

# The Diclque Representation and Decomposition of Binary Relations

ROBERT M. HARALICK

*The University of Kansas, Lawrence, Kansas*

**ABSTRACT.** The binary relation is often a useful mathematical structure for representing simple relationships whose essence is a directed connection. To better aid in interpreting or storing a binary relation we suggest a diclque decomposition. A diclque of a binary relation  $R$  is defined as an ordered pair  $(I, O)$  such that  $I \times O \subseteq R$  and  $(I, O)$  is maximal. In this paper, an algorithm is described for determining the diclques of a binary relation; it is proved that the set of such diclques has a nice algebraic structure. The algebraic structure is used to show how diclques can be coalesced, the relationship between cliques and diclques is discussed, and an algorithm for determining cliques from diclques is described.

**KEY WORDS AND PHRASES:** graphs, directed graphs, digraphs, cliques, lattices, clustering, diclques, monoids, systems

**CR CATEGORIES:** 3.29, 3.39, 3.63, 5.32

## 1. Introduction

The binary relation is often a useful mathematical structure for representing simple relationships whose essence is a directed connection. One important application for the binary relation occurs in the field of general systems theory where a binary relation  $R$  can be constructed from a particular discipline theory or constructed empirically from a data set, where the relation  $R$  is considered to be the set of all pairs of variables  $(x, y)$  where variable  $x$  in some sense influences, controls, or dominates variable  $y$ . Typically, there are three problems associated with such a binary relation: (1) drawing the binary relations as a digraph, (2) storing the binary relation in the computer, and (3) using the binary relation to get clues concerning which sets of variables form subsystems and for each subsystem determining the input variables and the output variables.

Each is a problem because in large systems the variables number in the hundreds or thousands and the connections number in the thousands or ten thousands. These three problems are really equivalent, for they are problems of organization. If I can organize the binary relation, I can draw it or store it and the organization itself will reveal the subsystems.

It is the purpose of this note to define a decomposition of the binary relation which can prove useful in such organizing tasks and to prove some theorems, which are constructive in nature, that will indicate the algorithms to carry out the decomposition.

The binary relation organization is based around a representation which we call the diclque representation; the parts to the decomposition we call diclques. The term di-

Copyright © 1974, Association for Computing Machinery, Inc. General permission to republish, but not for profit, all or part of this material is granted provided that ACM's copyright notice is given and that reference is made to the publication, to its date of issue, and to the fact that reprinting privileges were granted by permission of the Association for Computing Machinery.

This work was supported by a University Research Grant from The University of Kansas general research fund.

Author's address: The University of Kansas Space Technology Center, 2291 Irving Hill Drive—Campus West, Lawrence, KS 66045.

$x$	$R(x)$
1	1, 2, 3, 4, 6, 7, 9, 10, 11, 12, 14, 15
2	1, 2, 3, 4, 6, 7, 9, 10, 11, 12, 14, 15
3	1, 2, 3, 4, 6, 7, 9, 10, 11, 12, 14, 15
4	8, 13, 16, 17, 18, 19, 20, 21, 22, 23, 24
5	1, 2, 3, 5, 6
6	4, 5, 7, 9, 10, 11, 12, 14, 15
7	4, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24
8	1, 2, 3, 6
9	4, 7, 9, 10, 11, 12, 13, 14, 15, 17, 18, 19, 20, 21, 22, 23
10	4, 7, 9, 10, 11, 12, 13, 14, 15, 17, 18, 19, 20, 21, 22, 23
11	8, 13, 16, 17, 18, 19, 20, 21, 22, 23, 24
12	5
13	1, 2, 3, 6, 13, 17, 18, 19, 20, 21, 22, 23
14	5
15	5
16	
17	5
18	13, 17, 18, 19, 20, 21, 22, 23
19	8, 13, 16, 24
20	13, 17, 18, 19, 20, 21, 22, 23
21	4, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 24
22	8, 13, 16, 24
23	8, 13, 16, 24
24	1, 2, 3, 6

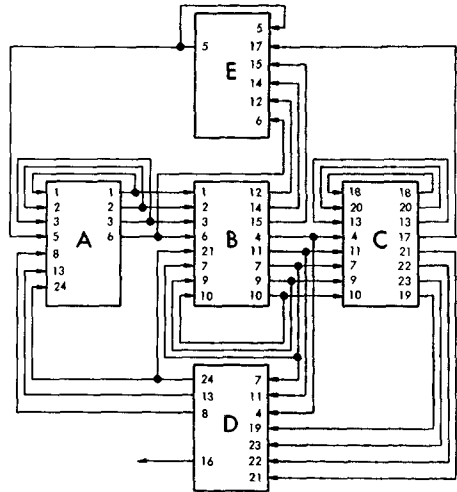


FIG. 1. Illustration of a complicated relation and a diagram of five of its dicliques which cover the relation

clique is used because a diclique is like a two-sided clique. Graph theorists will identify a diclique with a maximally complete bipartite subgraph [1].

To motivate our development we provide the example of Figure 1. The table of Figure 1 indicates that there are 188 ordered pairs in the example binary relation. We can well imagine that the digraph corresponding to the example relation would be hard to draw and messy to interpret. Yet if the digraph is organized according to five of its dicliques (there are many more than five dicliques for this relation) which cover the relation, we obtain the system diagram shown in Figure 1. This system diagram has on it exactly every connection which the digraph would have. Notice that the complicated system has been decomposed into subsystems. All subsystem interconnections including feedback loops are readily seen. Furthermore, it takes only 62 storage places to store the relation in the diclique format whereas 188 places are required to store the relation in standard format (this doesn't account for pointer storage); this three times savings is obtained from a digraph of only 24 nodes. Digraphs of 50 or 100 nodes could generally bring much better savings.

