A GRAPH-THEORETIC APPROACH TO LINEARLY ORDERABLE SETS

by

Kapali P. Eswaran

Memorandum No. ERL-M369

28 November 1972

A GRAPH-THEORETIC APPROACH TO LINEARLY ORDERABLE SETS

by

Kapali P. Eswaran

Memorandum No. ERL-M369

28 November 1972

ELECTRONICS RESEARCH LABORATORY

College of Engineering
University of California, Berkeley
94720

A GRAPH-THEORETIC APPROACH TO LINEARLY ORDERABLE SETS

by

Kapali P. Eswaran

Department of Electrical Engineering and Computer Sciences
and the Electronics Research Laboratory
University of California, Berkeley, California 94720

ABSTRACT: This memo studies a family of sets which can be represented by intervals in real-line. Such a family is said to be linearly orderable. Relationships between the intersection graph of a family that is linearly orderable and the interval graph have been developed. Necessary and sufficient conditions for a family of sets to be linearly orderable are derived. It is shown that the Consecutive Retrieval Storage Organization is a direct application of the property of linear ordering.

Research sponsored by the Naval Electronic Systems Command, Contract N00039-71-C-0255.

Section 1: Interval Graphs and Linearly Orderable Sets

In this section we define as to what we mean by linearly orderable sets. We will also show the relations between interval graphs and the intersection graph of a family of sets that is linearly orderable.

Let $Q = \{q_2, q_2, \ldots, q_m\}$ be a family of distinct, non-empty, finite sets. The intersection graph of Q is denoted by $\Omega(Q)$ and is defined as follows: for each set $q_i \in Q$, there exists a corresponding node $(q_i) \in \Omega(Q)$ and vice versa and for $i \neq j$, (q_i) is connected with (q_j) iff $q_i \cap q_j \neq \phi$.

Example 1:

Let $Q = \{q_1, q_2, q_3, q_4, q_5\}$

where $q_1 = \{a_1, a_4, a_6, a_7\}$

$q_2 = \{a_1, a_2, a_5\}$

$q_3 = \{a_1, a_6, a_7\}$

$q_4 = \{a_1, a_2, a_3, a_4, a_5\}$

and $q_5 = \{a_2, a_3, a_5\}$

The intersection graph $\Omega(Q)$ of Q is given in figure 1.

Let G be any graph. If it is possible to assign to each node (a_i) of G, a distinct interval I_i in the real line such that I_i overlaps with I_j iff nodes (a_i) and (a_j) are connected, then G is called an interval graph. The intervals may be open or closed. In the sequel, we shall assume that the intervals are closed. In particular, if $\Omega(Q)$ is an interval graph, then we see the family of sets Q may be represented by a set of intervals on the real line.

Example 2:

The graph $\Omega(Q)$ in example 1 is an interval graph.

Let $I_1 = [0,4]$

$I_2 = [2,7]$

$I_3 = [-2,3]$

$I_4 = [1,8]$

$I_5 = [5,7]$ where I_i, $1 \leq i \leq 5$ corresponds to node $\textcircled{q_i}$.

Hereafter Q will denote the family of sets $\{q_1, q_2, \ldots, q_m\}$ and S the set $\bigcup_{\substack{a_i \in q_j \\ q_j \in Q}} a_i = \{a_1, a_2, \ldots, a_n\}$. Elements belonging to the set $(S-q_i)$ are called foreign w.r.t. q_i. Suppose there exists a 1-1 function f that maps the elements of S into (points in) the real line R such that for each $q_i \in Q$, there exists an interval I_i containing images of all elements $\in q_i$ but not images of any foreign elements w.r.t. q_i. Then we say that the family Q possesses the property of linear ordering or Q is linearly orderable. The intersection graph $\Omega(Q)$ is called a linearly orderable graph (w.r.t. this particular family Q) or in short a L.O. graph. Any interval that contains images of all elements $\in q_i$ is said to correspond to q_i.

Theorem 1:

If $\Omega(Q)$ is a L.O. graph, then it is an interval graph.

Proof: Let $Q = \{q_1, q_2, \ldots, q_m\}$. Since $\Omega(Q)$ is a L.O. graph, there exists a 1-1 function f and intervals $I_1, I_2, \ldots, I_i, \ldots, I_m$ where $I_i = [\underset{a_p \in q_i}{\text{Min}}(f(a_p)), \underset{a_p \in q_i}{\text{Max}}(f(a_p)) + \delta_i]$. I_i, for $1 \leq i \leq m$, corresponds to q_i. Further, for δ_i small enough, I_i does not contain images of foreign elements w.r.t. q_i. (Note that the increment δ_i is added to I_i to take care of the situation that q_i may be a singleton.) Then, I_i overlaps with I_j, $j \neq i$, iff $q_i \cap q_j \neq \phi$. But $q_i \cap q_j \neq \phi$ iff (q_i) and (q_j) are connected in $\Omega(Q)$. QED

Theorem 2:

If G is an interval graph, then there exists a family of sets Q for which G is a L.O. graph. In other words every interval graph is a L.O. graph for some family of sets.

Proof: Let {(q_1), (q_2),..., (q_m)} be the set of nodes of the interval graph G and $I = \{I_1, I_2, \ldots, I_m\}$ be a set of distinct intervals s.t. interval I_i corresponds to node (q_i). We shall assume that the intervals are finite. If I_i is an infinite interval, we can always choose a finite subinterval I_i' of I_i s.t. I_i' overlaps with I_j iff I_i overlaps with I_j. I_i may be replaced by I_i' in I.

