or some function of it, as opposed to the C matrix used by Chakrabarti [3] and the treatment-block chains of Bose [1] who originated the concept of connectedness. The association matrix, NN', and the block characteristic matrix, N'N, are the functions of N used in the algorithms for theorems 3.2 and 3.3, respectively. Two examples are given to demonstrate the mechanics of the algorithms for theorems 3.2, 3.3, and 3.4. In addition, numerous corollaries, rules, and remarks appear throughout section 3 and, although some may be trivial, they can be helpful in many cases to establish local connectedness.

Paper No. BU-410-M in the Biometrics Unit Mimeograph Series, Department of Plant Breeding and Biometry, Cornell University, Ithaca, New York 14850.

In dealing with global connectedness, one's attention is focused on each experimental unit of the design rather than a treatment as in the case of local connectedness. Thus characterization of globally connected designs is a difficult and somewhat tricky task, as demonstrated by theorem 4.1. This theorem gives a set of necessary and sufficient conditions, which must hold simultaneously, for a design to be globally connected as defined by Hedayat [5]. Two examples are given, one of which gives rise to the introduction of a further type of connected design, namely a pseudo-globally connected design, which is in effect a compromise between globally and locally connected designs. A set of necessary and sufficient conditions which, as in theorem 4.1, must hold simultaneously are given in theorem 5.1 of section 5. Further discussion of pseudo-globally connected designs in comparison with locally and globally connected designs and an example are also presented.

Some invariance properties of both locally and globally connected designs are presented in section 6. In addition we have considered the proposition of combining connected designs so that the newly composed design has some connected nature.

There is a strong analogy between some graph theory concepts and experimental design theory. Several theorems and concepts from Harary [4] and Busacker and Saaty [2] yield graph theoretic analogies to several of the theorems in section 3 and further properties and notions applicable to experimental (block) design.

2. Background

Let $r = \{1, 2, \cdots, v\}$ be a set of v treatments. By a block design with parameters v; b; $r_1, r_2, \cdots, r_v$; $k_1, k_2, \cdots, k_b$; and incidence structure N denoted by $BD(v; b; r_1, \cdots, r_v; k_1, \cdots, k_b; N)$ on Ω we shall mean an allocation of elements of Ω, one on each of the $m = k_1 + \cdots + k_b$ experimental units arranged in b blocks or groups of experimental units designated by B_j, $j = 1, 2, \cdots, b$ with B_j of size k_j such that i is assigned into r_i experimental units. $N = [n_{ij}]$ is the $v \times b$ incidence matrix where n_{ij} denotes the number of experimental units in the j^{th} block receiving the i^{th} treatment.

There are numerous special cases of the above block design. If for all i, $r_i = r$ (constant) then we have $BD(v; b; r; k_1, \cdots, k_b; N)$ which is called an equi-replicated block design. A proper block design is one that for all j, $k_j = k$ (constant), i.e., $BD(v; b; r_1, \cdots, r_v; k; N)$. Thus a block design that has for all i, $r_i = r$ and for all j, $k_j = k$ is denoted by $BD(v, b, r, k, N)$.

A block design is said to be pairwise balanced if

$$NN' = T + \lambda J$$

where T is a diagonal matrix, λ is a scalar, and J is a matrix of ones. For a proper block design that is pairwise balanced we write $BD(v, b, r_1, \cdots, r_v, k, \lambda)$ and if the design is also equireplicated then we write $BD(v, b, r, k, \lambda)$ which is the classical balanced incomplete block design with parameters v, b, r, k, and λ, i.e., $BIBD(v, b, r, k, \lambda)$.

A block that contains only one treatment, though possibly more than one experimental unit, is called a singleton, i.e., $B_j = [i, i, \cdots, i]$ or $[i]$. In experimental design theory NN' is called the association matrix and $N'N$ the block characteristic matrix.

In the following sections of this paper the statistical model underlying the discussion is one with fixed block effects. If the goal of an experiment is to estimate elementary treatment contrast, $i - j$, unbiasedly then connectedness is most important since every elementary treatment contrast is estimable if and only if the design is connected. The analysis of interest is the intrablock.

Model:

$$E(y_{ij}) = \mu + \beta_j + t_i \qquad i = 1, 2, \cdots, v$$
$$j = 1, 2, \cdots, b.$$

β_j = effect of B_j (fixed), t_i = effect of i, μ = general or mean effect.

From the normal equations we have

$$C\hat{\underline{t}} = Q \tag{2.1}$$

where

$\hat{\underline{t}}$ is the vector of estimated treatment effects,

$$C = \text{diag}(r_1, r_2, \cdots, r_v) - N\left\{\text{diag}[k_1^{-1}, k_2^{-1}, \cdots, k_b^{-1}]\right\}N', \tag{2.2}$$

and

$$Q = T - N\left\{\text{diag}[k_1^{-1}, k_2^{-1}, \cdots, k_b^{-1}]\right\}B \tag{2.3}$$

for

T = column vector of treatment totals, and

B = column vector of block totals.

Equation (2.1) is known as the equation for estimating the treatment effect and the matrix defined by (2.2) is the well-known C-matrix. One can easily verify that the linear function $\ell_1 t_1 + \ell_2 t_2 + \cdots + \ell_v t_v$ is estimable if and only if $\ell_1 + \ell_2 + \cdots + \ell_v = 0$, in which case the linear function $\underline{\ell}'\underline{t}$ is called a contrast. Elementary contrasts are those of the form $t_i - t_j$. The best linear unbiased estimator of $t_i - t_j$ is $\hat{t}_i - \hat{t}_j$ where $\hat{t}_i$, i = 1, 2, $\cdots$, v, is given by the solution of equation (2.1). Thus, obviously, the C-matrix plays a decisive role in the estimation of contrasts and hence the connectedness of a design.