*Notation.*

- $X$  is a nonempty set and domain for the binary relation  $R$ .
- $Y$  is a nonempty set and range of the binary relation  $R$ .
- $R$  is a nonempty binary relation from  $X$  to  $Y$ ;  
 $R \subseteq X \times Y$ .
- $I$  is a subset of  $X$ .

- $O$  is a subset of  $Y$ .
- $x$  is an element of  $X$ .
- $i$  is an element of  $I$ .
- $y$  is an element of  $Y$ .
- $o$  is an element of  $O$ .
- $R(x) = \{y \in Y \mid (x, y) \in R\}$ .
- $R^{-1}(y) = \{x \in X \mid (x, y) \in R\}$ .

*Definition.* Let  $R \subseteq X \times Y$  and  $I \subseteq X$  and  $O \subseteq Y$ . The pair  $(I, O)$  is a block of  $R$  if and only if  $I \times O \subseteq R$ . The set  $I$  of the block is called the input set of the block and the set  $O$  is called the output set of the block.

*Definition.* Let  $R \subseteq X \times Y$ ,  $I \subseteq X$ , and  $O \subseteq Y$ . The pair  $(I, O)$  is a diclique of  $R$  if and only if (1)  $(I, O)$  is a block of  $R$ , and (2)  $I' \supseteq I$  and  $O' \supseteq O$  and  $(I', O')$  a block of  $R$  implies  $(I', O') = (I, O)$ . In other words, dicliques are maximal blocks.

Our development will be in three parts. In the first part we prove some lemmas leading to the Diclique Representation Theorem, which by its constructive nature indicates

how to determine the dicliques of any binary relation. In the second part we prove some more properties about dicliques which indicate that the set of all dicliques of a given binary relation is characterized by being an idempotent commutative monoid. This algebraic structure provides a quick way via substitution property partitions [2, 3] of homomorphisms to coalesce dicliques and we present an algorithm for doing so. In part three we indicate a relationship between clique and diclique and show how to obtain cliques from dicliques.

## 2. Diclique Representation Theorem

Our procedure will be to prove three lemmas to establish some basic facts about binary relations which are needed to establish the Diclique Representation Theorem. The first lemma states that if we start with a given subset  $A$  in the domain of the relation  $R$  and travel through the relation  $R$  to that subset  $O$ , in the range of the relation, each of whose elements are reachable by each element in  $A$ , then when we travel back through the relation  $R^{-1}$  to that subset  $I$ , in the domain of the relation, each of whose elements are reachable by each element in  $O$ , we necessarily find that  $I$  must contain  $A$ .

Lemma 2 states that if we now travel from the subset  $I$  through the relation  $R$  to that subset  $O'$ , in the range of the relation, each of whose elements are reachable by each element in  $I$ , then we necessarily find that  $O' = O$ . In other words, in any one round trip through the relation we cannot lose any members and might in fact pick up some; but trips additional to the first round trip cannot pick up or lose any members.

Lemma 3 states that the maximality property of the dicliques implies that if  $I$  is the input set to some diclique of a relation  $R$ , then one trip through the relation  $R$  will land us at the output set  $O$  of the diclique.

The Diclique Representation Theorem states that if we start with a given subset  $A$  in the domain of the relation  $R$  and travel through the relation  $R$  to that subset  $O$ , each of whose members are reachable from each member of  $A$ , and then round trip it back through  $R^{-1}$  from  $O$  to that subset  $I$ , each of whose members are reachable from each member of  $O$ , then since trips beyond the first round trip do not add any members,  $(I, O)$  must be a diclique of  $R$ ; and the converse is true.

The Diclique Representation Theorem immediately leads to the algorithm for finding dicliques of a binary relation  $R \subseteq X \times Y$ .

### DICLIQUE ALGORITHM

Step 1. First form the set  $T$  of distinct subsets  $R(x)$ ,  $x \in X$ . This set will be the generating set for the diclique output sets.

Step 2. Enlarge this set by generating all possible intersections between output sets of  $T$ .

Step 3. If the set  $T$  has increased in size go to step 2, else go to step 4.

Step 4. Corresponding to each output set  $O$  in  $T$ , compute its corresponding input set  $I$  by  $I = \bigcap_{o \in O} R^{-1}(o)$ . Each pair  $(I, O)$  is then a diclique of  $R$ .

LEMMA 1. Let  $R \subseteq X \times Y$ ,  $A \subseteq X$ ,  $O = \bigcap_{i \in A} R(i)$ , and  $I = \bigcap_{o \in O} R^{-1}(o)$ . Then  $A \subseteq I$ .

PROOF. Let  $i \in A$ . Then for every  $o \in O$ ,  $(i, o) \in R$ , since  $O = \bigcap_{i \in A} R(i)$ . So  $i \in R^{-1}(o)$  for every  $o \in O$  or  $i \in \bigcap_{o \in O} R^{-1}(o) = I$ .

COROLLARY (Dual). Let  $R \subseteq X \times Y$ ,  $B \subseteq Y$ ,  $I = \bigcap_{o \in B} R^{-1}(o)$ ,  $O = \bigcap_{i \in I} R(i)$ . Then  $B \subseteq O$ .

LEMMA 2. Let  $R \subseteq X \times Y$ ,  $A \subseteq X$ ,  $O = \bigcap_{i \in A} R(i)$ ,  $I = \bigcap_{o \in O} R^{-1}(o)$ , and  $O' = \bigcap_{i \in I} R(i)$ . Then  $O' = O$ .