Define i_{min} = Minimum of interval I_i.

i_{max} = Maximum of interval I_i.

We, now, define sets $q_1, q_2, \dots q_m$.

For $1 \le i \le m$, $q_i = \{i_{min}, i_{max}\} \cup \{j_{min} | i_{min} \le j_{min} \le i_{max}\}$

$$\cup \{j_{max} | i_{max} \le j_{max} \le i_{max}\}$$

Since $\{i_{min}, i_{max}\} \neq \{j_{min}, j_{max}\}$, it follows that $q_i \neq q_j$ for $i \neq j$. If (q_i) and (q_j) are connected in G, then intervals I_i and I_j overlap. Then we have $\{i_{min}, i_{max}, j_{min}\} \subseteq q_i$ or $\{i_{min}, i_{max}, j_{max}\} \subseteq q_i$. Since $\{j_{min}, j_{max}\} \subseteq q_j$, $q_i, \cap q_j \neq \phi$.

Let $Q = \{q_1, q_2, \dots q_m\}$. We first observe that G is $\Omega(Q)$. Let $S = \bigcup_{\substack{a_i \in q_j \\ q_j \in Q}} a_i$. Note that the elements of S are either minimum or maximum of some interval $\in$ I.

We need to prove that Q has L.O. property. Define $S_{min} = \underset{a_i \in S}{Min}(a_i)$ and $S_{max} = \underset{a_i \in S}{Max}(a_i)$. Let f be the identity function that maps S into points in $[S_{min}, S_{max}]$. We claim that the function f and the set of intervals I imply that Q is linearly orderable. To see this, assume to the contrary that f and I do not imply the L.O. property of Q. Then, there exists an interval, say I_i, s.t. I_i contains the image of at least one foreign element, say k, w.r.t. q_i, the set corresponding to I_i. I_i is $[i_{min}, i_{max}]$. Then k is s.t. $i_{min} \le k \le i_{max}$. Since $k \in S$, k is either j_{min} or j_{max} for some interval I_j, $j \neq i$. This implies that I_j overlaps I_i and $k \in q_i$. Then k is not foreign to q_i. Contradiction. QED.

<u>Example 3</u>:

Let $q_1 = \{b, c, g, h, a\}$

$q_2 = \{a, e, d\}$

$q_3 = \{h, a, e, d\}$

$q_4 = \{c, g, h, e, a\}$

and $q_5 = \{e, d\}$

Let $Q = \{q_1, q_2, q_3, q_4, q_5\}$

The intersection graph $\Omega(Q)$ is shown in figure 2. $\Omega(Q)$ is an interval graph. Let the intervals corresponding to the nodes be:

$I_1 = [1, 5]$

$I_2 = [5, 7]$

$I_3 = [4, 7]$

$I_4 = [2, 6]$

$I_5 = [6, 7]$. Interval I_i corresponds to node $\textcircled{q_i}$ for $1 \leq i \leq 5$. The intervals are represented pictorially in figure 3. The node that an interval represents is given in parenthesis in the figure.

$\Omega(Q)$ is also a L.O. graph. Let a function f be:

$f(b) = 1$

$f(c) = 2$

$f(g) = 3$

$f(h) = 4$

$f(a) = 5$

$f(e) = b$ and $f(d) = 7$

The function f and the intervals I_1, I_2, I_3, I_4 and I_5 imply that $\Omega(Q)$ is a L.O. graph. The pre-image of i for i = 1, 2, ... 7, is given in parenthesis next to i in figure 3.

If G is any graph, then the complement of G, denoted by G^c, is a graph that has the same nodes as G with an edge connecting a pair of distinct nodes in G^c iff that edge does not occur in G. Let G be an undirected graph defined by [V, R], where V is a finite non-empty set of nodes of G and R is a irreflexive relation on V s.t. $\forall$ $ⓐ_i$, $ⓐ_j$ $\in$ V, $i \neq j$, $ⓐ_i$ R $ⓐ_j$ iff $ⓐ_i$ is connected with $ⓐ_j$ in G. An undirected graph [V, R] is transitive orientable iff there exists a directed graph $\tilde{G} = [V, \tilde{R}]$ s.t. for $\forall$ $ⓐ_i$, $ⓐ_j$, $ⓐ_k$ $\in$ V, if $ⓐ_i$ R $ⓐ_j$ then either $ⓐ_i$ $\tilde{R}$ $ⓐ_j$ or $ⓐ_j$ $\tilde{R}$ $ⓐ_i$ and $ⓐ_i$ $\tilde{R}$ $ⓐ_j$, $ⓐ_j$ $\tilde{R}$ $ⓐ_k$ $\Rightarrow$ $ⓐ_i$ $\tilde{R}$ $ⓐ_k$. $ⓐ_i$ $\tilde{R}$ $ⓐ_j$ iff there is an edge from $ⓐ_i$ to $ⓐ_j$ in $\tilde{G}$. In other words, it is possible to assign directions to the edges of G s.t. G is transitive. A transitive orientable graph is sometimes called a comparable graph. The following theorem is due to Gilmore and Hoffman and is given here for the sake of completeness. The proof may be found in reference [1].

<u>Theorem 3</u>:

A graph G is an interval graph iff every quadrilateral in G has a diagonal and G^c is transitive orientable.

Pnueli et al. [2] give an algorithm to check if a graph is transitive orientable.

