Dicliques: Finding Needles in Haystacks

Robert M. Haralick
Computer Science, Graduate Center
City University of New York
365 Fifth Avenue
New York, NY 10016

1 Introduction

An important process in counter-terrorism is the information processing of large data bases that have been assembled from a variety of sources such as internet exchanges, email, bank records, telephone records, credit card records, travel records, and observations of many sorts. The bits and pieces of information come from many sources and the pieces do not all tightly connect together. Some (possibly disconnected) pieces tightly connect to some other (possibly disconected) pieces. This is the nature of terrorist cell operations. Those in one cell do not know those in another cell. The person who routes money to the cells does not know or have any dealings with the one who handles and supervises the terrorists. And the one who handles them does not know or have any dealings with the one who provides safe houses or the one who provides explosives etc.

The data bases are very large. The number of records that pertain to any kind of terrorist activity is very small. The problem is to find the needle in the haystack. Yet despite the difficulty, the needle we are looking for involves connections between different types of relevant data. In this paper we discuss what dicliques are and how they can be used to find related records that constitute the needle in the haystack.

2 The Binary Relation

Our model begins with the binary relation. A connection between information piece A and information piece B is labeled. For example IP address A and IP address B "chatted."

$$(A, B, "chatted")$$

Let X be a set. A *binary relation* R is a subset of $X \times X$. Let L be a set of labels. A *labeled binary relation* R is a subset of $X \times X \times L$.

The set X can include:

Names
Addresses
Telephone Numbers
Bank Account Numbers
Bank Names
IP address
Passport Numbers
Places

The labels can include:

Visited
Communicated with
Was seen with
Was at
Traveled to
Telephoned
Transfered money to
Received money from

For an abstract example we consider a set X defined by

$$X = \{a, b, c, d, x, y, z, t, u\}$$

The binary relation R on X is defined by

$$\begin{aligned} R \quad = \quad & \{(a, x), (a, y), (a, z), (b, x), (b, y), (b, z), \\ & (b, t), (b, u), (c, x), (c, y), (d, u), (d, t), (d, y)\} \end{aligned}$$

We can represent R in a shorter list form by

R

	R
a	x, y, z
b	x, y, z, t, u
c	x, y
d	u, t, y

And we can visualize R as a digraph. In the digraph the elements of X are shown as nodes. A pair (i, j) in R is drawn as an arrow going from i to j. This is shown in figure 2. By changing the physical location of the nodes, different drawings of the same relation are possible.

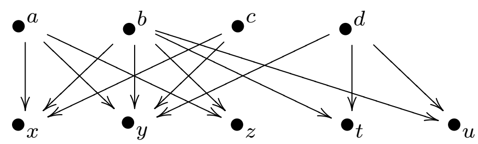

Figure 1: Shows the digraph corresponding to the example relation R.

It is clear from examining the digraph that there are connections but not everything involved in the connections are connected to one another. a, b, c, d have no connections to one another. x, y, z, t, u have no connections to one another.

In order to form a description of what information the relation contains, we can try to reorder the digraph – redraw it as in Figure 2.

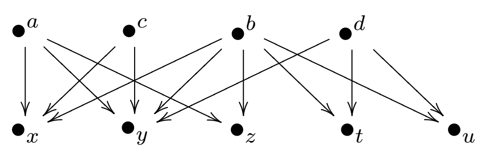

Figure 2: Shows a reordered drawing of the digraph of figure 2.

What we want however is not a technique that is interactive, but one which is systematic and provides a description of the relation in a short and more understandable form. We want a technique that produces structures from the relation where we have in the structure itself maximally relevant information pieces. This structure is called a diclique. The diclique structure and some of its properties were first introduced by Haralick.[1]

3 The Diclique

A pair (P, S) is called a *Diclique* of relation $R \subseteq X \times X$ if and only if

- Containment: $P \times S \subseteq R$
- Maximality: $P' \times S' \subseteq R$ and $P' \supseteq P$ and $S' \supseteq S$ imply $P = P'$ and $S = S'$

The containment condition restricts the diclique description of the relation to only include pairs actually in the relation and not include any pairs not in the relation. The maximality condition means that the diclique structure itself contains maximally relevant information pieces.

The set $\mathcal{D}$ of dicliques are of the example relation R are:

$$\begin{aligned} D_1 &= (\{b\}, \{x, y, z, t, u\}) \\ D_2 &= (\{a, b\}, \{x, y, z\}) \\ D_3 &= (\{a, b, c\}, \{x, y\}) \\ D_4 &= (\{b, d\}, \{y, t, u\}) \\ D_5 &= (\{a, b, c, d\}, \{y\} \end{aligned}$$

$$\mathcal{D} = \{D_1, D_2, D_3, D_4, D_5\}$$

The primary or input side of the diclique can be interpreted as cause and the secondary or output side of the diclique can be interpreted as effect. For example consider the diclique D_3 of our example relation.

$$D_3 = (\{a, b, c\}, \{x, y\})$$

In the counter-terrorism application we might have:

a is a funder
b is a handler
c is the explosive provider
x is terrorist 1
y is terrorist 2

But dicliques have as well other interpretations. For example they can be used the define events.