PROOF. By Lemma 1,  $A \subseteq I'$ . By the corollary to Lemma 1,  $O \subseteq O'$ . But  $O' = \bigcap_{i \in I'} R(i) \subseteq \bigcap_{i \in A} R(i) = O$ , since  $A \subseteq I'$ . Therefore,  $O' = O$ .

COROLLARY (Dual). Let  $R \subseteq X \times Y$ ,  $B \subseteq Y$ ,  $I = \bigcap_{o \in B} R^{-1}(o)$ ,  $O = \bigcap_{i \in I} R(i)$ , and  $I' = \bigcap_{o \in O} R^{-1}(o)$ . Then  $I' = I$ .

LEMMA 3. Let  $R \subseteq X \times Y$  and  $(I, O)$  be a dichique of  $R$ . Then  $O = \bigcap_{i \in I} R(i)$ .

PROOF. Let  $o \in O$ . Since  $(I, O)$  is a dichique of  $R$ ,  $I \times O \subseteq R$ . So for every  $i \in I$ ,  $o \in R(i)$ . Hence,  $o \in \bigcap_{i \in I} R(i)$  and  $O \subseteq \bigcap_{i \in I} R(i)$ . Next note that  $I \times \bigcap_{i \in I} R(i) \subseteq R$ . Let  $(x, y) \in I \times \bigcap_{i \in I} R(i)$ . Then for every  $i \in I$ ,  $y \in R(i)$ . Since  $x \in I$  we have  $y \in R(x)$  or  $(x, y) \in R$ . Finally,  $I \supseteq I$ ,  $\bigcap_{i \in I} R(i) \supseteq O$ ,  $I \times \bigcap_{i \in I} R(i) \subseteq R$ , and  $(I, O)$  a dichique implies  $O = \bigcap_{i \in I} R(i)$ .

COROLLARY (Dual). Let  $R \subseteq X \times Y$  and  $(I, O)$  be a dichique of  $R$ . Then  $I = \bigcap_{o \in O} R^{-1}(o)$ .

DICHIQUE REPRESENTATION THEOREM. Let  $R \subseteq X \times Y$ ,  $I \subseteq X$ , and  $O \subseteq Y$ . Then  $(I, O)$  is a dichique of  $R$  if and only if for some  $A$ ,  $A \subseteq X$ ,  $O = \bigcap_{i \in A} R(i)$ , and  $I = \bigcap_{o \in O} R^{-1}(o)$ .

PROOF. Suppose for some set  $A \subseteq X$ ,  $O = \bigcap_{i \in A} R(i)$  and  $I = \bigcap_{o \in O} R^{-1}(o)$ . First we show that  $(I, O)$  is a block of  $R$ . Let  $(i, o) \in I \times O$ . By Lemma 2,  $O = \bigcap_{i \in I} R(i)$ . Hence, for every  $i \in I$  and for every  $o \in O$ ,  $(i, o) \in R$ .

Second we show that  $(I, O)$  is maximal. Suppose  $I' \supseteq I$  and  $O' \supseteq O$  and  $I' \times O' \subseteq R$ . Since  $I = \bigcap_{o \in O} R^{-1}(o)$ , there exists no  $x \in X \cap I^c$  such that  $(x, o) \in R$  for every  $o \in O$ . We already know that  $I \times O \subseteq R$ . And nothing in addition to those elements in  $I$  will go through  $R$  to all those elements of  $O$ . Therefore,  $I \times O \subseteq I' \times O \subseteq I' \times O' \subseteq R$  must imply  $I = I'$ . Since  $O = \bigcap_{i \in I} R(i)$ , there exists no  $y \in Y \cap O^c$  such that  $(i, y) \in R$  for every  $i \in I$ . We already know that  $I \times O \subseteq R$  and nothing in addition to those elements in  $O$  will go through  $R^{-1}$  to all those elements of  $I$ . Therefore,  $I \times O \subseteq I \times O' \subseteq R$  must imply  $O = O'$ . Because  $(I, O)$  is a maximal block of  $R$ , it is by definition a dichique of  $R$ .

Suppose  $(I, O)$  is a dichique of  $R$ . We are looking for some set  $A$  such that

$$O = \bigcap_{i \in A} R(i) \quad \text{and} \quad I = \bigcap_{o \in O} R^{-1}(o).$$

By Lemma 3 and its dual corollary,  $O = \bigcap_{i \in I} R(i)$  and  $I = \bigcap_{o \in O} R^{-1}(o)$ . We have found such a set:  $I$  itself.

COROLLARY (Dual). Let  $R \subseteq X \times Y$ ,  $I \subseteq X$ , and  $O \subseteq Y$ . Then  $(I, O)$  is a dichique of  $R$  if and only if for some set  $B$ ,  $B \subseteq Y$ ,  $I = \bigcap_{o \in B} R^{-1}(o)$ , and  $O = \bigcap_{i \in I} R(i)$ .

### 3. Coalescing of Dichiques

The definition of the binary relation  $R$ , whose dichiques we wish to find, usually involves some judgment, subjectivity, and error on the part of the definer when the definition depends on experimental data. In other words, there are probably some pairs of variables which do not belong to  $R$  which should, an error of omission, and others which do belong to  $R$  which should not, an error of commission. In the process of system synthesis and modeling, those variables involved in errors of omission cause serious problems, for it means that certain important relationships will not be taken into account. On the other hand, those variables involved in errors of commission cause little problem except to make the system synthesis a little more difficult. Furthermore, sensitivity analysis of the model will indicate to us later that large changes in these variables have little effect on the values of those variables they are supposed to control. Hence, we will discover them after the model is set up. So our problem at the time we wish to define subsystems is to make reasonably sure that all the necessary input and output variables are included in each subsystem and that the subsystems are simply described and reasonably sized pieces of the system.