Example 4:

Consider the graph in figure 2. We saw that it was an interval graph. The quadrilaterals (q_1, q_2, q_3, q_4), (q_1, q_2, q_3, q_5), (q_1, q_2, q_4, q_5), (q_1, q_3, q_4, q_5), (q_2, q_3, q_4, q_5) have at least one diagonal and the complement of the graph is transitive orientable. See figure 4.

Section 2: Singleton sets in a family of sets.

In this section we show that the singleton sets in a family of sets do not influence the linear ordering property of the family.

We shall first introduce a notation. Q and S are as defined before. Set f be a 1 - 1 function that maps S into R and $I_1, I_2, \ldots I_m$ be a set of intervals on R such that I_i corresponds to q_i. Further let I_i, for $1 \leq i \leq m$, not contain images of any foreign elements w.r.t. q_i. Then we say that $(f; I_1, I_2, \ldots I_m)$ implies that Q has L.O. property.

Lemma 1:

If Q is linearly orderable, then $Q' \subseteq Q$ is linearly orderable.

Proof:

Since Q is linearly orderable, there exists a function f and intervals $I_1, I_2, \ldots I_m$ implying the L.O. property of Q.

Let $S' = \bigcup_{\substack{a_i \in q_j \\ q_j \in Q'}} a_i$ where $Q' \subseteq Q$

Now, define f': $f'(a_i) = f(a_i) \quad \forall a_i \in S'$

$I'_i = I_i$ for $\forall q_i \in Q'$

f' and $\{I'_i \mid q_i \in Q'\}$ imply the L.O. property of Q'. QED.

Lemma 2:

Let $Q = \{q_1, q_2, \ldots q_m\}$, $S = \bigcup_{\substack{a_i \in q_j \\ q_j \in Q}} a_i = \{a_1, a_2, \ldots a_n\}$

and $\overline{q}_j = \{a_j\}$ for $1 \leq j \leq n$. Then $\Omega(Q)$ is a L.O. graph iff $\Omega(\overline{Q})$ is a L.O. graph where $\overline{Q} = Q \cup \{\overline{q}_i\}$, $i \in \{1, 2, \ldots n\}$.

Proof: If Q is L.O., then there exists a function f and intervals $I_1, I_2, \ldots I_m$ s.t. they satisfy the L.O. property of Q. Let $\overline{I}_i = [f(a_i), f(a_i) + \delta_i]$ for $\forall\ \overline{q}_i \in \overline{Q}$. For δ_i small enough, $\overline{I}_i$ does not contain images of any elements other than a_i. Then f and $\{I_i\} \cup \{\overline{I}_i \mid \overline{q}_i \in \overline{Q}\}$ imply that $\Omega(\overline{Q})$ is a L.O. graph.

$\Leftarrow$. If $\overline{Q}$ is linearly orderable, then by Lemma 1, $Q \subseteq \overline{Q}$ is linearly orderable. QED.

We can, therefore, assume that as far as linear ordering is concerned, no set in Q is a singleton.

Section 3: Directed Semantic Graphs And L.O. Graphs.

In this section, we derive a number of results regarding the L.O. property of Q when $\Omega(Q)$ is a complete graph. We establish necessary and sufficient conditions for an intersection graph, that is complete, to be a L.O. graph.

A graph G is complete iff every pair of distinct nodes is joined by an edge in G; i.e. no more edge can be added to G.

Lemma 3:

If $\Omega(Q)$ is complete and is a L.O. graph, then $I = \bigcap_{q_i \in Q} q_i \neq \phi$

Proof: Let $(f; I_1, I_2, \ldots, I_m)$ imply the L.O. property of Q. As $\Omega(Q)$ is complete, we have $q_i \cap q_j \neq \phi$ for $1 \leq i, j \leq m$. This implies that $I_i \cap I_j \neq \phi$ and there exists an $a_{ij} \in S$ s.t. $f(a_{ij}) \in (I_i \cap I_j)$ for $i, j = 1, 2, \ldots m$.

We shall assume that all the intervals are finite (see proof of theorem 2). Let $I_i = [i_{min}, i_{max}]$ for $1 \leq i \leq m$. Let I_p be s.t. $p_{max} = \underset{1 \leq j \leq m}{\text{Min}} [j_{max}]$ and I_k be s.t. $k_{min} = \underset{1 \leq j \leq m}{\text{Max}} [j_{min}]$. We first observe that $k_{min} \leq p_{max}$. For, if $k_{min} > p_{max}$, intervals I_p and I_k would not overlap which would be a contradiction. Since $\Omega(Q)$ is complete and is an interval graph and $i_{min} \leq k_{min}$ and $i_{max} \geq p_{max}$ for $\forall\, i = 1, 2, \ldots m$, all intervals contain the subinterval $[k_{min}, p_{max}]$ which is $I_p \cap I_k$ (see figure 5). But, we know that the intersection of every pair of intervals contains the image of at least one element $\in S$. Then there exists an $a_{pk} \in S$ s.t. $f(a_{pk}) \in I_p \cap I_k \Rightarrow f(a_{pk})$ belongs to all intervals $I_1, I_2, \ldots, I_m$. Since $\Omega(Q)$ is a L.O. graph, a_{pk} is not foreign to any set $\in Q$, i.e. $\bigcap_{q_i \in Q} q_i \neq \phi$.