Chakrabarti [3] defines a design to be connected if its C-matrix has rank v - 1. However, Bose [1] the originator of the concept of connectedness, defined connectedness in the following way:

> "A treatment and block are said to be associated if the treatment is contained in the block. Two treatments, two blocks, or a treatment and a block may be said to be connected if it is possible to pass from one to the other by means of a chain consisting alternately of blocks and treatments such that any two members of a chain are associated. A design (or a portion of a design) is said to be a connected design (or a connected portion of a design) if every block or treatment of the design (or a portion of the design) is connected to every other."

Unbiased estimators of an elementary treatment contrast can be obtained directly from the chains connecting the treatments of the contrast. For example consider the chain $1B_3 2$, where $1B_3$ denotes the observed response of 1 in block 3, y_{13}, and $B_3 2$ denotes the observed response of 2 in block 3 and is given a negative sign, $-y_{23}$. Thus $1B_3 2$ means $y_{13} - y_{23}$ which is an unbiased estimator of 1 - 2.

<u>Example 2.1.</u> Consider the design BD(3,3,2,2,1) $r = \{1,2,3\}$

B_1	B_2	B_3
1	1	2
2	3	3

The C matrix of this design is $\begin{bmatrix} 1 & -\frac{1}{2} & -\frac{1}{2} \\ -\frac{1}{2} & 1 & -\frac{1}{2} \\ -\frac{1}{2} & -\frac{1}{2} & 1 \end{bmatrix}$, which has rank = 2, i.e., v - 1.

Let y_{ij} be the observed response of treatment i in block j; then we have the following chains and their unbiased estimators.

Contrast	Chains		Estimators	
1 - 2	(i) $1B_1 2$	(ii) $1B_2 3B_3 2$	(i) $y_{11} - y_{12}$	(ii) $y_{12} - y_{32} + y_{33} - y_{23}$
1 - 3	(i) $1B_2 3$	(ii) $1B_1 2B_3 3$	(i) $y_{12} - y_{32}$	(ii) $y_{11} - y_{21} + y_{23} - y_{33}$
2 - 3	(i) $2B_3 3$	(ii) $2B_1 1B_2 3$	(i) $y_{23} - y_{33}$	(ii) $y_{21} - y_{11} + y_{12} - y_{32}$

Thus the design is connected under both definitions.

Example 2.2. Consider the following block design

$$\begin{array}{ccc} B_1 & B_2 & B_3 \\ \boxed{\begin{matrix}1\\2\\3\end{matrix}} & \boxed{\begin{matrix}1\\2\\3\end{matrix}} & \boxed{\begin{matrix}4\\5\\5\end{matrix}} \end{array}$$

$$C = \begin{bmatrix} \frac{16}{9} & -\frac{2}{9} & -\frac{2}{9} & 0 & 0 \\ -\frac{2}{9} & \frac{16}{9} & -\frac{2}{9} & 0 & 0 \\ -\frac{2}{9} & -\frac{2}{9} & \frac{16}{9} & 0 & 0 \\ 0 & 0 & 0 & \frac{8}{9} & \frac{16}{9} \\ 0 & 0 & 0 & \frac{16}{9} & \frac{14}{9} \end{bmatrix}$$

which has rank = 3, i.e., $\neq v - 1$. One cannot construct a chain to pass from 4 or 5 to 1, 2, or 3. Thus under both Chakrabarti's and Bose's definitions the design is not connected. The equivalence, in general, of Bose's and rank C definitions has been proved in [3].

The obvious drawback of both the above definitions of connectedness is that they are difficult and time consuming to use. The goal of this paper is to develop simpler theorems and procedures than the above for determining whether or not a design is connected. In addition we have extended the concept of connectedness to further classify a connected design as either locally or globally connected. The terms locally and globally connected are defined in the following paragraph. Firstly let us consider local connectedness. Two treatments, i and j, $i \neq j$, of a design are said to be locally connected if one can construct a chain, as defined by Bose [1], between a replicate of i and a replicate of j.

Definition 2.1. A design is said to be locally connected if every pair of treatments is locally connected. [N.B. Our definition of locally connected designs is the same as the connected designs of Bose.]

If we allow a treatment to be locally connected to itself the above definition still holds and the relationship (R), i locally connected to j, defines an equivalence relation on Ω since

(i) $(ii) \in R$, for all $i \in \Omega$, i.e., every treatment is locally connected to itself,

(ii) $(ij) \in R$ then $(ji) \in R$, for all i and $j \in \Omega$, i.e., if i is locally connected to j then j is locally connected to i, and

(iii) $(ij) \in R$ and $(jk) \in R$ then $(ik) \in R$, for all i,j,k, i.e., if i is locally connected to j and j is locally connected to k, then i is locally connected to k.

Lemma 2.1. *A design is locally connected if by the above equivalence relation, R, there is only one equivalence class.*

Hedayat [5] defined two treatments i and j, $i \neq j$, in D to be globally connected if for any replication of i and any replication of j one can construct a chain, as defined by Bose [1], to pass from i to j. The replications of i and j to be connected can each appear only once in the chain between them, namely at the beginning and end, respectively.

Definition 2.2. A design is said to be globally connected if every pair of treatments is globally connected.

The relationship i globally connected to j defines an equivalence relation, R^*, on Ω. The proof is analogous to the locally connected case.

Lemma 2.2. A design is globally connected if, by the above equivalence relation, R^*, there is only one equivalence class.

3. Results for Locally Connected Designs

Theorems and algorithms for determining the local connectedness of a design are presented in this section. The incidence matrix, N, association matrix, NN', and block characteristic matrix, N'N, are utilized in the following theorems and algorithms. Some corollaries, remarks, and rules for special cases are also given, along with a few examples. From this point on we shall denote the general block design with no restrictions on any of its parameters by $D = [B_1, B_2, \cdots, B_b]$.

Theorem 3.1. Design, D, is locally connected if and only if its incidence matrix, N, cannot be partitioned as follows:

$$N = \begin{bmatrix} N_1 & & & 0 \\ & N_2 & & \\ & & \ddots & \\ 0 & & & N_a \end{bmatrix}, \; 1 < a \le v, \; N_i \text{ are matrices}$$

N_i are connected subsets of the set of treatments.