One possible way to assume that all variable corrections which should be in the relation are in and the subsystems are reasonably sized is to combine or coalesce dichiques having corresponding input and output sets of variables which have a large proportion of members in common. For example if  $(\{1, 3, 5\}, \{2, 4, 8\})$  and  $(\{3, 5, 6\}, \{4, 7, 8\})$  are two dichiques of some binary relation we might want to combine them since they have 2 out of 3 members of the input sets matching and 2 out of 3 members of the output sets

matching. The combinations would be the coalesced pair  $(\{1, 3, 5, 6\}, \{2, 4, 7, 8\})$ . This coalescing has the effect of adding the two pairs  $(1, 7)$ ,  $(6, 2)$  to the binary relation since  $(\{1, 3, 5\} \times \{2, 4, 8\}) \cup (\{3, 5, 6\} \times \{4, 7, 8\}) \cup \{(1, 7), (6, 2)\} = \{1, 3, 5, 6\} \times \{2, 4, 7, 8\}$ .

Unfortunately coalescing of dicliques cannot be done blindly or arbitrarily. For if we indeed add elements to the binary relation, it is possible that dicliques other than those we coalesced will be affected. Since we do not want to recompute the set of dicliques of the expanded relation, we must determine the effect of the members added to the relation on all its dicliques. Since the members we wish to add to the relation are obtained by coalescing dicliques and since coalescing coarsens the initial identity partition on the set of dicliques and since, as we show in this section, the set of dicliques has a nice algebraic structure, it is possible to trace the effect of the coalescing from the dicliques themselves. In essence, we coalesce the dicliques we initially wanted to coalesce and we then coalesce all other pairs of dicliques which are affected by the initial coalescing and we keep this up until there are no more effects. The coarsened partition produced by this procedure is called a partition with the substitution property [2, 3].

Lemma 4 establishes that when the input sets to dicliques are intersected, the result is also an input set to a diclique. Hence the input sets to the dicliques of a relation form a lattice under the intersection operation [4]. Lemma 5 takes the appropriately defined intersection of dicliques as a binary operation on the set of dicliques of a given binary relation and proves that the set of dicliques with the intersection operation forms not only a groupoid but an idempotent commutative monoid.

Lemmas 6 and 7 review some basic relationships between homomorphisms, homomorphic images, and substitution property partitions of groupoids. Lemmas 8 and 9 prove that substitution property partitions are characterized by a congruence relation. Use of this characterization is made by the algorithm for coalescing.

#### COALESCING ALGORITHM

To find the largest homomorphic image of a commutative groupoid having elements  $a$  and  $b$  coalesced:

- Step 1. Set up a partition in which each cell has one member except for the cell containing  $a$  and  $b$ .
- Step 2. In the multiplication table, replace all references of " $b$ " by " $a$ ."
- Step 3. If there exist two columns labeled the same go to step 4; otherwise stop.
- Step 4. If for any row of these same labeled columns, the pair of entries is not in the same cell, coalesce the cells and in the multiplication table replace all references of the second cell by references to the first cells. Continue doing this until one pair of these same labeled columns are identical.
- Step 5. Delete one column of the pairs of identical same labeled columns. Delete the corresponding row and go to step 3.

LEMMA 4. Let  $R \subseteq X \times Y$  and  $(I_1, O_1)$  and  $(I_2, O_2)$  be dicliques of  $R$ . Then  $(I_1 \cap I_2, \bigcap_{x \in I_1 \cap I_2} R(x))$  is a diclique of  $R$ .

PROOF. By the corollary dual to the Diclique Representation Theorem,  $B \subseteq Y$  implies  $(\bigcap_{b \in B} R^{-1}(b), \bigcap_{x \in \bigcap_{b \in B} R^{-1}(b)} R(x))$  is a diclique of  $R$ . Take  $B = O_1 \cup O_2$ . Then  $\bigcap_{b \in B} R^{-1}(b) = \bigcap_{y \in O_1 \cup O_2} R^{-1}(y) = [\bigcap_{y \in O_1} R^{-1}(y)] \cap [\bigcap_{y \in O_2} R^{-1}(y)]$ . But  $(I_1, O_1)$  and  $(I_2, O_2)$  are dicliques of  $R$  and by Lemma 3 we know that  $I_1 = \bigcap_{y \in O_1} R^{-1}(y)$  and  $I_2 = \bigcap_{y \in O_2} R^{-1}(y)$ . Hence,  $\bigcap_{b \in B} R^{-1}(b) = I_1 \cap I_2$ .

COROLLARY (Dual). Let  $R \subseteq X \times Y$  and  $(I_1, O_1)$  and  $(I_2, O_2)$  be dicliques of  $R$ . Then  $(\bigcap_{y \in O_1 \cap O_2} R^{-1}(y), O_1 \cap O_2)$  is a diclique of  $R$ .

LEMMA 5. Let  $R \subseteq X \times Y$  and  $\mathfrak{D} = \{(I, O) \mid (I, O) \text{ is a diclique of } R\}$ . Define the operation  $\cdot$  by  $(I_1, O_1) \cdot (I_2, O_2) = (I_1 \cap I_2, \bigcap_{x \in I_1 \cap I_2} R(x))$ . Then  $(\mathfrak{D}, \cdot)$  is a commutative idempotent monoid.

PROOF. (1) By Lemma 4,  $(I_1, O_1)$  and  $(I_2, O_2) \in \mathfrak{D}$  imply

$$(I_1 \cap I_2, \bigcap_{x \in I_1 \cap I_2} R(x)) \in \mathfrak{D}.$$

Hence the operation  $\cdot$  is closed and  $(\mathfrak{D}, \cdot)$  is a groupoid.