Define a directed semantic graph $\overline{G} = [V, R, I]$. V is a finite non-empty set of nodes. R is a irreflexive relation on V s.t. $\forall\, a_i, a_j \in V$, $i \neq j$, $a_i R\, a_j \Leftrightarrow$ there is an edge from a_i to a_j in $\overline{G}$. R is called the connectivity relation of $\overline{G}$. An undirected edge between nodes a_i and a_j in $\overline{G}$ means that $a_i\ R\ a_j$ and $a_j\ R\ a_i$. I is a subset of V. Nodes $\in$ I are called direction-changer nodes and are denoted by an * mark in $\overline{G}$. (a_i, a_j) denotes the edge between a_i and a_j, ignoring the direction on the edge.

$\langle a_i, a_j \rangle$ denotes the directed edge from node a_i to node a_j. A path in a directed semantic graph (DSG) $\overline{G}$ is a sequence of distinct nodes $a_0, a_1, \ldots, a_i, a_{i+1}, \ldots a_k$ of $\overline{G}$ s.t. for $0 \le i \le k-1$, $\langle a_i, a_{i+1} \rangle$ is an edge of $\overline{G}$ when in direct mode and $\langle a_{i+1}, a_i \rangle$ is an edge of $\overline{G}$ when in reverse mode, where the modes are defined as follows: If a path starts with a non-direction-changer node, then the mode is direct. If it starts with a direction-changer-node, the mode is reverse. Whenever a direction-changer node is reached from a non-direction-changer node, the mode is switched. (If a direction-changer node is reached from a direction-changer node, no change of mode occurs.) We shall enclose the sequence of nodes defining a path in angle brackets, $\langle$ and $\rangle$. If $P = \langle a_0, a_1, \ldots, a_k \rangle$ is a path of $\overline{G}$, then a_0 is called the starting node of P, a_k the end node of P and $a_1, a_2, \ldots a_{k-1}$ the intermediate nodes of P. Note that if $I = \phi$, then our definition of a path is the same as the usual definition of a directed-path in a directed graph. Because of the presence of direction changer nodes in $\overline{G}$, there is some semantics in the definitions regarding $\overline{G}$. Hence the name directed semantic graph.

Example 5:

Consider the DSG, $\overline{G}$, shown in figure 6.

We have $V = \{a_0, a_1, a_2, a_3, a_4, a_5, a_6, a_7\}$

$$I = \{a_2, a_5\}$$

and R is the connectivity relation. $\langle a_0, a_1, a_2 \rangle$, $\langle a_1, a_2, a_3 \rangle$, $\langle a_4, a_3, a_2 \rangle$, $\langle a_2, a_3, a_4, a_5, a_6 \rangle$, $\langle a_0, a_1, a_2, a_3, a_4, a_5, a_6, a_7 \rangle$, $\langle a_1, a_2, a_3, a_4, a_5, a_0 \rangle$ are some of the paths in $\overline{G}$.

A Hamiltonian path in a directed semantic graph $\overline{G}$ is a path that passes through all the nodes of $\overline{G}$.

Example 6:

Consider the graph $\overline{G}$ in figure 6. $\overline{G}$ has only one Hamiltonian path which is $\langle a_0, a_1, a_2, a_3, a_4, a_5, a_6, a_7 \rangle$.

We now define the DSG of a family of sets. Let $I = \bigcap_{q_i \in Q} q_i$.

Let $\overline{R}$ be an irreflective relation defined on S as follows: $a_i \overline{R} a_j$ iff $i \neq j$ and for $\forall q_k \in Q$, $a_i \in q_k \Rightarrow a_j \in q_k$. Note that $\overline{R}$ is transitive. The directed semantic graph of Q is $\overline{G}(Q) = [S', \overline{R}, I']$. S' is the set of nodes of $\overline{G}(Q)$ and is $\{(a_1), (a_2), \ldots, (a_i), \ldots, (a_n)\}$ where node (a_i) corresponds to element $a_i \in S$ and vice versa. $(a_i) \overline{R} (a_j)$ iff $a_i \overline{R} a_j$. We use the same symbol $\overline{R}$ for a relation between two elements of S and two elements $\in S'$ since there is no confusion. $\overline{R}$ is the connectivity relation of $\overline{G}(Q)$. I' is the set of direction-changer nodes of $\overline{G}(Q)$ with $(a_p) \in I'$ iff $a_p \in I$.

Example 7:

Let $q_1 = \{a_2, a_3\}$

$q_2 = \{a_1, a_2, a_3\}$

$q_3 = \{a_2, a_3, a_4\}$

and $q_4 = \{a_3, a_4, a_5\}$

we have $Q = \{q_1, q_2, q_3, q_4\}$ and $S = \{a_1, a_2, a_3, a_4, a_5\}$

$$I = \bigcap_{q_i \in Q} q_i = \{a_3\}$$

$\overline{R}$: $a_1 \overline{R} a_2$, $a_1 \overline{R} a_3$

$a_2 \overline{R} a_3$

$a_4 \overline{R} a_3$

$a_5 \overline{R} a_4$, $a_5 \overline{R} a_3$

$\overline{G}(Q) = [S', \overline{R}, I']$ where

$S' = \{ⓐ_1, ⓐ_2, ⓐ_3, ⓐ_4, ⓐ_5\}$ and $I' = \{ⓐ_3\}$.

$\overline{R}$ is the connectivity relation. $\overline{G}(Q)$ is given in figure 7. $\langle ⓐ_1, ⓐ_2, ⓐ_3, ⓐ_4, ⓐ_5 \rangle$, $\langle ⓐ_5, ⓐ_4, ⓐ_3, ⓐ_2, ⓐ_1 \rangle$ are Hamiltonian paths in $\overline{G}(Q)$.