Proof: If N cannot be partitioned as above then there is only one equivalence class of the relationship of connectedness, and vice versa.

<u>Corollary 3.1.</u> NN' <u>and</u> N'N <u>can be partitioned similar to</u> N <u>if and only if</u> N <u>can be partitioned as in theorem 4.1.</u>

<u>Remark 3.1.</u> The treatments of a disconnected design can be grouped into connected subsets as reflected by N, NN', and N'N, which is analogous to the breaking up of a Markov chain into closed states or sets of states.

<u>Remark 3.2.</u> N can be replaced by C and theorem 4.1 still holds.

It is obvious that theorem 4.1 is of only limited practical value. However, numerous conditions that are either sufficient or necessary can be found; though some may be trivial, they can be used in many cases to establish local connectedness.

The following are some rules to help in establishing whether or not a design is locally connected:

1. D is locally connected if N has a row or column with no zero elements, i.e., if a treatment appears in every block or a block contains every treatment, then D is locally connected.
2. D is locally connected if N has at least one non-zero element in row i, i = 2, 3, ···, v, below the non-zero elements of the 1^{st} row.
3. D is locally connected if N has more than v - d non-zero elements in row i, i = 2, 3, ···, v, and there are only d non-zero elements in the 1^{st} row.
4. D is NOT locally connected if NN' or N'N has a row with only one non-zero element.

N can be replaced by NN' and N'N in the above rules. In [1] there is a similar set of rules concerning the C matrix.

In an effort to find simpler necessary and sufficient conditions for local connectedness, NN' and N'N were investigated. The following theorems and algorithms are the results of that investigation.

<u>Theorem 3.2.</u> <u>D is locally connected if and only if there exists a set</u>

$$D^* = \left[B_1^*, B_2^*, \cdots, B_b^* \middle| B_s^* \in D \ \forall \ s = 1, 2, \cdots, b \ \underline{\text{and}} \ \exists \ \underline{\text{a}} \ q < p \ \underline{\text{such that}}\right.$$

$$\left. B_p^* \cap B_q^* \neq \phi \ \forall \ p = 2, 3, \cdots, b \right].$$

<u>Proof:</u> (i) <u>D^* exists:</u> D^* is just a reordering of the elements of D and $\bigcup_{s=1}^{b} B_s^* = \bigcup_{s=1}^{b} B_s \supseteq \Omega$. D^* implies that every treatment must appear in a block that contains at least two treatments. Thus each B_s^* must intersect with a B_r^*, $r \neq s$, that contains at least two treatments and the union of all blocks containing two treatments contains Ω. Hence we can construct a chain that passes through all the blocks containing two or more treatments and thus pass through every treatment.

(ii) <u>D^* does not exist:</u> If D^* does not exist then there is a B_p^* for which no B_q^* exists such that $B_p^* \cap B_q^* \neq \phi$, $q < p$, and the B_s^*'s can be grouped into disjoint sets of B_s^*. Thus the treatments contained in these disjoint sets of B_s^* form subsets of connected treatments and D is not locally connected.

Let us consider the set T_i, which has as elements the blocks that contain treatment i, and denote $\mathcal{T} = [T_1, T_2, \cdots, T_v]$.

<u>Theorem 3.3.</u> <u>D is locally connected if and only if there exists a set</u>

$$\mathcal{T}^* = \left[T_1^*, T_2^*, \cdots, T_v^* \middle| T_i^* \in \mathcal{T} \ \forall \ i = 1, 2, \cdots, v \ \underline{\text{and}} \ \exists \ \underline{\text{a}} \ j < i \ \underline{\text{such that}}\right.$$

$$\left. T_i^* \cap T_j^* \neq \phi \ \forall \ i = 2, 3, \cdots, v \right].$$

<u>Proof:</u> This proof is analogous to that of theorem 3.2.

If treatment i and j are connected by a chain we write this as (ij). Define the operator $\cdot$ (dot) by $(ij) \cdot (jk) = (ik)$; i.e., if i and j are connected and j and k are connected then, obviously, i and k are connected by a chain. Also, if i and j are connected by a chain then j and i are connected by a chain; i.e., (ij) = (ji). It should be noted that if a design is locally connected then there are v(v - 1) chains, excluding the chains of (ii).

Theorem 3.4. *D is locally connected if and only if there is a set, $\mathcal{U}$, with v - 1 elements each of the form (ij) ∈ D, such that under the dot operator, as defined above, the v(v - 1) necessary and sufficient chains can be generated.*

Before proving the theorem, note that if $\mathcal{U}$ exists every treatment appears in at least one element of $\mathcal{U}$. Under the dot operator each element gives rise to v - 1 other chains plus its reverse; i.e., (ij) = (ji). Thus total number of chains is v(v - 1) since there are (v - 1)(v - 2) chains by dot operator plus 2(v - 1) from the elements of $\mathcal{U}$ and their reverses.

Proof: (i) Sufficiency obvious.

(ii) If D is locally connected then every treatment is connected to every other treatment and $\mathcal{U}$ can be easily constructed.

The non-zero elements of NN' represent chains of the form iB_rj, which is the ij element. Thus $(NN')^2$ is in essence the result of the dot operation between the chains represented by non-zeros in NN' and in general $(NN')^a$, $2 \le a \le v - 1$, is equivalent to the dot operation between the non-zero elements of $(NN')^{a-1}$ and those of NN'. The longest possible chain between any two treatments is one which contains all the treatments; such a chain could be constructed by the dot operation between v - 1 chains of the form iB_rj with distinct B_r's. Thus the non-zero elements of $(NN')^{v-1}$ represent those pairs of treatments that are locally connected. We now have the following theorem.

Theorem 3.5. A necessary and sufficient condition for a design to be locally connected is that $(NN')^{v-1}$ have no zero entries.