(2) Suppose  $(I_1, O_1), (I_2, O_2),$  and  $(I_3, O_3) \in \mathfrak{D}$ . Then

$$\begin{aligned} ((I_1, O_1) \cdot (I_2, O_2)) \cdot (I_3, O_3) &= (I_1 \cap I_2, \bigcap_{x \in I_1 \cap I_2} R(x)) \cdot (I_3, O_3) \\ &= ((I_1 \cap I_2) \cap I_3, \bigcap_{x \in (I_1 \cap I_2) \cap I_3} R(x)) \\ &= (I_1 \cap (I_2 \cap I_3), \bigcap_{x \in I_1 \cap (I_2 \cap I_3)} R(x)) \\ &= (I_1, O_1) \cdot (I_2 \cap I_3, \bigcap_{x \in I_2 \cap I_3} R(x)) \\ &= (I_1, O_1) \cdot ((I_2, O_2) \cdot (I_3, O_3)). \end{aligned}$$

Hence the operation is associative and  $(\mathfrak{D}, \cdot)$  is a semigroup.

(3) Note that  $(X, \bigcap_{x \in X} R(x)) \in \mathfrak{D}$ . Since by Lemma 1  $X \subseteq \bigcap_{y \in \bigcap_{x \in XR(x)} R^{-1}(y)}$  and since  $R \subseteq X \times Y, R^{-1}(y) \subseteq X$  for every  $y \in Y$ . Thus  $\bigcap_{y \in \bigcap_{x \in XR(x)} R^{-1}(y)} \subseteq X$  and  $X = \bigcap_{y \in \bigcap_{x \in XR(x)} R^{-1}(y)}$ . By the representation theorem,  $(\bigcap_{y \in \bigcap_{x \in XR(x)} R^{-1}(y)}, \bigcap_{x \in X} R(x))$  is a diclique of  $R$ . But  $\bigcap_{y \in \bigcap_{x \in XR(x)} R^{-1}(y)} = X$  implies  $(X, \bigcap_{x \in X} R(x))$  is a diclique of  $R$ . Now suppose  $(I, O)$  is also a diclique of  $R$ . Consider  $(I, O) \cdot (X, \bigcap_{x \in X} R(x)) = (I \cap X, \bigcap_{x \in I \cap X} R(x)) = (I, \bigcap_{x \in I} R(x)) = (I, O)$ . Hence  $(X, \bigcap_{x \in X} R(x))$  is an identity and  $(\mathfrak{D}, \cdot)$  is a monoid.

(4) Suppose  $(I, O)$  is a diclique of  $R$ . Consider

$$(I, O) \cdot (I, O) = (I \cap I, \bigcap_{x \in I \cap I} R(x)) = (I, \bigcap_{x \in I} R(x)) = (I, O).$$

Hence,  $(\mathfrak{D}, \cdot)$  is idempotent.

(5) Suppose  $(I_1, O_1)$  and  $(I_2, O_2)$  are dicliques of  $R$ . Consider  $(I_1, O_1) \cdot (I_2, O_2) = (I_1 \cap I_2, \bigcap_{x \in I_1 \cap I_2} R(x)) = (I_2 \cap I_1, \bigcap_{x \in I_2 \cap I_1} R(x)) = (I_2, O_2) \cdot (I_1, O_1)$ . Hence,  $(\mathfrak{D}, \cdot)$  is commutative.

*Definition.* Let  $(G, \cdot)$  be a groupoid and  $\pi = \{P_k\}_{k=1}^K$  be a partition over  $G$  having the property for every  $x \in P_i$  and for every  $y \in P_j, xy \in P_k$ , for some  $P_k \in \pi, P_k$  depending only on  $i$  and  $j$ . Then  $\pi$  is called a substitution property partition of  $G$ .

*Definition.* Let  $(G, \cdot)$  and  $(H, \cdot)$  be groupoids and  $f$  be a function,  $f: G \rightarrow H$ , having the property for every  $x, y \in G, f(xy) = f(x)f(y)$ , then  $f$  is a homomorphism.

LEMMA 6. Let  $(G, \cdot)$  and  $(H, \cdot)$  be groupoids and  $f: G \rightarrow H$  be a homomorphism. Then  $\pi = \{f^{-1}(h) \mid h \in H\}$  is a substitution property partition.

PROOF. Fix  $a, b \in H$  and let  $c = ab$ . Choose  $d \in f^{-1}(a), e \in f^{-1}(b)$ . We will show  $de \in f^{-1}(c)$ . Since  $d \in f^{-1}(a), f(d) = a$ . Since  $e \in f^{-1}(b), f(e) = b$ . Then  $c = ab = f(d)f(e) = f(de)$  and  $de \in f^{-1}(c)$ .

LEMMA 7. Let  $(G, \cdot)$  be a groupoid and  $\pi = \{P_k\}_{k=1}^k$  be a substitution property partition of  $G$ . Then  $f: G \rightarrow \pi$  is a homomorphism from  $G$  to  $\pi$  where  $f$  is defined  $f(g) = P_i$  where  $g \in P_i$ .

PROOF. First we note that  $(\pi, \odot)$  is a groupoid. The multiplication  $\odot$  is defined by  $P_i \odot P_j = P_k$  where for some  $x \in P_i$  and some  $y \in P_j, x \cdot y \in P_k$ . The multiplication is well defined since for every  $x \in P_i$ , and for every  $y \in P_j, xy \in P_k$  by definition of  $\pi$ .