- Dicliques provide a window to events
- Overlapping Dicliques provide different views of the same event.

- When does an event begin?
- When does an event end?
- What does an event include?

Dicliques can be used in document information extraction. Define the relation R by

$$R = \{(word1, word2) \mid \text{word1 stands in relation to word 2}\}$$

The concept of stands in relation can mean: word1 occurs within k words after word 2 in j documents; or word1 is a noun subject, word2 is a noun object, of a given verb in j documents. In these applications a diclique is a word usage pattern or a meaning usage.

3.1 Diclique Cover

Once we have the constructed the set of dicliques of a relation, we may construct from the set of dicliques a diclique cover of the relation. Let R be a binary relation on X and $\mathcal{D}$ be the set of dicliques of R. A ***Diclique Cover*** of R is a subset $\mathcal{C}$ of dicliques of R satisfying

$$\bigcup_{(P,S)\in\mathcal{C}} P \times S = R$$

A Diclique Cover $\mathcal{C}$ of the example relation R is $\mathcal{C} = \{D_2, D_3, D_4\}$ where

$$\begin{aligned} D_2 &= (\{a, b\}, \{x, y, z\}) \\ D_3 &= (\{a, b, c\}, \{x, y\}) \\ D_4 &= (\{b, d\}, \{y, t, u\}) \end{aligned}$$

A diclique cover $\mathcal{C}$ of $R \subseteq X$ can be visualized by a system diagram where the rectangles represent dicliques and the connecting lines represent the elements of the set X. This is shown in figure 3.1.

The system diagram itself has multiple interpretations many of which are centered as a cause and effect interpretation. The primary set of the diclique is the cause and the secondary set of the diclique is the effect. In the system diagram of Figure 3.1, we have the following inferences.

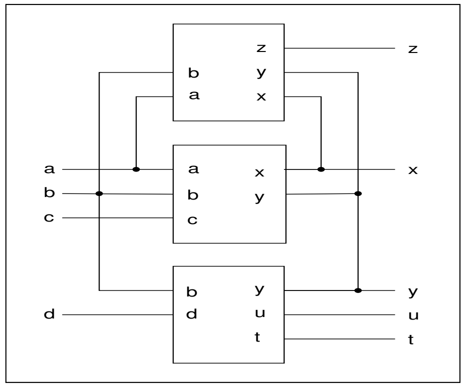

Figure 3: Shows the set of dicliques of a relation represented by a system diagram.

Cause and Effect
b works with a and c
x, y and z are terrorists
a and b work with x, y and z on event 1
a,b and c work with x and y on event
b and d work with x,u and t on event 3
b is a key person

4 Diclique Properties

In this section we summarize the basic mathematical properties of the diclique. These properties are important because they will lead us to an algorithm for computing all the dicliques of a relation. We begin with the diclique intersection theorem.

Theorem 1 *Let $R \subseteq X \times X$. Let (P_1, S_1) and (P_2, S_2) be dicliques of R. Then*

$$(1) \qquad (P_1 \cap P_2, \bigcap_{x \in (P_1 \cap P_2)} R(x))$$

$$(2) \qquad (\bigcap_{y \in (S_1 \cap S_2)} R^{-1}(y), S_1 \cap S_2)$$

are dicliques of R.

The diclique intersection theorem leads us to see that the set of dicliques form an idempotent commutative monoid. Actually there are two idempotent commutative monoids formed. One on the basis of the input or primary sets and one on the basis of the output sets or secondary sets.

Theorem 2 *Let $R \subseteq X \times X$ and let $\mathcal{D}$ be the set of dicliques of R. Let $(P_1, S_1) \in \mathcal{D}$ and $(P_2, S_2) \in \mathcal{D}$. Define the operation $\circ$ on $\mathcal{D}$ by*

$$(P_1, S_1) \circ (P_2, S_2) = (P_1 \cap P_2, \bigcap_{x \in P_1 \cap P_2} R(x))$$

Then $(\mathcal{D}, \circ)$ is an idempotent commutative monoid.

Theorem 3 *Let $R \subseteq X \times X$ and let $\mathcal{D}$ be the set of dicliques of R. Let $(P_1, S_1) \in \mathcal{D}$ and $(P_2, S_2) \in \mathcal{D}$. Define the operation $\circ$ on $\mathcal{D}$ by*

$$(P_1, S_1) \circ (P_2, S_2) = (\bigcap_{y \in S_1 \cap S_2} R^{-1}(y), S_1 \cap S_2)$$

Then $(\mathcal{D}, \circ)$ is an idempotent commutative monoid.

The algebraic structure of the set of dicliques is more than just monoids. The set of dicliques have a partial ordering.

Let (A, B) and (C, D) be dicliques of R. Define the $\leq$ relation by

$$(A, B) \leq (C, D) \text{ if and only if } A \subseteq C$$

Theorem 4 Let $\mathcal{D}$ be the set of all dicliques of R. Then $(\mathcal{D}, \leq)$ is a partially ordered set.