Let $h = \langle ⓐ_0, ⓐ_1, \ldots, ⓐ_i, ⓐ_{i+1}, \ldots, ⓐ_{j-1}, ⓐ_j, \ldots, ⓐ_k \rangle$ be a path in a DSG. A subpath h' of h is $\langle ⓐ_i, ⓐ_{i+1}, \ldots, ⓐ_{j-1}, ⓐ_j \rangle$ where $i \geq 0$ and $j \leq k$. A set of nodes I" is said to be between $ⓐ_i$ and $ⓐ_j$ in a path P in a DSG iff all and only the nodes $\in$ I" are in the subpath of P from $ⓐ_i$ to $ⓐ_j$. For example let $P = \langle ⓐ_4, ⓐ_1, ⓐ_2, ⓐ_3 \rangle$. The set of nodes between $ⓐ_1$ and $ⓐ_3$ is $\{a_2\}$ and the set of nodes between $ⓐ_3$ and $ⓐ_1$ is null.

<u>Lemma 4</u>: Let $\overline{G}(Q)$ be the DSG of Q and h be any Hamiltonian path in $\overline{G}(Q)$. Then, there does not exist a subpath h' of h s.t. the starting and end

nodes of h' are direction-changer nodes and the intermediate nodes are non-direction-changer nodes.

<u>Proof</u>: Let $I = \bigcap_{q_i \in Q} q_i$. $\textcircled{a_i}$ is a direction-changer node of $\overline{G}(Q)$ iff $a_i \in I$. $\textcircled{a_i}$ is a non-direction-changer node of $\overline{G}(Q)$ iff $a_i \in (S-I)$. Assume to the contrary that there exists a subpath h' of h s.t. h' = $\langle \textcircled{a_i}, \textcircled{a_{i+1}}, \ldots, \textcircled{a_{j-1}}, \textcircled{a_j} \rangle$ where $\textcircled{a_i}$ and $\textcircled{a_j}$ are direction-changer nodes and $\textcircled{a_{i+1}}, \textcircled{a_{i+2}}, \ldots, \textcircled{a_{j+1}}$ are not. Since $a_i \in I$ and $a_{i+1} \in (S-I)$ we have $a_{i+1}\ \overline{R}\ a_i$ where $\overline{R}$ is the connectivity relation of $\overline{G}(Q)$. Similarly $a_{j-1}\ \overline{R}\ a_j$. Hence we have edge $(\textcircled{a_{i+1}}, \textcircled{a_i})$ directed from $\textcircled{a_{i+1}}$ to $\textcircled{a_i}$ and edge $(\textcircled{a_{j-1}}, \textcircled{a_j})$ directed from $\textcircled{a_{j-1}}$ to $\textcircled{a_j}$ - i.e. $\langle \textcircled{a_{i+1}}, \textcircled{a_i} \rangle$ and $\langle \textcircled{a_{j-1}}, \textcircled{a_j} \rangle$ are edges of $\overline{G}(Q)$. (see Fig. 8)

In the subpath h', since there are no direction-changer nodes between $\textcircled{a_i}$ and $\textcircled{a_j}$ we should have either (i) edges $\langle \textcircled{a_i}, \textcircled{a_{i+1}} \rangle$ and $\langle \textcircled{a_{j-1}}, \textcircled{a_j} \rangle$ or (ii) edges $\langle \textcircled{a_{i+1}}, \textcircled{a_i} \rangle$ and $\langle \textcircled{a_j}, \textcircled{a_{j-1}} \rangle$. In either case, we have a situation that contradicts the earlier statement that $\langle \textcircled{a_{i+1}}, \textcircled{a_i} \rangle$ and $\langle \textcircled{a_{j-1}}, \textcircled{a_j} \rangle$ are edges of $\overline{G}(Q)$. QED.

<u>Corollary 1</u>: Any Hamiltonian path in the DSG of Q should have a subpath of the form $\langle \textcircled{a_i}, \textcircled{a_{i+1}}, \ldots, \textcircled{a_{j-1}}, \textcircled{a_j} \rangle$ where $\{a_i, a_{i+1}, \ldots, a_{j+1}, a_j\} = I = \bigcap_{q_i \in Q} q_i$.

<u>Lemma 5</u>: If there exists a Hamiltonian path in $\overline{G}(Q)$, then $\bigcap_{q_i \in Q} q_i \neq \phi$, i.e. there exists at least one direction changer node in $\overline{G}(Q)$.

Proof: Suppose to the contrary that the set of direction-changer nodes in $\overline{G}(Q) = \phi$. Let $h = \langle (a_0), (a_1), \ldots, (a_n) \rangle$ be a Hamiltonian path of $\overline{G}(Q)$. Then $\langle (a_0), (a_1) \rangle, \langle (a_1), (a_2) \rangle, \ldots \langle (a_{n-1}), (a_n) \rangle$ are among the edges of $\overline{G}(Q)$. Relation $\overline{R}$ is transitive. Then for $0 \leq p \leq n-1$, $\langle (a_p), (a_n) \rangle$ is an edge of $\overline{G}(Q)$. This means that $a_n \in q_j$ for $\forall q_j \in Q \Rightarrow \bigcap_{q_j \in Q} q_j \neq \phi$. Contradiction.

The following theorem gives another necessary condition for the existence of a Hamiltonian path in $\overline{G}(Q)$.

Theorem 4: Let there exist a Hamiltonian path in $\overline{G}(Q)$. If $S_1 \subseteq S$ is a set of incomparable elements (w.r.t. $\overline{R}$), then $\# \{S_1\} \leq 2$.