Now we show  $f$  is a homomorphism. Let  $a, b \in G$  and  $c = ab$ . Let  $P_i = f(a), P_j = f(b)$ , and  $P_k = f(c)$ . Then  $P_k = f(c) = f(ab)$ . Consider  $f(a)f(b) = P_i \odot P_j$ . Since  $a \in P_i$  and  $b \in P_j, ab = c \in P_k$  by definition of  $\odot, P_i \odot P_j = P_k$ . Hence,  $f(ab) = f(c) = f(a) \odot f(b)$  and  $f$  is a homomorphism.

LEMMA 8. Let  $(G, \cdot)$  be a groupoid and  $\pi$  be a partition over  $G$  having the property  $x, y \in P \in \pi$  implies for every  $z \in G$ , there exists  $Q \in \pi$  such that  $xz, yz \in Q$ , and there exists  $Q' \in \pi$  such that  $xz, zy \in Q'$ . Then  $\pi$  is a substitution property partition.

PROOF. Let  $x \in P_i, y \in P_j, z = xy \in P_k$ . Let  $x' \in P_i$  and  $y' \in P_i$  and  $z' = x'y'$ . We show that  $z' \in P_k$ . Since  $x, x' \in P_i$  we must have  $xy$  and  $x'y$  in the same cell. But  $xy \in P_k$  implies  $x'y \in P_k$ . Since  $y, y' \in P_j$ , we must have  $x'y$  and  $x'y'$  in the same cell. But  $x'y \in P_k$  implies  $x'y' \in P_k$ . Hence  $z' \in P_k$ .

LEMMA 9. Let  $(G, \cdot)$  be a groupoid and  $\pi$  be a substitution property partition over  $G$ . Then  $x, y \in P$  implies for every  $z \in G$  there exists  $Q \in \pi$  such that  $xz, yz \in Q$ , and there exists  $Q' \in \pi$  such that  $xz, zy \in Q'$ .

PROOF. Let  $x, y \in P_i$  and  $z \in G$ . Since  $\pi$  covers  $G$ , there exists a cell  $P_j$  such that  $z \in P_j$ . Since  $\pi$  is a substitution property partition, for every  $a \in P_i$ , for every  $b \in P_j$ , there exist  $P_k$  such that  $ab \in P_k$  where  $P_k = P_i \odot P_j$ . Hence, if  $xz \in P_k$ , then  $yz \in P_k$ ; if  $zx \in P_k$ , then  $zy \in P_k$ .

Let  $\pi_0$  be a partition on a finite groupoid  $G$ . We define an iterative procedure which generates a finite length sequence of partitions  $\pi_1, \pi_2, \dots, \pi_N$  such that  $\pi_N$  is the substitution property partition of  $G$ , having the largest number of cells, and of which  $\pi_0$  is a refinement.

Suppose  $\pi_n$  has already been determined. Either  $\pi_n$  is a substitution property partition or not. If not, then by Lemma 8, there must exist some cell  $P$  of  $\pi_n$  and elements  $x, y \in P$  such that for some  $z \in G, Q_1, Q_2 \in \pi_n$ , either  $xz \in Q_1$  and  $yz \in Q_2$  and  $Q_1 \neq Q_2$  or  $zx \in Q_1$  and  $zy \in Q_2$  and  $Q_1 \neq Q_2$ .

Define a coarsened partition  $\pi_{n+1}$  by  $\pi_{n+1} = \{P \mid (P \in \pi_n \text{ and } P \neq Q_1 \text{ and } P \neq Q_2) \text{ or } P = Q_1 \cup Q_2\}$ .

Since at each iteration step the number of cells of the current partition decreases by one, the procedure must come to a stop in a finite number of steps. Since the final partition  $\pi_N$  satisfies the property that  $x, y \in P$  and  $z \in G$  implies  $xz$  and  $yz$  are in the same class and  $zx$  and  $zy$  are in the same class,  $\pi_N$  is a substitution property partition. Since elements were coalesced or put into the same cell only by necessity at each step,  $\pi_N$  is that substitution property partition having the largest number of cells. Since each cell of the partition  $\pi_N$  is the union of some of the cells of  $\pi_0$ ,  $\pi_0$  is a refinement of  $\pi_N$ .

We illustrate determining the coalescing of dicles with the digraph of Figure 2. The binary relation corresponding to the digraph is given by the table of Figure 2. To obtain the dicles we form a list of the distinct sets  $R(x): \{2, 6, 9\}, \{10, 11\}, \{6, 9\}, \{2, 6\}, \{10, 12\}, \{10, 11, 12\}, \{1, 3\}$ . Then we extend this list adding all possible non-trivial intersections of combinations of these distinct sets. This extended list then represents all possible sets of the form  $\bigcap_{x \in A} R(x)$  for some  $A \subseteq X$ . These sets represent the possible output sets for dicles and as illustrated below the input set corresponding to each output set is easily obtained from  $\bigcap_{y \in \bigcap_{x \in A} R(x)} R^{-1}(y)$ ,

All possible output sets $\bigcap_{x \in A} R(x)$	Corresponding input sets $\bigcap_{y \in \bigcap_{x \in A} R(x)} R^{-1}(y)$	Diclique label
2, 6, 9	1, 3, 7	A
10, 11	2, 9	B
6, 9	1, 3, 4, 7	C
2, 6	1, 3, 5, 7	D
10, 12	6, 8, 9	E
10, 11, 12	9	F
1, 3	10, 11, 12	G
10	2, 6, 9, 8	H
6	1, 3, 4, 5, 7	I