For example, consider the following block design:

B_1	B_2	B_3
1	2	3
2	3	4

$$N = \begin{bmatrix} 1 & 0 & 0 \\ 1 & 1 & 0 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{bmatrix}, \qquad NN' = \begin{bmatrix} 1 & 1 & 0 & 0 \\ 1 & 2 & 1 & 0 \\ 0 & 1 & 2 & 1 \\ 0 & 0 & 1 & 1 \end{bmatrix},$$

$$(NN')^2 = \begin{bmatrix} 2 & 3 & 1 & 0 \\ 3 & 6 & 4 & 1 \\ 1 & 4 & 6 & 3 \\ 0 & 1 & 3 & 2 \end{bmatrix}, \qquad (NN')^3 = \begin{bmatrix} 5 & 9 & 5 & 1 \\ 9 & 19 & 15 & 5 \\ 5 & 15 & 19 & 9 \\ 1 & 5 & 9 & 5 \end{bmatrix}.$$

Thus the above design is locally connected.

Corollary 3.2. In the above theorem $(NN')^{v-1}$ can be replaced by $(N'N)^{b-1}$ and the condition remains necessary and sufficient.

Corollary 3.3. If any power of NN' or N'N have a row or column with no zero elements, then the design is locally connected.

The following algorithms for each of the preceding theorems, except theorem 3.5, will help clarify the theorems and demonstrate their applicability.

Algorithm for theorem 3.2.

Consider N'N, if (r,s) element of N'N matrix is not zero then $\Leftrightarrow B_r \cap B_s \neq \phi$,

$r \neq s$, or $B^*_m \cap B^*_n \neq \emptyset$, $m \neq n$ for some m,n. Thus inspect the 1^{st} row of N'N then all B_1 and B_r such that $B_1 \cap B_r \neq \emptyset$, become the first elements of D^*. For all r such that $B_1 \cap B_r \neq \emptyset$, go to the r^{th} row of N'N then all B_s such that $B_r \cap B_s \neq \emptyset$ become the next part of D^*. Continue until:

1. all blocks belong to $D^* \Leftrightarrow D$ is locally connected,
2. all treatments are contained in the blocks of a partially formed D^*, then the remaining blocks can be added in any order $\Leftrightarrow D$ is locally connected,
3. a B_r is found such that $B_r \cap B_s = \emptyset$ for all $s < r$ and $B_s \in D^*$, and also any B_u such that $B_u \cap B_r \neq \emptyset$ with $B_u \cap B_s = \emptyset \Leftrightarrow D$ is not locally connected.

Algorithm for theorem 3.3.

This is analogous to the above algorithm with NN' in place of N'N, T_i in place of B_r, and $\mathcal{J}^*$ in place of D^*.

Method for using theorem 3.4.

The easiest chains to construct are those between treatments in the same block, i.e., $iB_sj = (ij)$. Every treatment must exist in at least one chain of this form. A constraint to apply to elements of $\mathcal{U}$ is that if $(ij) \in \mathcal{U}$ and $(jk) \in \mathcal{U}$ then $(ik) \in \mathcal{U}$, i.e., don't include in $\mathcal{U}$ a chain that is inferred by the dot operator between two elements of $\mathcal{U}$.

Example 3.1. $\Omega = [1, 2, 3, 4, 5, 6, 7, 8]$

Design D:

B_1	B_2	B_3	B_4	B_5	B_6	B_7	B_8
1	2	1	4	7	6	8	3
3	7	3	5	7	6	7	4
4		5	6	8	3	2	5
					3		6

$$N = \begin{bmatrix} 1 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 1 & 0 \\ 1 & 0 & 1 & 0 & 0 & 2 & 0 & 1 \\ 1 & 0 & 0 & 1 & 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 1 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 1 & 0 & 2 & 0 & 1 \\ 0 & 1 & 0 & 0 & 2 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 1 & 0 \end{bmatrix}$$

<u>Example of algorithm for theorem 3.3.</u>

$$NN' = \begin{bmatrix} 2 & 0 & 2 & 1 & 1 & 0 & 0 & 0 \\ 0 & 2 & 0 & 0 & 0 & 0 & 2 & 1 \\ 2 & 0 & 7 & 1 & 1 & 5 & 0 & 0 \\ 1 & 0 & 1 & 3 & 2 & 2 & 0 & 0 \\ 1 & 0 & 1 & 2 & 3 & 2 & 0 & 0 \\ 0 & 0 & 5 & 2 & 2 & 6 & 0 & 0 \\ 0 & 2 & 0 & 0 & 0 & 0 & 6 & 3 \\ 0 & 1 & 0 & 0 & 0 & 0 & 3 & 6 \end{bmatrix}$$

From row 1 we have treatments 1, 3, 4, 5 connected; row 3 yields treatment 6; row 4 yields 6; and row 5 yields 6. Now row 6 adds no new treatments. <u>Thus the design is not locally connected</u>. (1, 3, 4, 5, 6) form a group of connected treatments and (2, 7, 8) also are a group of connected treatments.

<u>Example of algorithm for theorem 3.2.</u>

$$N'N = \begin{bmatrix} 3 & 0 & 2 & 1 & 0 & 2 & 0 & 2 \\ 0 & 2 & 0 & 0 & 2 & 0 & 2 & 0 \\ 2 & 0 & 3 & 1 & 0 & 2 & 0 & 2 \\ 1 & 0 & 1 & 3 & 0 & 2 & 0 & 3 \\ 0 & 2 & 0 & 0 & 5 & 0 & 3 & 0 \\ 2 & 0 & 2 & 2 & 0 & 8 & 0 & 4 \\ 0 & 2 & 0 & 0 & 3 & 0 & 3 & 0 \\ 2 & 0 & 2 & 3 & 0 & 4 & 0 & 4 \end{bmatrix}$$

Row 1 yields B_1, B_3, B_4, B_6, B_8 as the first part of D^*. But rows 3, 4, 6, and 8 add no new blocks, thus the design is not locally connected. B_2, B_5, and B_7 intersect pairwise and form a set disjoint from B_1, B_3, B_4, B_6, and B_8.