Proof: By contradiction. Suppose there exists a set $S_1 = \{a_i, a_j, a_k\}$ of incomparable elements and $S_1 \subseteq S$. $S_1' = \{(a_i), (a_j), (a_k)\}$ is the set of nodes of $\overline{G}(Q)$ that correspond to S_1. Let h be a Hamiltonian path of $\overline{G}(Q)$.

Let $I = \bigcap_{q_i \in Q} q_i$. Since all the elements of S are $\overline{R}$ - related with the elements $\in I$, we have $S_1 \cap I = \phi$, i.e. none of the nodes $\in S_1'$ is a direction-changer node. Without loss of generality, we shall assume that in the Hamiltonian path h of $\overline{G}(Q)$, (a_i) precedes (a_j) and (a_j) precedes (a_k).

By lemma 4, in any Hamiltonian path non-direction-changer nodes are not present between any two direction-changer nodes. We then have only the following cases for h:

Case (i) h passes through all the direction-changer nodes after leaving

(a_j). Then $\langle (a_i), (a_{i+1}) \rangle$, $\langle (a_{i+1}), (a_{i+2}) \rangle$, ..., $\langle (a_{j-1}), (a_j) \rangle$ are all edges of $\overline{G}(Q)$. $\overline{R}$ is transitive. Thus we have that $\langle (a_i), (a_j) \rangle$ is an edge of $\overline{G}(Q)$. But this is not possible since a_i and a_j are not comparable w.r.t. $\overline{R}$.

<u>Case (ii)</u> h passes through all the direction-changer nodes before reaching (a_i). After an argument similar to that of case (i), we get $\langle (a_j), (a_i) \rangle$ which again contradicts the incomparability of a_i and a_j.

<u>Case (iii)</u> h visits all the direction-changer nodes between (a_i) and (a_j). After visiting (a_j), h needs to visit (a_k) and this is the same as case (ii) which leads to a contradiction.

By cases (i), (ii) and (iii), we see that h can not visit all (a_i), (a_j) and (a_k). Then h is not a Hamiltonian path which is a contradiction. QED.

<u>Corollary 2</u>: For all $a_i, a_j \in S$, $i \neq j$, a_i and a_j are incomparable (w.r.t. $\overline{R}$) only if in any Hamiltonian path of $\overline{G}(Q)$, (a_i) and (a_j) exist on the opposite sides of the subpath which passes through all the direction-changer nodes.

The following lemma leads us to the connection between the linear ordering property of a family Q and the existence of a Hamiltonian path in $\overline{G}(Q)$, when $\Omega(Q)$ is complete.

<u>Lemma 6</u>: Let $h = \langle (a_1), \ldots, (a_i), (a_{i+1}), \ldots, (a_{j-1}), (a_j), \ldots, (a_n) \rangle$ be a Hamiltonian path of $\overline{G}(Q)$. If $\{a_i, a_j\} \subseteq q_p \in Q$ then $\{a_i, a_{i+1}, a_{i+2}, \ldots a_{j-1}, a_j\} \subseteq q_p$.

Proof: Let $I = \bigcap_{q_i \in Q} q_i$. We have three situations.

(1) Both (a_i) and (a_j) are direction changer nodes. By lemma 4, all the nodes between (a_i) and (a_j) are also direction-changer nodes. Then the elements corresponding to (a_{i+1}), (a_{i+2}), ..., (a_{j-1}) belong to I. Hence the lemma.

(2) Both (a_i) and (a_j) are not direction-changer nodes. For this situation, we have the following possible cases similar to the ones we had in the proof of theorem 3.

Case (i): h visits all the direction-changer nodes after leaving (a_j). Then $\langle (a_i), (a_{i+1}) \rangle$, $\langle (a_{i+1}), (a_{i+2}) \rangle$, ..., $\langle (a_{j-1}), (a_j) \rangle$ are edges of $\overline{G}(Q)$. Since $\langle (a_\ell), (a_m) \rangle \Rightarrow (a_\ell \in q_k \Rightarrow a_m \in q_k$, for $\forall q_k$ in Q), we have $a_i, a_{i+1}, \ldots, a_{j-1}, a_j \in q_p$.

Case (ii): Let h visit all the direction-changer nodes before reaching (a_i). A similar argument as in case (i) leads us to the conclusion that $a_j, a_{j-1}, \ldots, a_{i-1}, a_i \in q_p$ when $a_i, a_j \in q_p$.

Case (iii): The direction-changer nodes are between (a_i) and (a_j) in h. Let (a_ℓ) and $(a_{\ell+k})$ be the starting and end nodes of the subpath of h that consists only of the direction-changer nodes (by lemma 4). Then the path h is

$$\langle (a_1), \ldots, (a_i), \ldots, \underbrace{(a_\ell), \ldots, (a_{\ell+k})}_{\text{direction-changers}}, \ldots, (a_j) \rangle$$

By case (i), all the elements that correspond to nodes between (a_i) and $(a_{\ell-1})$ in h belong to q_p and by case (ii), all the elements corresponding to nodes between $(a_{\ell+k+1})$ and (a_j) in h belong to q_p. The direction-changer nodes correspond to the elements of I which is a subset of all sets in Q. Hence we have the lemma.

(3) Either (a_i) or (a_j) is a direction-changer node. Suppose (a_i) is. Let (a_{i+k}) be the end node of the subpath of h that consists only of the direction-changer nodes. Then the path h is $\langle (a_1), \ldots, (a_i), \ldots, (a_{i+k}), (a_{i+k+1}), \ldots, (a_{j-1}), (a_j) \rangle$. By case (ii) of (2), $\{a_j, a_{j-1}, \ldots, a_{i+k+1}\} \subseteq q_p$. We know that $\{a_i, a_{i+1}, \ldots, a_{i+k}\} \subseteq I \subseteq q_p$. Hence, $\{a_i, a_{i+1}, \ldots, a_{j-1}, a_j\} \subseteq q_p$.