This partial ordering put together with the two idempotent commutative monoids makes the partial ordering a lattice. The set of dicliques forms a lattice.

Theorem 5 $(\mathcal{D}, \leq)$ is a lattice with the meet $\wedge$ and join $\vee$ operators defined by

$$\begin{aligned} (A, B) \wedge (C, D) &= (A \cap C, \bigcap_{x \in A \cap C} R(x)) \\ (A, B) \vee (C, D) &= (\bigcap_{y \in B \cap D} R^{-1}(y), B \cap D) \end{aligned}$$

With these definitions and we see that if $\bigcap_{x \in X} R(x) = \emptyset$, then $(X, \emptyset)$ is a diclique and if $\bigcap_{y \in X} R^{-1}(y) = \emptyset$, then $(\emptyset, X)$ is a diclique.

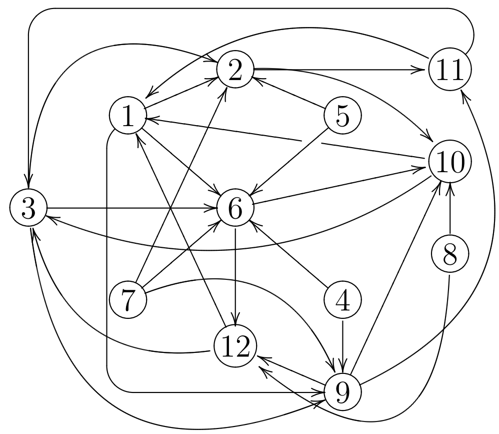

Figure 4: Shows a digraph visualization of a more complicated relation.

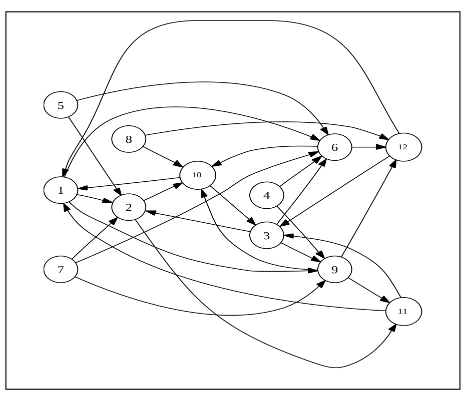

Figure 5: Shows a different digraph visualization of a more complicated relation.

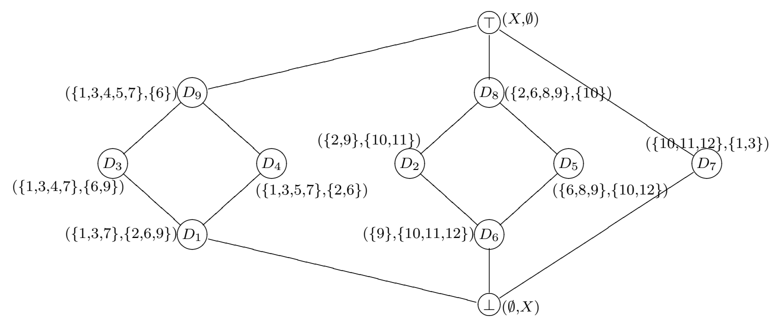

Figure 6: Shows the lattice of the dicliques of the more complicated relation.

5 Computing Dicliques

In this section we summarize some theorems which establish the algorithm for constructing all the dicliques of a relation.

Theorem 6 *If (A,B) is a diclique of R, then*

$$\begin{aligned} A &= \bigcap_{b\in B} R^{-1}(b) \\ B &= \bigcap_{a\in A} R(a) \end{aligned}$$

Indeed this idea of starting with the input or primary set of the dicique and from it determining the output or secondary set and vica-versa can be generalized. We can start with any subset as the input set. From that subset determine the corresponding output set and from the output set determine a new enlarged input set. Once we have weaved our way through the relation in the forward sense and then again in the backward sense, there is no need to continue the process. We have reached a fixed point.

Theorem 7 *Let $A \subseteq X$ be given. Define B, A' and B' by*

$$\begin{aligned} B &= \bigcap_{a\in A} R(a) \\ A' &= \bigcap_{b\in B} R^{-1}(b) \\ B' &= \bigcap_{a\in A'} R(a) \end{aligned}$$

Then $B = B'$.

This fixed point theorem has a dual by starting with the output sets and weaving through the relation to the input set and then again to the output set.

Theorem 8 *Let* $B \subseteq Y$ *be given. Define* A, B' *and* A' *by*

$$
\begin{aligned}
A &= \bigcap_{b \in B} R^{-1}(b) \\
B' &= \bigcap_{a \in A} R(a) \\
A' &= \bigcap_{b \in B'} R^{-1}(b)
\end{aligned}
$$

Then $A = A'$.

The fixed point theorems lead us to the diclique representation theorem.