<u>Example of method for theorem 3.4.</u>

From B_1 we get (1,3) and (1,4); do not include (3,4) since $(1,3) \cdot (1,4) = (3,4)$. From B_2 we get (2,7), B_3 gives (1,5), B_4 gives (5,6), B_5 gives (7,8), and B_6, B_7, and B_8 give nothing that cannot be inferred by the dot operator. Thus we have (1,3), (1,4), (2,7), (1,5), (5,6), (7,8), only six elements, and seven are required. Also many chains are impossible to construct; e.g., (5,7), (5,8), etc. Thus the design is not locally connected.

<u>Example 3.2.</u> $\Omega = [1, 2, 3, 4, 5, 6, 7, 8]$

Design D:

B_1	B_2	B_3	B_4	B_5	B_6	B_7	B_8	B_9
1	4	7	2	7	1	2	1	2
2	5	8	4	6	7	5	3	3
3	6		8	5		6	7	
						8		

$$N = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 1 & 0 & 1 & 0 \\ 1 & 0 & 0 & 1 & 0 & 0 & 1 & 0 & 1 \\ 1 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 \\ 0 & 1 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 1 & 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 & 1 & 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 & 1 & 1 & 0 & 1 & 0 \\ 0 & 0 & 1 & 1 & 0 & 0 & 1 & 0 & 0 \end{bmatrix}$$

Example of algorithm for theorem 3.2.

$$N'N = \begin{bmatrix} 3 & 0 & 0 & 1 & 0 & 1 & 1 & 2 & 2 \\ 0 & 3 & 0 & 1 & 2 & 0 & 2 & 0 & 0 \\ 0 & 0 & 2 & 1 & 1 & 1 & 1 & 1 & 0 \\ 1 & 1 & 1 & 3 & 0 & 0 & 2 & 0 & 1 \\ 0 & 2 & 1 & 0 & 3 & 1 & 2 & 1 & 0 \\ 1 & 0 & 1 & 0 & 1 & 2 & 0 & 2 & 0 \\ 1 & 2 & 1 & 2 & 2 & 0 & 4 & 0 & 1 \\ 2 & 0 & 1 & 0 & 1 & 2 & 0 & 3 & 1 \\ 2 & 0 & 0 & 1 & 0 & 0 & 1 & 1 & 2 \end{bmatrix}$$

Row 1 yields B_1, B_4, B_6, B_7, B_8, B_9 as the first part of D^*, row 4 yields B_2 and B_3, and row 6 yields B_5. Thus

$$D^* = [B_1, B_4, B_6, B_7, B_8, B_9, B_2, B_3, B_5] \Rightarrow B_1^* = B_1, B_2^* = B_4, B_3^* = B_6,$$

and so on, and D is locally connected.

Example of algorithm for theorem 3.3.

$$NN' = \begin{bmatrix} 3 & 1 & 2 & 0 & 0 & 0 & 2 & 0 \\ 1 & 4 & 2 & 1 & 1 & 1 & 0 & 2 \\ 2 & 2 & 3 & 0 & 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 2 & 1 & 1 & 0 & 1 \\ 0 & 1 & 0 & 1 & 3 & 3 & 2 & 0 \\ 0 & 1 & 0 & 1 & 3 & 3 & 1 & 1 \\ 2 & 0 & 1 & 0 & 2 & 1 & 4 & 1 \\ 0 & 2 & 0 & 1 & 0 & 1 & 1 & 3 \end{bmatrix}$$

Row 1 yields T_1, T_2, T_3, T_7 and row 2 yields T_4, T_5, T_6, T_8. Thus

$$\mathcal{T}^* = [T_1, T_2, T_3, T_7, T_4, T_5, T_6, T_8] \Rightarrow T_1^* = T_1, T_2^* = T_2, T_3^* = T_3, T_4^* = T_7,$$

and so on, and D is locally connected.

Example of method for theorem 3.4.

B_1 yields (1,2) and (1,3)

B_2 yields (4,5) and (4,6)

B_3 yields (7,8)

B_4 yields (2,4) and (2,8).

Thus $\mathcal{U} = \left[(1,2)\ (1,3)\ (4,5)\ (4,6)\ (7,8)\ (2,4)\ (2,8)\right]$, 7 elements. Every treatment appears in at least one of the elements and a chain can be constructed that passes through every treatment. Also no element of $\mathcal{U}$ can be formed by the dot operation between any two elements of $\mathcal{U}$. Thus D is locally connected.

4. Results for Globally Connected Designs

An advantage of globally connected designs is that when estimating the elementary contrast between the effects of i and j, every replicate participates, yielding $r_i \times r_j$ estimates of i - j or j - i. The following theorem characterizes globally connected designs.

Theorem 4.1. A design, D, is globally connected if and only if the following conditions hold simultaneously.

(1) D is locally connected.

(2) Every block of D contains at least two treatments that appear in more than one block; i.e., for all $B_s \in D$ there exists an i and $j \in B_s$ such that $i \in B_r$ and $j \in B_u$, $u \neq s$ and $r \neq s$.

(3) If there exists a $B_s \in D$ such that $B_s = \{i,j\}$ or $\{i,j,k \cdots\}$ where all experimental units of B_s, except i and j, appear in B_s only, then i and j must each belong to two other blocks of D. That is, if i and j are the only treatments contained in B_s that appear in other B_r's, $r \neq s$, then i and j must be such that $i \in B_r$ and B_u, and $j \in B_m$ and B_n, $r \neq s$, $u \neq s$, $m \neq s$, and $n \neq s$.