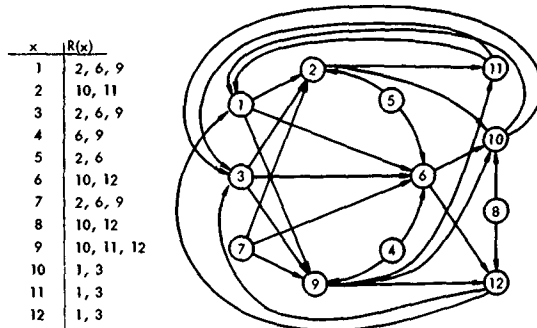


FIG. 2. Illustration of a digraph and corresponding relation having nontrivial dicles ( $\{1, 3, 7\}, \{2, 6, 9\}$ ), ( $\{2, 9\}, \{10, 11\}$ ), ( $\{1, 3, 4, 7\}, \{6, 9\}$ ), ( $\{1, 3, 5, 7\}, \{2, 6\}$ ), ( $\{6, 8, 9\}, \{10, 12\}$ ), ( $\{9\}, \{10, 11, 12\}$ ), ( $\{10, 11, 12\}, \{1, 3\}$ ), ( $\{2, 6, 9, 8\}, \{10\}$ ), and ( $\{1, 3, 4, 5, 7\}, \{6\}$ )





TABLE V. THE RESULT AFTER ALL REFERENCES TO H IN TABLE IV ARE REPLACED BY E AND ONE OF THE SAME LABELED E IDENTICAL COLUMNS AND ROWS IS DELETED

	A	B	E	F	G	X	∅
A	A	∅	∅	∅	∅	A	∅
B	∅	B	E	B	∅	B	∅
E	∅	E	E	E	∅	E	∅
F	∅	B	E	F	∅	F	∅
G	∅	∅	∅	∅	G	G	∅
X	A	B	E	F	G	X	∅
∅	∅	∅	∅	∅	∅	∅	∅

TABLE VI. THE RESULT AFTER E AND F OF TABLE V ARE COALESCED

	A	B	G	X	∅
A	A	∅	∅	A	∅
B	∅	B	∅	B	∅
G	∅	∅	G	G	∅
X	A	B	G	X	∅
∅	∅	∅	∅	∅	∅

sets, we coalesce them. Replace all references to H by E in Table IV. Since the two columns labeled E are identical we may get rid of one and obtain the collapsed Table V.

Dicliques E and F have a fair number of elements in common. Putting these two together forces them both to coalesce with B and we obtain Table VI. Finally we summarize what the final nontrivial dicliques are:

$$\begin{aligned}
 A &= (\{1, 3, 7\}, \{2, 6, 9\}) \cup (\{1, 3, 4, 7\}, \{6, 9\}) \cup (\{1, 3, 5, 7\}, \{2, 6\}) \\
 &\quad \cup (\{1, 3, 4, 5, 7\}, \{6\}) = (\{1, 3, 4, 5, 7\}, \{2, 6, 9\}), \\
 B &= (\{2, 9\}, \{10, 11\}) \cup (\{6, 8, 9\}, \{10, 12\}) \cup (\{9\}, \{10, 11, 12\}) \\
 &\quad \cup (\{2, 6, 8, 9\}, \{10\}) = (\{2, 6, 8, 9\}, \{10, 11, 12\}), \\
 G &= (\{10, 11, 12\}, \{1, 3\}).
 \end{aligned}$$

The digraph can now be drawn in diclique or subsystem form as illustrated in Figure 3. Compare how much more easily the subsystem diagram of Figure 3 is understood than the digraph of Figure 2. Every connection on Figure 2 is on Figure 3. The only difference is that Figure 3 has 5 additional connections: (4, 2), (5, 9), (2, 12), (6, 11), (8, 11).

4. *Dicliques and Cliques*

As the names would imply, diclique and clique are similar concepts: diclique is to digraph what clique is to graph. This similarity would lead us to believe that there is a close relationship between the two structures. In this section we show how the cliques of a relation can be quickly obtained from the dicliques of the relation.

*Definition.* Let  $R \subseteq X \times Y$  and  $C \subseteq X, C' \subseteq Y$ . The subset  $C$  is called a clique of  $R$  if and only if (1)  $C \times C \subseteq R$ , and (2)  $C' \supset C$  and  $C' \times C' \subseteq R$  imply  $C' = C$ .

Lemma 10 states that corresponding to each clique  $C$ , there exists a diclique  $(I, O)$  such that  $C = I \cap O$ . So if we are looking for the cliques of a relation  $R'$ , we determine the symmetric core  $R$  of  $R'$  by  $R = R' \cap R'^{-1}$ . (This serves only to reduce computations and is not really necessary.) Then we find the dicliques of  $R$  and take the intersection of the input and output sets of each diclique. By the lemma, we know that among the sets in the collection of diclique input sets intersected with their corresponding output sets each and every clique is represented. If we eliminate those sets in the collection which are subsets of other sets in the collection, the remaining sets will constitute the entire collection of cliques of the relation  $R$ .

The following example illustrates how cliques can be obtained from the set of cliques. Consider the graph of Figure 4. Clearly there are 4 cliques: {1, 2, 3, 4}, {1, 2, 7}, {2, 4, 5}, {3, 4, 6}. The binary relation corresponding to the graph is given by the table in

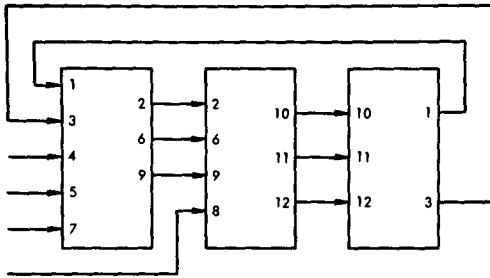
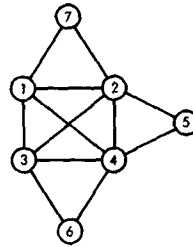