Theorem 9 (P, S) *is a diclique of* $R \subseteq X \times X$ *if and only if for some* $A \subseteq X$,

$$
\begin{aligned}
S &= \bigcap_{x \in A} R(x) \\
P &= \bigcap_{y \in S} R^{-1}(y)
\end{aligned}
$$

The diclique representation theorem leads us to the diclique finding algorithm.

1. $n = 0,\ T_n = \{R(x) \mid x \in X\}$
2. Repeat until no change
3. Compute Q the set of all possible intersections between pairs of sets in T_n
4. $T_{n+1} = T_n \cup Q$
5. At fixed point $\mathcal{D} = \{(\bigcap_{y \in S} R^{-1}(y), S), S \in T_N\}$

6 Coalescing Dicliques

The data collection and aggregation forming the initial relation R may not be complete. This is probably the case most of the time – some facts do not get gathered. The relation R we have at hand may miss some (x, y) pairs. It is possible by working with the dicliques to form hypotheses about which (x, y) pairs may be missing. Such hypotheses are of course an enormous benefit to the operational personal as to where to specifically look or gather additional information to confirm the hypothesis. The technique used to form such hypotheses is called coalescing of dicliques.

Dicliques eligible for coalescing are those that have corresponding primary and secondary sets with large overlap. For example consider the following pair of dicliques.

$$D_1 = \{\{1, 3, 5\}, \{2, 4, 8\}\}$$
$$D_2 = \{\{3, 5, 6\}, \{4, 7, 8\}\}$$

Neglecting whatever else might be in the relation from which these dicliques come, if these dicliques were to be coalesced the new coalesced diclique would be $D = \{\{1, 3, 5, 6\}, \{2, 4, 7, 8\}\}$. And this would indeed be a diclique of a new relation based on the original relation but including the added pairs $(1, 7)$ and$(6, 2)$.

Now adding pairs to the relation to effect a coalescing of some dicliques affects other dicliques. To maintain the consistency between the augmented relation and the dicliques, we must determine the effect of pairs added to relation for all dicliques. Our tool for doing this is based on groupoid homomorphisms.

Let $(G, \circ)$ and $(H, *)$ be groupoids. A function $f : G \to H$ is a *Homomorphism* from $(G, \circ)$ to $(H, *)$ if and only if

$$x, y \in G \text{ implies } \quad f(x \circ y) = f(x) * f(y)$$

We will be wanting to do everything in the framework of groupoid homomorphisms. To do so we will need a characterization of the coalescing that takes place through the homomorphism. This characterization is based on what has historically been called the substitution property partition.

Let $(G, \circ)$ be a groupoid and $\Pi = \{\pi_k\}_{k=1}^{K}$ be a partition over G. Π is called a *Substitution Property Partition* if and only if for every π_i and π_j, there exists a π_k such that

$$x \in \pi_i \text{ and } \quad y \in \pi_j \text{ implies } \quad x \circ y \in \pi_k$$

Groupoid homomorphisms create substitution property partitions as stated in the following theorems.

Theorem 10 *Let $(G, \circ)$ and $(H, *)$ be groupoids and $f : G \rightarrow H$ a homomorphism from $(G, \circ)$ to $(H, *)$. Then*

$$\Pi = \{f^{-1}(h) \mid h \in H\}$$

is a substitution property partition.

Theorem 11 *Let $(G, \circ)$ be a groupoid and Π be a substitution property partition on G. Then $(\Pi, *)$ is a groupoid where the multiplication $*$ is defined by*

$$\pi_i * \pi_j = \pi_k$$

where if $x \in \pi_i$ and $y \in \pi_j$, $x \circ y \in \pi_k$.

Theorem 12 *Let $(G, \circ)$ be a groupoid and Π be a substitution property partition on G. Then $f : G \rightarrow \Pi$ is a homomorphism from $(G, \circ)$ to $(H, *)$*

Theorem 13 *Let $(G, \circ)$ be a groupoid and Π be a partition over G. Π is a substitution property partition if and only if $x, y \in P \in \Pi$ implies for every $z \in G$*

1. *there exists $Q \in \Pi$ such that $x \circ z, y \circ z \in Q$*
2. *there exists $Q' \in \Pi$ such that $z \circ x, z \circ y \in Q'$*

These theorems lead us to the coalescing algorithm in which the algorithm constructs the largest homomorphic image of a commutative groupoid G having elements a and b coalesced.

1. Set up a partition of G in which each cell has one member except for the cell containing both a and b
2. In the multiplication table replace all references of b by a
3. If there exist two columns labeled the same, go to (4)
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 of the first cell. Continue doing this until one pair of these same labeled columns are identical
5. Delete one column of the pairs of identical same labeled columns. Delete the corresponding row. Go to (3)

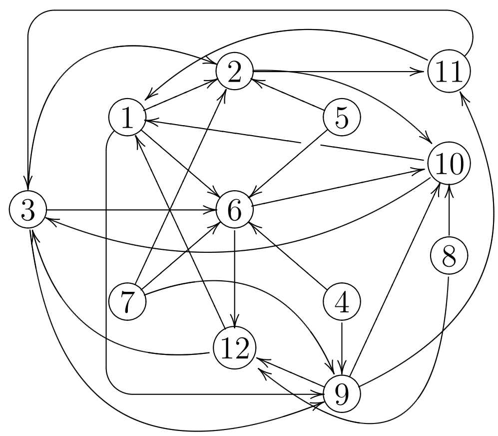

Figure 7: The Example Relation R.