(4) Any treatment, i say, that appears in two or more blocks (but not all blocks) must do so in blocks that contain

(i) a treatment that appears in two blocks containing i, and two not containing i. That is, $i \in B_r$ and B_s and there exists a $j \in B_r$, B_s, B_m, and B_n where $i \notin B_m$ and $i \notin B_n$,

or

(ii) two treatments each appearing in a block containing i, and a block not containing i. That is, i and $j \in B_r$, i and $k \in B_s$, then $j \in B_m$ and $k \in B_n$ with $i \notin B_m$ and $i \notin B_n$.

Some of these conditions may seem redundant; however, with a few simple examples we will show that this is not the case.

Case 1: Let D_1 and D_2 be globally connected designs of the treatment sets Ω_1 and Ω_2, respectively, and $\Omega_1 \cap \Omega_2 = \phi$. Thus theorem 4.1 holds for each design. Consider the design $D_1 \cup D_2$ of the set of treatments $\Omega_1 \cup \Omega_2$; conditions (2), (3), and (4) of theorem 4.1 will still hold. However, the design $D_1 \cup D_2$ is not locally connected, i.e., condition (1) no longer holds and the design is not globally connected.

Case 2: Let (1), (3), and (4) of theorem 4.1 hold, but not (2).

	B_1	B_2	B_3	B_4	B_5	B_6
D:	1	1	1	1	3	3
		2	2	2	4	4
				3		
				4		

D is not globally connected since the replicate of treatment 1 in B_1 is not connected to any other treatment.

<u>Case 3</u>: Let (1), (2), and (4) of theorem 4.1 hold, but not (3) for treatment 2.

D:

B_1	B_2	B_3	B_4	B_5
1	1	1	3	3
2	4	3	4	4
	2		5	5

D is not globally connected since the replicate of treatment 1 in B_1 is not connected to the replicate of treatment 2 in B_2.

<u>Case 4</u>: Let (1), (2), and (3) of theorem 4.1 hold, but not (4).

D:

B_1	B_2	B_3	B_4
1	1	3	4
2	2	4	5
3	3	5	6
		6	

D is not globally connected since, for example, the replicate of treatment 3 in B_3 is not connected to treatment 1 or 2.

<u>Proof of theorem 4.1.</u>

<u>Necessity</u>

(i) Condition (1) is obvious.

(ii) If condition (2) is violated then D has a singleton. The treatment belonging to the singleton cannot be connected by a chain to any other treatment and so it follows that D is not globally connected.

(iii) If condition (3) is violated by treatment i of block B_r, then i occurs in only one other block, B_s. A chain between any treatment in B_r and i in B_s cannot be constructed. Thus D is not globally connected.

(iv) If condition (4) does not hold for treatment i say, then there is a treatment j which occurs in at least two blocks containing i, and exactly one not containing i, B_r. It follows that one cannot construct chains between all the replications of j and i, namely the replicate of $j \in B_r$ and any replicate of i. Thus D is not globally connected.

Sufficiency

Consider any replicate of any treatment, say replicate x of treatment i, and denote as i^x. Then given that the conditions hold, can i^x be connected by a chain to any replicate of any other treatment, say k^y? Now by condition (2), if $i^x \in B_s$ then there exists a $j \in B_s$ such that we have $i^x B_s j$. Since the design is locally connected we can construct a chain between j and k. If j is connected to k^y, then we are finished. However, if j is connected to k^z, $z \neq y$, then since the blocks containing k^z and k^y satisfy the conditions (2), (3), and (4), a chain between k^z and k^y can be constructed. This completes the proof.

Corollary 4.1. If two treatments appear in every block, then the design is globally connected. (The design must have at least three blocks.)

Corollary 4.2. If N has no zero elements, then D is globally connected. (If N has no zero elements, then NN' and N'N have no zero elements.)

Example 4.1. Example 2 of section 3.

	B_1	B_2	B_3	B_4	B_5	B_6	B_7	B_8	B_9
	1	4	7	2	5	1	2	1	2
D:	2	5	8	4	6	7	5	3	3
	3	6		8	7		6	7	
							8		

(1) From section 4 we know that <u>D</u> is <u>locally</u> <u>connected</u>.

(2) Condition (2) is satisfied; i.e., all blocks have at least two treatments that appear in more than one block.

B_1:	for	example	$1 \in B_6$	and	$2 \in B_4$
B_2:	"	"	$4 \in B_4$	"	$5 \in B_5$
B_3:	"	"	$7 \in B_5$	"	$8 \in B_4$
B_4:	"	"	$4 \in B_2$	"	$2 \in B_1$
B_5:	"	"	$5 \in B_7$	"	$6 \in B_7$
B_6:	"	"	$1 \in B_1$	"	$7 \in B_8$
B_7:	"	"	$2 \in B_1$	"	$5 \in B_5$
B_8:	"	"	$1 \in B_1$	"	$3 \in B_1$
B_9:	"	"	$2 \in B_1$	"	$3 \in B_1$

(3) The only blocks concerned with condition (3) are B_3, B_6, and B_9.

(i) 7 and $8 \in B_3$, $7 \in B_5$, B_6, and B_8, $8 \in B_4$ and B_7

(ii) 1 and $7 \in B_6$, $7 \in B_3$, B_5, and B_8, $1 \in B_1$ and B_8

(iii) 2 and $3 \in B_9$, $2 \in B_1$, B_4, and B_7, $3 \in B_1$ and B_8

Thus condition (3) is satisfied.

(4) All treatments appear in two or more blocks, thus they all must satisfy condition 4.

(i) $1 \in B_1$, B_6, and B_8; note that $3 \in B_1$ and B_9, $2 \in B_1$ and B_4, but $1 \notin B_4$ or B_9.

(ii) $2 \in B_1$, B_4, B_7, and B_9; note that $5 \in B_5$ and B_7, $6 \in B_5$ and B_7, but $2 \notin B_5$.

(iii) $3 \in B_1$, B_8, and B_9; note that $2 \in B_1$ and B_9, and also $2 \in B_4$ and B_7, but $3 \notin B_4$ or B_7.