FIG 3. Illustration of the subsystem diagram corresponding to the coalesced dichiques of the binary relation of Figure 2



x	R(x)
1	1 2 3 4 7
2	1 2 3 4 5 7
3	1 2 3 4 6
4	1 2 3 4 5 6
5	2 4 5
6	3 4 6
7	1 2 7

FIG. 4. Illustration of a graph and its corresponding relation having cliques {1, 2, 3, 4}, {1, 2, 7}, {2, 4, 5}, and {3, 4, 6}. Although not drawn on the graph, we assume each node is connected to itself

Figure 4. First we form the list of the distinct sets  $R(x)$ : {1, 2, 3, 4, 7}, {1, 2, 3, 4, 5, 7}, {1, 2, 3, 4, 6}, {1, 2, 3, 4, 5, 6}, {2, 4, 5}, {3, 4, 6}, {1, 2, 7}. Then we extend this list adding all possible nontrivial intersections. This extended list represents all possible output sets of the form  $\bigcap_{x \in A} R(x)$  for some  $A \subseteq X$ . We obtain the corresponding input sets, then intersect the output sets with the input sets and note that the cliques are the distinct maximal intersections.

All possible output sets $\bigcap_{x \in A} R(x)$	Corresponding input sets $\bigcap_{y \in \bigcap_{x \in A} R(x)} R^{-1}(y)$	Intersection of output and input sets $\bigcap_{x \in A} R(x)$ and $\bigcap_{y \in \bigcap_{x \in A} R(x)} R^{-1}(y)$	Distinct maximal intersection (cliques)
1, 2, 3, 4, 7	1, 2	1, 2	
1, 2, 3, 4, 5, 7	2	2	
1, 2, 3, 4, 6	3, 4	3, 4	
1, 2, 3, 4, 5, 6	4	4	
2, 4, 5	2, 4, 5	2, 4, 5	2, 4, 5
3, 4, 6	3, 4, 6	3, 4, 6	3, 4, 6
1, 2, 7	1, 2, 7	1, 2, 7	
1, 2, 3, 4	1, 2, 3, 4	1, 2, 3, 4	1, 2, 3, 4
1, 2, 3, 4, 5	2, 4	2, 4	
4	1, 2, 3, 4, 5, 6	4	
2	1, 2, 3, 4, 5, 7	2	
2, 4	1, 2, 3, 4, 5	2, 4	
3, 4	1, 2, 3, 4, 6	3, 4	
1, 2	1, 2, 3, 4, 7	1, 2	

LEMMA 10. Let  $R \subseteq X \times Y$  be a binary relation and  $C$  be a clique of  $R$ . Then there exists some dichique  $(I, O)$  of  $R$  such that  $C = I \cap O$ .

PROOF. Our proof is constructive. Take  $I = \bigcap_{x \in \bigcap_{y \in C} R^{-1}(y)} R(x)$  and  $O = \bigcap_{x \in C} R(x)$ . By the dichique representation theorem,  $(I, O)$  is a dichique of  $R$ .

We now show that  $C = I \cap O$ . Since  $C$  is a clique of  $R$ ,  $(C, C)$  is certainly a block of  $R$ .  $(C, C)$  a block of  $R$  implies  $C \subseteq \bigcap_{y \in C} R^{-1}(y)$  and this in turn implies that  $\bigcap_{x \in C} R(x) \supseteq \bigcap_{x \in \bigcap_{y \in C} R^{-1}(y)} R(x)$ . By Lemma 1,  $C \subseteq \bigcap_{x \in \bigcap_{y \in C} R^{-1}(y)} R(x)$ . Putting these two relationships together we have  $C \subseteq \bigcap_{x \in \bigcap_{y \in C} R^{-1}(y)} R(x) \subseteq \bigcap_{x \in C} R(x)$ . Hence,  $C \subseteq I \subseteq O$ , which implies  $C \subseteq I \cap O$ .

Now note that  $(I \cap O) \times (I \cap O) \subseteq R$  since  $(I \cap O) \times (I \cap O) \subseteq I \times O$  and  $(I, O)$  by virtue of being a dichique is a block of  $R$ . The maximality of  $C$  asserts that if  $I \cap O \supseteq C$  and  $(I \cap O) \times (I \cap O) \subseteq R$ , then  $C = I \cap O$ .

### 5. Conclusions

We have discussed the usefulness of the dichique concept and an algorithm for determining dichiques. We have briefly examined the relationship between clique and dichique

and have described an algorithm for determining cliques from the set of dicliques. There are still some basic questions to be answered: (1) When determining which dicliques ought to be coalesced, what criteria should be used? Why? (2) How does the clique finding algorithm suggested in Section 3 compare with others such as Bierstone's Algorithm for generating cliques [5, 6]?

ACKNOWLEDGMENTS. I would like to acknowledge the helpful discussions with Professors Creese, Adams, and Bavel as well as the comments made by the referee.

#### REFERENCES

1. HARARY, FRANK. *Graph Theory*. Addison-Wesley, Reading, Mass., 1969, p. 17.
2. HARTMANIS, J. On the state assignment problem for sequential machines I. *IRE Trans. EC-10* (1961), 157-165.
3. HARTMANIS, J. On the state assignment problem for sequential machines II. *IRE Trans. EC-10* (1961), 593-603.
4. MACLANE, S., AND BIRKHOFF, G. *Algebra*. Macmillan Company, London, 1967.
5. AUGUSTSON, J. G., AND MINKE, J. An analysis of some graph theoretical cluster techniques. *J. ACM* 17, 4 (Oct. 1970), 571-588.
6. MULLIGAN, G. D., AND CORNEIL, D. G. Corrections to Bierstone's algorithm for generating cliques. *J. ACM* 19, 2 (April 1972), 244-247.

RECEIVED JANUARY 1973; REVISED JULY 1973