Diclique Label	Primary Set	Secondary Set
A	1,3,7	2,6,9
B	2,9	10,11
C	1,3,4,7	6,9
D	1,3,5,7	2,6
E	6,8,9	10,12
F	9	10,11,12
G	10,11,12	1,3
H	2,6,8,9	10
I	1,3,4,5,7	6
X	ϕ	X
ϕ	X	ϕ

Table 1: The Dicliques for the relation R of Figure 7.

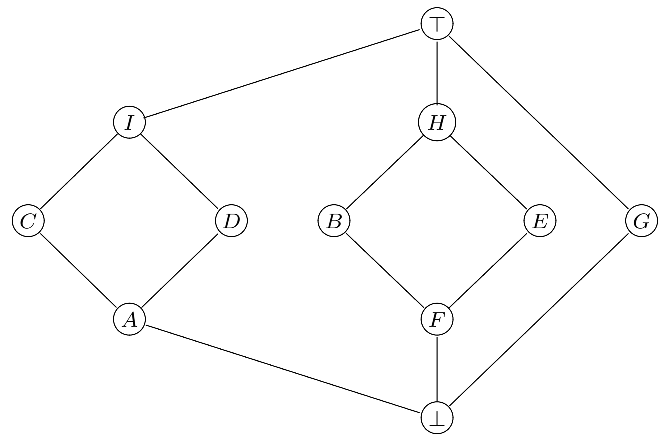

Figure 8: The diclique lattice for the dicliques of the relation R of Figure 7.

	A	B	C	D	E	F	G	H	I	X	ϕ
A	A	ϕ	C	D	ϕ	ϕ	ϕ	ϕ	I	A	ϕ
B	ϕ	B	ϕ	ϕ	H	B	ϕ	H	ϕ	B	ϕ
C	C	ϕ	C	I	ϕ	ϕ	ϕ	ϕ	I	C	ϕ
D	D	ϕ	I	D	ϕ	ϕ	ϕ	ϕ	I	D	ϕ
E	ϕ	H	ϕ	ϕ	E	E	ϕ	H	ϕ	E	ϕ
F	ϕ	B	ϕ	ϕ	E	F	ϕ	H	ϕ	F	ϕ
G	ϕ	ϕ	ϕ	ϕ	ϕ	ϕ	G	ϕ	ϕ	G	ϕ
H	ϕ	H	ϕ	ϕ	H	H	ϕ	H	ϕ	H	ϕ
I	I	ϕ	I	I	ϕ	ϕ	ϕ	ϕ	I	I	ϕ
X	A	B	C	D	E	F	G	H	I	X	ϕ
ϕ	ϕ	ϕ	ϕ	ϕ	ϕ	ϕ	ϕ	ϕ	ϕ	ϕ	ϕ

Table 2: Multiplication Table of Commutative Groupoid for the relation R.

Diclique Label	Primary Set	Secondary Set
C	1,3,4,7	6,9
D	1,3,5,7	2,6
New C	1,3,4,5,7	2,6,9

Table 3: Coalescing Dicliques C and D corresponds to adding (4,2) and (5,9) to the relation R.

Now, coalesce E and H by adding (2,12) to the relation R.

$$B = (\{2,9\},\{10,11\}) \subseteq (\{2,6,8,9\},\{10,11,12\}) = NewE$$

Diclique Lattice

	A	B	C	C	E	F	G	H	I	X	ϕ
A	A	ϕ	C	C	ϕ	ϕ	ϕ	ϕ	I	A	ϕ
B	ϕ	B	ϕ	ϕ	H	B	ϕ	H	ϕ	B	ϕ
C	C	ϕ	C	I	ϕ	ϕ	ϕ	ϕ	I	C	ϕ
C	C	ϕ	I	C	ϕ	ϕ	ϕ	ϕ	I	C	ϕ
E	ϕ	H	ϕ	ϕ	E	E	ϕ	H	ϕ	E	ϕ
F	ϕ	B	ϕ	ϕ	E	F	ϕ	H	ϕ	F	ϕ
G	ϕ	ϕ	ϕ	ϕ	ϕ	ϕ	G	ϕ	ϕ	G	ϕ
H	ϕ	H	ϕ	ϕ	H	H	ϕ	H	ϕ	H	ϕ
I	I	ϕ	I	I	ϕ	ϕ	ϕ	ϕ	I	I	ϕ
X	A	B	C	C	E	F	G	H	I	X	ϕ
ϕ	ϕ	ϕ	ϕ	ϕ	ϕ	ϕ	ϕ	ϕ	ϕ	ϕ	ϕ

Table 4: Coalescing Algorithm: C and D overlap so replace D by C.