If (a_j) is a direction-changer node and (a_i) is not, a similar argument as above can be applied and the lemma proved. QED.

The theorem that follows gives the necessary and sufficient conditions for an intersection graph of a family Q, that is complete, to be a L.O. graph.

Theorem 5: Let $\Omega(Q)$ be complete. $\Omega(Q)$ is a L.O. graph iff there exists a Hamiltonian path in $\overline{G}(Q)$.

Proof: The sufficiency part of the theorem is easy to prove. We have $Q = \{q_1, q_2, \ldots q_m\}$ and $S = \{a_1, a_2, \ldots a_n\}$. Let h be a Hamiltonian path of $\overline{G}(Q)$. We can consider h as a n-tuple. Define a set of functions, $\{k_1, k_2, \ldots, k_n\}$, where k_i, $1 \leq i \leq n$, maps any n-tuple to the i^{th} member of the tuple, i.e. $k_i(\langle x_1, x_2, \ldots, x_i, \ldots, x_n \rangle) = x_i$.

Corresponding to Hamiltonian path h, let f_h be a 1-1 function that maps S into R s.t. for $\forall a_i \in S$, $f_h(a_i) = j$ where $k_j(h) = (a_i)$. (Note that f_h maps elements of S onto integers from 1 to n.)

Now, for $\forall q_i \in Q$, define $I_i = [\min_{a_p \in q_i} (f_h(a_p)), \max_{a_p \in q_i} (f_h(a_p))]$

It can be observed that $I_i \neq \phi$ and contains the images of all elements $\in q_i$. I_i, for $i = 1, 2, \ldots, m$, does not contain images of foreign elements w.r.t. q_i. To see this, suppose to the contrary that there exists an interval I_i containing images of foreign element(s) w.r.t. q_i. Then there exists $a_b, a_{c_1}, a_{c_2}, \ldots, a_{c_k}, \ldots, a_{c_j}, a_d$ belonging to S with $a_{c_k} \notin q_i$ and $f_h(a_{c_k})$ between $f_h(a_b)$ and $f_h(a_d)$ and $\{a_b, a_d\} \subseteq q_i$. This by the definition of f_h implies that a_{c_k} is between a_b and a_d which contradicts Lemma 6. Hence $(f_h; I_1, I_2, \ldots, I_m)$ implies that Q has the L.O. property.

Before we prove the necessity part of the theorem, we shall give an example to illustrate the above proof.

Example 8:

Let $q_1 = \{a_2, a_3, a_4, a_6\}$

$q_2 = \{a_1, a_2, a_3, a_6, a_5\}$

$q_3 = \{a_1, a_3, a_6\}$

and $q_4 = \{a_1, a_2, a_3, a_4, a_6\}$

we have $Q = \{q_1, q_2, q_3, q_4\}$

$S = \{a_1, a_2, a_3, a_4, a_5, a_6\}$

$I = \bigcap_{q_i \in Q} q_i = \{a_3, a_6\}$

The intersection graph of Q is given in figure 9. We note that $\Omega(Q)$ is complete. The DSG of Q is $\overline{G}(Q) = [S', \overline{R}, I']$ where $S' = \{ⓐ_1, ⓐ_2, ⓐ_3, ⓐ_4, ⓐ_5, ⓐ_6\}$ $I' = \{ⓐ_3, ⓐ_6\}$ and $\overline{R}$ is the connectivity relation. $\overline{G}(Q)$ is given in figure 10. $h = \langle ⓐ_4, ⓐ_2, ⓐ_3, ⓐ_6, ⓐ_1, ⓐ_5\rangle$ is a Hamiltonian path of $\overline{G}(Q)$ and is shown in solid lines in Fig. 10.

Define f_h: $f_h(a_4) = 1$ as $k_1(h) = ⓐ_4$

$f_h(a_2) = 2$ as $k_2(h) = ⓐ_2$

$f_h(a_3) = 3$ as $k_3(h) = ⓐ_3$

$f_h(a_6) = 4$ as $k_4(h) = ⓐ_6$

$f_h(a_1) = 5$ as $k_5(h) = ⓐ_1$

and $f_h(a_5) = 6$ as $k_6(h) = ⓐ_5$

we then define the intervals I_1, I_2, I_3 and I_4.

$$I_1 = [\min_{a_i \in q_1} (f_h(a_i)), \max_{a_i \in q_1} (f_h(a_i))]$$

$$= [f_h(a_4), f_h(a_6)] = [1, 4]$$

Similarly $I_2 = [2, 6]$, $I_3 = [3, 5]$ and $I_4 = [1, 5]$. The intervals are shown pictorially in figure 11.

To prove the necessity part of theorem 5, we need the following lemma which is a counterpart of lemma 4.

Lemma 7: Let $\Omega(Q)$ be a complete and L.O. graph. Let $I = \bigcap_{q_i \in Q} q_i$ and $(f; I_1, I_2, \ldots, I_m)$ imply the L.O. property of Q. Then there does not exist a_b, a_c, a_d s.t. $a_b, a_d \in I$ and $a_c \in (S-I)$ and $f(a_c)$ is between $f(a_b)$ and $f(a_d)$.