(iv) $4 \in B_2$ and B_4; 5 and $6 \in B_2$ and B_7, but $4 \notin B_7$.

(v) $5 \in B_4$, B_5, and B_7; $2 \in B_4$ and B_1, $8 \in B_7$ and B_3, but $5 \notin B_1$ and B_3.

(vi) $6 \in B_2$, B_5, and B_7; $2 \in B_7$ and B_4, $4 \in B_2$ and B_4, but $6 \notin B_4$.

(vii) $7 \in B_3$, B_5, B_6, and B_8; $8 \in B_3$ and B_4, $3 \in B_8$ and B_9, but $7 \notin B_4$ or B_9.

(viii) $8 \in B_3$, B_4, and B_7; $2 \in B_7$, B_4, B_1, and B_9, but $8 \notin B_1$ or B_9.

All four conditions of theorem 4.1 are satisfied and thus the design is globally connected. Most of the calculations in the above example can be done mentally; it has been presented in an attempt to clarify any misunderstandings or confusion the reader may have had.

Example 4.2.

D:

B_1	B_2	B_3
1^1	1^2	2^2
2^1	3^1	3^2

Condition (3) of theorem 4.1 does not hold and the design is not globally connected. However, for any two treatments of the design, say 1 and 2, every replicate of 1 is connected by a chain to at least one replicate of 2. That is,

$1^1B_12^1$; $1^2B_23B_32^2$ $\Rightarrow$ each chain yields an estimator of 1 - 2

also $1^1B_12B_33^2$; $1^2B_23^1$ $\Rightarrow$ " " " " " " 1 - 3

and $2^1B_11B_23^1$; $2^2B_33^2$ $\Rightarrow$ " " " " " " 2 - 3.

Any two treatments connected in this fashion are said to be pseudo-globally connected. A more concise definition and some discussion of pseudo-globally connectedness is given in the next section.

5. Pseudo-Globally Connected Designs

Pseudo-globally connected designs assures one that in estimating elementary contrasts each replicate of the treatments involved is utilized. When estimating elementary treatment contrasts, globally connected designs maximize the use of all replicates of the treatments whereas pseudo-globally connected designs guarantee that no replicates are "wasted". That is, every replicate of each treatment in the contrast is involved at least once in the estimation.

From example 4.2 in section 4 we have the following definition:

Definition 5.1. Two treatments i and j, $i \neq j$, in D are said to be pseudo-globally connected if every replicate of i is connected by a chain to at least one replicate of j, and vice-versa. D is said to be pseudo-globally connected if every pair of treatments is pseudo-globally connected.

The relationship i pseudo-globally connected to j defines an equivalence relation on D, and D is pseudo-globally connected if there is only one equivalence class for all treatments. This class of connected designs covers the ground between local and global connectedness. Pseudo-globally connected designs use each replicate of treatments i and j, $i \neq j$, at least once in the estimation of the contrast i - j, for all i and $j \in \Omega$.

Theorem 5.1. Conditions (1), (2), and (4) of theorem 4.1 are necessary and sufficient for a design to be pseudo-globally connected.

Proof: This proof is analogous to that of theorem 4.1.

Example 5.1.

	B_1	B_2	B_3	B_4	B_5
D:	1	1	2	3	4
	2	5	5	5	5
	3				
	4				

D satisfies theorem 5.1 and the required chains can be constructed. Thus D is pseudo-globally connected.

6. Some Further Properties of Connected Designs

A. Invariance Properties: If a design D is locally (globally) connected then any of the following can occur and D will remain locally (globally) connected.

(a) For D locally connected:

(i) any treatment belonging to Ω can be added to any block of D.

(ii) any new treatment(s) can be added in any block of D.

(b) For D globally connected:

(i) any block belonging to D can be repeated any number of times.

(ii) if a treatment appears in a block it can be replicated any number of times within that block.

(iii) if a block does not contain treatment j say, then j can be added to that block.

Recall that if a design is globally connected then it is pseudo-globally connected, which also implies that the design is locally connected. Thus the facts in (b) above apply to pseudo-globally and also locally connected designs. Similarly, in the remainder of this section the results for globally connected designs are inferred to pseudo-globally connected designs which are not specifically mentioned.

B. The Composition of Connected Designs: Let us consider the proposition of composing designs that are locally or globally connected.

(a) Compositions that yield locally connected designs:

(i) If D_1 and D_2 are locally connected designs on the sets of treatments Ω_1 and Ω_2, respectively, and $\Omega_1 \cap \Omega_2 = \phi$, then the design

$\overline{D}_\ell = D_1 \cup D_2 \cup B$ is locally connected, where B is a block containing at least two treatments, i and j say, such that $i \in \Omega_1$ and $j \in \Omega_2$. The block B forms the link between the two designs D_1 and D_2. Since i is connected to all treatments in Ω_1 and j to all in Ω_2 then the chain iBj locally connects every pair of treatments of $\Omega_1 \cup \Omega_2$.

(ii) Let D_1 and D_2 be locally connected designs on Ω_1 and Ω_2, respectively, and if $\Omega_1 \cap \Omega_2 \neq \phi$, i.e., Ω_1 and Ω_2 have at least one element in common, then $D_1 \cup D_2$ is a locally connected design.

(b) Compositions that yield globally connected designs:

(i) Consider D_1 and D_2 to be globally connected designs of treatment sets Ω_1 and Ω_2, respectively, $\Omega_1 \cap \Omega_2 = \phi$. As before, $\overline{D}_g = D_1 \cup D_2 \cup B$ where B, as above, is locally connected. However, if B contains four treatments (i, j, k, and ℓ), such that i and $j \in \Omega_1$ and k and $\ell \in \Omega_2$, also i and j each appear in at least two blocks of D_1 and similarly k and ℓ in D_2, then $\overline{D}_g$ is globally connected. Moreover, if B contains three treatments of Ω_1 and three of Ω_2 then $\overline{D}_g$ is globally connected. It is easily shown that $\overline{D}_g$, with the above B's, satisfies theorem 4.1.