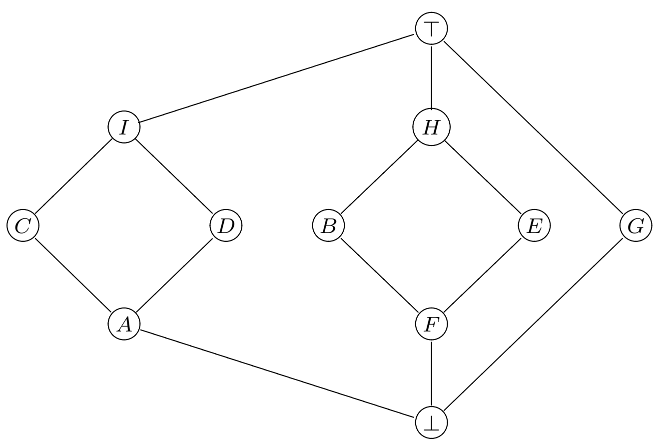

Figure 9: The new diclique lattice after the merging of dicliques C and D.

	A	B	C	C	E	F	G	H	C	X	ϕ
A	A	ϕ	C	C	ϕ	ϕ	ϕ	ϕ	C	A	ϕ
B	ϕ	B	ϕ	ϕ	H	B	ϕ	H	ϕ	B	ϕ
C	C	ϕ	C	C	ϕ	ϕ	ϕ	ϕ	C	C	ϕ
C	C	ϕ	C	C	ϕ	ϕ	ϕ	ϕ	C	C	ϕ
E	ϕ	H	ϕ	ϕ	E	E	ϕ	H	ϕ	E	ϕ
F	ϕ	B	ϕ	ϕ	E	F	ϕ	H	ϕ	F	ϕ
G	ϕ	ϕ	ϕ	ϕ	ϕ	ϕ	G	ϕ	ϕ	G	ϕ
H	ϕ	H	ϕ	ϕ	H	H	ϕ	H	ϕ	H	ϕ
C	C	ϕ	C	C	ϕ	ϕ	ϕ	ϕ	C	C	ϕ
X	A	B	C	C	E	F	G	H	C	X	ϕ
ϕ	ϕ	ϕ	ϕ	ϕ	ϕ	ϕ	ϕ	ϕ	ϕ	ϕ	ϕ

Table 5: By the merging of C and D we are now forced to merge I and C. Replace I by C.

	A	B	C	E	F	G	H	X	ϕ
A	A	ϕ	C	ϕ	ϕ	ϕ	ϕ	A	ϕ
B	ϕ	B	ϕ	H	B	ϕ	H	B	ϕ
C	C	ϕ	C	ϕ	ϕ	ϕ	ϕ	C	ϕ
E	ϕ	H	ϕ	E	E	ϕ	H	E	ϕ
F	ϕ	B	ϕ	E	F	ϕ	H	F	ϕ
G	ϕ	ϕ	ϕ	ϕ	ϕ	G	ϕ	G	ϕ
H	ϕ	H	ϕ	H	H	ϕ	H	H	ϕ
X	A	B	C	E	F	G	H	X	ϕ
ϕ	ϕ	ϕ	ϕ	ϕ	ϕ	ϕ	ϕ	ϕ	ϕ

Table 6: The coalesced groupoid table.

	A	B	A	E	F	G	H	X	ϕ
A	A	ϕ	A	ϕ	ϕ	ϕ	ϕ	A	ϕ
B	ϕ	B	ϕ	H	B	ϕ	H	B	ϕ
A	A	ϕ	A	ϕ	ϕ	ϕ	ϕ	A	ϕ
E	ϕ	H	ϕ	E	E	ϕ	H	E	ϕ
F	ϕ	B	ϕ	E	F	ϕ	H	F	ϕ
G	ϕ	ϕ	ϕ	ϕ	ϕ	G	ϕ	G	ϕ
H	ϕ	H	ϕ	H	H	ϕ	H	H	ϕ
X	A	B	A	E	F	G	H	X	ϕ
ϕ	ϕ	ϕ	ϕ	ϕ	ϕ	ϕ	ϕ	ϕ	ϕ

Table 7: Replace diclique C by A.

	A	B	E	F	G	H	X	ϕ
A	A	ϕ	ϕ	ϕ	ϕ	ϕ	A	ϕ
B	ϕ	B	H	B	ϕ	H	B	ϕ
E	ϕ	H	E	E	ϕ	H	E	ϕ
F	ϕ	B	E	F	ϕ	H	F	ϕ
G	ϕ	ϕ	ϕ	ϕ	G	ϕ	G	ϕ
H	ϕ	H	H	H	ϕ	H	H	ϕ
X	A	B	E	F	G	H	X	ϕ
ϕ	ϕ	ϕ	ϕ	ϕ	ϕ	ϕ	ϕ	ϕ

Table 8: The coalesced table.