(ii) For $D_1 \cup D_2$ to be globally connected, it is sufficient for D_1 and D_2 each to be globally connected and one of the following:

(1) $\Omega_1 \cap \Omega_2 = [i]$ and i appears in two blocks of D_1 and two of D_2.

(2) $\Omega_1 \cap \Omega_2 = \begin{bmatrix} i \\ j \end{bmatrix}$ and i appears in at least one block of D_1 and two of D_2, while j appears in at least one block of D_2 and two of D_1.

(3) $\Omega_1 \cap \Omega_2 = \begin{bmatrix} i \\ j \\ k \end{bmatrix}$.

Since the proofs of the above are simple and straightforward, we will not bother to present them here. Also it is interesting to note that two designs, D_1 and D_2, can each be not locally connected but their union $D_1 \cup D_2$ may be locally connected. This is obvious since given a locally connected design, D, one can usually partition D into locally disconnected subsets.

7. Graph Theoretical Analogy

A graph G is a mathematical system consisting of two sets V and E. V is a finite nonempty set of p vertices and E is a prescribed set of q unordered pairs of distinct vertices of V. Each pair $e = \{u,v\}$ of vertices in E is an edge of G and e is said to join u and v. We write $e = uv$ and say that u and v are adjacent vertices, vertex u and edge e are incident with each other, as are v and e. Two distinct edges incident with a common vertex are said to be adjacent edges.

A walk of a graph is an alternating sequence of vertices and edges beginning and ending with vertices in which each line is incident with the two vertices immediately preceding and following it. A trail is a walk with all edges distinct and a path is one with all vertices distinct. Harary [4] defines a graph to be connected if every pair of vertices are joined by a path.

We define the treatments of a design to be the vertices of a graph G, and two vertices are incident if the two treatments belong to the same block. A walk of a graph is equivalent to a treatment-block chain as defined by Bose [1]. Thus, if every pair of vertices is connected by a walk then the design will be locally connected, and vice versa. Also we can define a design to be locally connected if and only if the graph G is connected.

<u>Analogue to theorem 3.1.</u> The design D is locally connected if and only if the graph G, as defined above, has only one connected component.

Define the graph G(D) to have as vertices the blocks of the design D and two vertices as incident if the two blocks have at least one treatment in common.

<u>Analogue to theorem 3.2.</u> D is locally connected if and only if every pair of vertices of G(D) is connected by a walk.

If the T_i, as defined in section 3, are the vertices of graph $G(\mathcal{T})$ and T_i and T_j, $i \neq j$, are incident if there is a $B_s \in T_i$ and a $B_r \in T_j$ such that $B_r \cap B_s \neq \emptyset$, then we have the following:

<u>Analogue to theorem 3.3.</u> D is locally connected if and only if every pair of vertices of $G(\mathcal{T})$ is connected by a walk.

As before, we can develop some simple rules, in graph theory terms, for determining the local connectedness of D.

(i) D is not locally connected if any of the above graphs has an isolation vertex.

(ii) If any vertex, v, of the above graphs has degree (number of edges incident with v) p - 1, where p is the number of vertices, then D is locally connected.

The removal or loss of treatments from a design obviously can affect the local connectedness of that design. Knowledge of treatments, which by their removal or loss cause the design to be not locally connected, would usually be of interest to the experimenter. A similar situation arises in graph theory. Busacker and Saaty [2] define a vertex v to be a <u>point of articulation</u> of a connected graph if the graph obtained by deleting v and all edges incident with v is disconnected. A graph is said to be separable if it has at least one articulation point.

Lemma 7.1. A necessary and sufficient condition for a vertex v to be a point of articulation is that v lie on all the paths connecting some pair of vertices.

Proof: See Busacker and Saaty [2].

A matrix of interest in graph theory is the vertex or adjacency matrix V. The element in the (i,j) position of V is the number of edges incident with both vertex i and vertex j. From Busacker and Saaty [2] we have the following theorem:

Theorem 7.1. The matrix V^n gives the number of walks of length n between any two vertices, where the length of a walk is the number of edges between the beginning and terminating vertices.

The analogous experimental design theory for the above terminology and theory is obvious. A treatment or block is said to be a point of articulation if the design obtained by deleting that treatment or block is not locally connected.

Lemma 7.2. A necessary and sufficient condition for a treatment or block to be a point of articulation is that it lie on all chains connecting some pair of treatments.

If we define the length of a chain to be the number of blocks that appear in the chain, then a treatment matrix can be defined similar to the vertex matrix of a graph. A treatment matrix Λ has as its (i,j) element the number of blocks that contain both treatments i and j.

Theorem 7.2. The matrix Λ^n gives the number of chains of length n between any two treatments of a design.

Globally and pseudo-globally connected designs were not considered in graph theory terms and it is doubtful if an analogy to theorems 4.1 and 5.1 would be of any use. However, graph theory is a powerful tool and in future research may yield some further facts and theorems on connectedness of designs in general.

8. Conclusion

The next step in the study of connected designs is to show, by some "optimality" criterion, that globally connected designs are better than locally connected designs. The criterion selected as the basis of comparison will, of course, be the determining factor. However, for any "reasonable" criterion it is difficult to conceive of a situation where a locally connected design would be better than all possible globally connected designs for a given set of parameters. The optimality criterion should be some sort of variance function.

REFERENCES

[1] Bose, R. C. Presidential Address, Proceedings of the 34^{th} Indian Science Congress (1947).

[2] Busacker, R. and Saaty, T. Finite Graphs and Networks. McGraw-Hill (1965).

[3] Chakrabarti, M. C. On the C-matrix in Design of Experiments. J.I.S.A., Vol. 1 (1963).

[4] Harary, F. Graph Theory. Addison-Wesley (1969).

[5] Hedayat, A. A set of conditions which every locally connected design must satisfy. BU-392-M, Cornell University (1971).