Diclique Label	Primary Set	Secondary Set
E	6,8,9	10,12
H	2,6,8,9	10
New E	2,6,8,9	10,12

Table 9: Showing the overlap between dicliques E and H which suggests that they should be coalesced.

	A	B	E	F	G	E	X	ϕ
A	A	ϕ	ϕ	ϕ	ϕ	ϕ	A	ϕ
B	ϕ	B	E	B	ϕ	E	B	ϕ
E	ϕ	E	E	E	ϕ	E	E	ϕ
F	ϕ	B	E	F	ϕ	E	F	ϕ
G	ϕ	ϕ	ϕ	ϕ	G	ϕ	G	ϕ
E	ϕ	E	E	E	ϕ	E	E	ϕ
X	A	B	E	F	G	E	X	ϕ
ϕ	ϕ	ϕ	ϕ	ϕ	ϕ	ϕ	ϕ	ϕ

Table 10: Replace diclique H by E.

	A	B	E	F	G	X	ϕ
A	A	ϕ	ϕ	ϕ	ϕ	A	ϕ
B	ϕ	B	E	B	ϕ	B	ϕ
E	ϕ	E	E	E	ϕ	E	ϕ
F	ϕ	B	E	F	ϕ	F	ϕ
G	ϕ	ϕ	ϕ	ϕ	G	G	ϕ
X	A	B	E	F	G	X	ϕ
ϕ	ϕ	ϕ	ϕ	ϕ	ϕ	ϕ	ϕ

Table 11: The Coalesced Table.

Diclique Label	Primary Set	Secondary Set
B	2,9	10,11
New E	2,6,8,9	10,12
F	9	10,11,12
New E	2,6,8,9	10,11,12

Table 12: Coalesce E and F by adding (6,11), (8,11).

	A	B	E	E	G	X	ϕ
A	A	ϕ	ϕ	ϕ	ϕ	A	ϕ
B	ϕ	B	E	B	ϕ	B	ϕ
E	ϕ	E	E	E	ϕ	E	ϕ
E	ϕ	B	E	E	ϕ	E	ϕ
G	ϕ	ϕ	ϕ	ϕ	G	G	ϕ
X	A	B	E	E	G	X	ϕ
ϕ	ϕ	ϕ	ϕ	ϕ	ϕ	ϕ	ϕ

Table 13: Replace F by E.

Diclique Label	Primary Set	Secondary Set
B	2,9	10,11
New E	2,6,8,9	10,11,12

Table 14: The overlap between B and the new E suggests that they be coalesced.

	A	B	B	B	G	X	ϕ
A	A	ϕ	ϕ	ϕ	ϕ	A	ϕ
B	ϕ	B	B	B	ϕ	B	ϕ
B	ϕ	B	B	B	ϕ	B	ϕ
B	ϕ	B	B	B	ϕ	B	ϕ
G	ϕ	ϕ	ϕ	ϕ	G	G	ϕ
X	A	B	B	B	G	X	ϕ
ϕ	ϕ	ϕ	ϕ	ϕ	ϕ	ϕ	ϕ

Table 15: Forced: Replace E by B.

	A	B	G	X	ϕ
A	A	ϕ	ϕ	A	ϕ
B	ϕ	B	ϕ	B	ϕ
G	ϕ	ϕ	G	G	ϕ
X	A	B	G	X	ϕ
ϕ	ϕ	ϕ	ϕ	ϕ	ϕ

Table 16: Coalesced

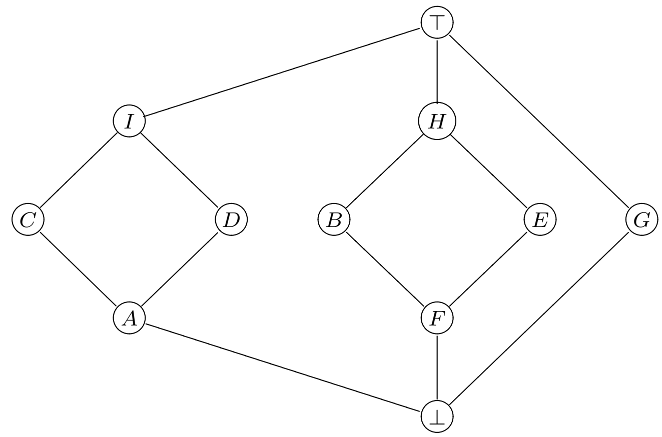

Figure 10: Diclique Lattice

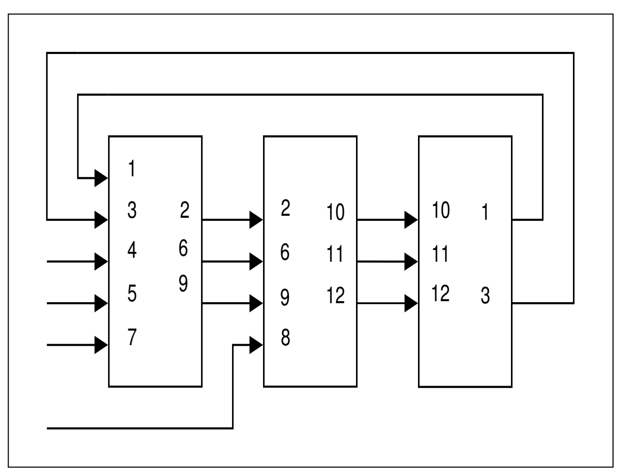

Figure 11: In the system diagram rectangles correspond to dicliques; lines correspond to elements of X.

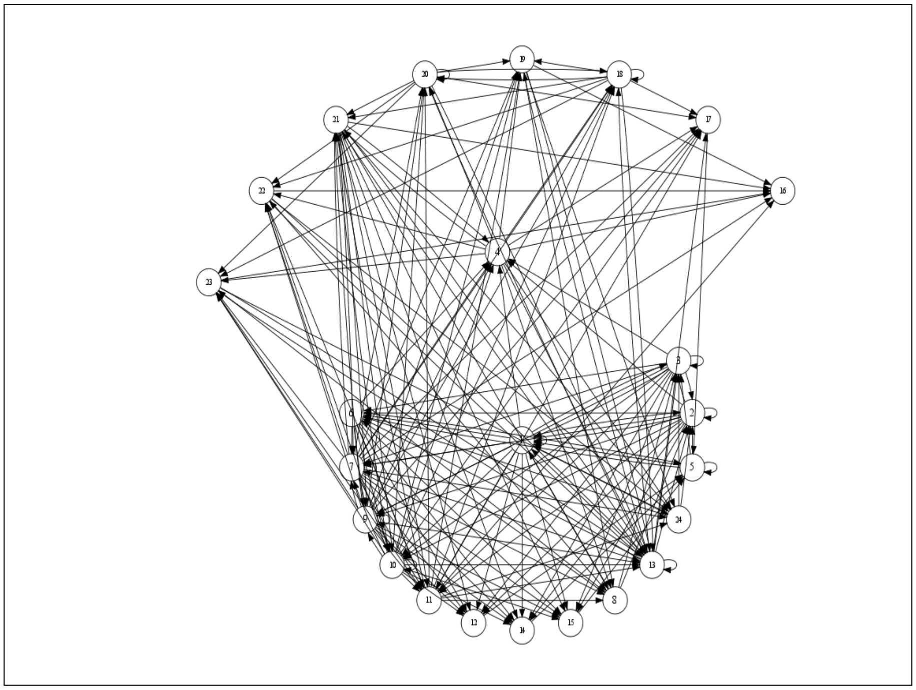

Figure 12: The graph of a more complex relation.

7 Conclusion

We have developed the concept of diclique and have shown how the dicliques of a relation maximally aggregate relation information pieces together. We have summarized the mathematical properties of dicliques showing that the set of dicliques forms a lattice and in two different ways forms commutative idempotent monoids. Finally, by investigating groupoid homomorphisms on the monoid structure we have shown how to coalesce overlaping dicliques and how coalesing a pair of dicliques may require coalescing other dicliques. We have stated the theorems governing these forced coalescings and have demonstrated all this in an example.

Future work will be in the use of time local relations where we believe that the dicliques correspond to events. In general relations, dicliques correspond to subsystems. We hope to do future research by developing a formal framework for events and subsystems. This formal framework would then permit us to algorithmically infer the event or subsystem from the observations.

The work that we have done for simple unlabled binary relations needs to be extended to labeled binary relations and needs to be extended to N-ary relations.

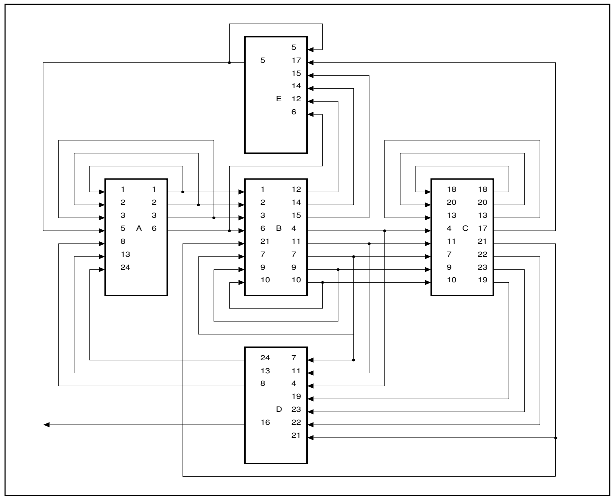

Figure 13: The dicliques of the more complex relation represented as a system diagram.

Although in practical application in counter-terrorism, it is more likely to have pairs of the relation missing then extraneous pairs in the relation, we would nevertheless like to have a theory for the removal of pairs in the relation that would simplify the diclique description of the relation. Finally, work needs to be done on developing fast algorithms for finding dicliques.

References

[1] R. Haralick. The diclique representation and decomposition of binary relations. *Journal of the Association for Computing Machinery*, pages 356–366, July 1974.