Proof: Assume to the contrary that there exist such a_b, a_c and a_d. Since $a_c \notin I$, there exists a set $q_i \in Q$ s.t. $a_c \notin q_i$. Since $\{a_b, a_c\} \subseteq I$, the interval I_i corresponding to q_i contains $f(a_b)$ and $f(a_d)$. If $f(a_c)$ is between $f(a_b)$ and $f(a_d)$, then I_i also contains $f(a_c)$. But a_c is foreign to q_i. This means that $(f; I_1, I_2, \ldots, I_m)$ does not imply the L.O. property of Q. Contradiction.

We are ready to prove the necessity part of theorem 5.

Proof: As $\Omega(Q)$ is a complete and L.O. graph, we have by lemma 3, $I = \bigcap_{q_i \in Q} q_i \neq \phi$. Let $I = \{a_p, a_{p+1}, \ldots, a_{p+\ell}\}$. There exist a function f_h and a set of intervals $\{I_1, I_2, \ldots, I_m\}$ s.t. $(f_h; I_1, I_2, \ldots, I_m)$ implies the linear ordering property of Q. For all $a_i, a_j \in S$, $i \neq j$, either $f_h(a_i) < f_h(a_j)$ or $f_h(a_j) < f_h(a_i)$. Then we can define a total ordering on the elements of S s.t. a_i precedes a_j iff $f_h(a_i) < f_h(a_j)$. By lemma 7, there does not exist a_b, a_c, a_d s.t. $a_b, a_d \in I$ and $a_c \in (S-I)$ and $f_h(a_c)$ is between $f_h(a_b)$ and $f_h(a_d)$. Without loss of generality we can assume that f_h is s.t.

$$f_h(a_1) < f_h(a_2) < \ldots < \underbrace{f_h(a_p) < f_h(a_{p+1}) < \ldots < f_h(a_{p+\ell})}_{\text{images of elements} \in I} < \ldots < f_h(a_n)$$

All the intervals contain the images of elements belonging to I. Hence, whenever an interval I_i contains $f_h(a_1)$ it has to contain $f_h(a_2)$, $f_h(a_3)$, ..., $f_h(a_{p+\ell})$. Then, $a_1 \in q_i \Rightarrow \{a_2, a_3, \ldots a_p, \ldots a_{p+\ell}\} \subseteq q_i$. For, if any of the elements $\in \{a_2, a_3, \ldots, a_{p-1}\}$ is foreign to q_i, we will have a contradiction that $(f_h; I_1, I_2, \ldots, I_m)$ does not imply the L.O. property of Q.

Thus we have $a_1 \overline{R} a_2$, $a_1 \overline{R} a_3$, ... $a_1 \overline{R} a_{p+\ell}$. In particular, $\langle \textcircled{a_1}, \textcircled{a_2} \rangle$ is an edge of $\overline{G}(Q)$. By considering intervals that contain $f_h(a_2)$, $f_h(a_3)$, ... $f_h(a_{p-1})$ and repeating the same argument as above, we see that $\langle \textcircled{a_2}, \textcircled{a_3} \rangle$, $\langle \textcircled{a_3}, \textcircled{a_4} \rangle$, ..., $\langle \textcircled{a_{p-1}}, \textcircled{a_p} \rangle$ are among the edges of $\overline{G}(Q)$.

A similar argument as above shows that $\langle \textcircled{a_n}, \textcircled{a_{n-1}} \rangle$, $\langle \textcircled{a_{n-1}}, \textcircled{a_{n-2}} \rangle$, ..., $\langle \textcircled{a_{p+\ell+1}}, \textcircled{a_{p+\ell}} \rangle$ are also edges of $\overline{G}(Q)$. Since the relation $\overline{R}$ is symmetric for I, every pair of nodes belonging to I' = $\{\textcircled{a_p}, \textcircled{a_{p+1}}, \ldots, \textcircled{a_{p+\ell}}\}$ is connected and directed both ways. But the nodes $\in$ I' are precisely the direction-changer nodes of $\overline{G}(Q)$. Hence $\langle \textcircled{a_1}, \textcircled{a_2}, \ldots, \textcircled{a_{p-1}}, \textcircled{a_p}, \ldots, \textcircled{a_{p+\ell}}, \textcircled{a_{p+\ell+1}}, \ldots, \textcircled{a_n} \rangle$ is a path of $\overline{G}(Q)$ which is Hamiltonian. <u>QED</u>.

Let $h = \langle \textcircled{a_1}, \textcircled{a_2}, \ldots, \textcircled{a_i}, \textcircled{a_{i+1}}, \ldots, \textcircled{a_n} \rangle$ be a Hamiltonian path in $\overline{G}(Q)$. We say that each $\textcircled{a_i}$ in h, for i = 2, 3, ..., (n-1), has both left and right neighbours. The left and right neighbours of $\textcircled{a_i}$ are $\textcircled{a_{i-1}}$ and $\textcircled{a_{i+1}}$ respectively. $\textcircled{a_1}$ has only the right neighbour namely $\textcircled{a_2}$ and $\textcircled{a_n}$ has only the left neighbour which is $\textcircled{a_{n-1}}$. The left neighbour of $\textcircled{a_1}$ is said to be empty and so is the right neighbour of $\textcircled{a_n}$.

Two paths P_1 and P_2 are equal (or non-distinct) iff the starting and end nodes of P_1 are the starting and end nodes of P_2 and for $\forall$ $a_i \in P_1$ s.t. (a_i) is not the starting node of P_1, the left neighbour of (a_i) in P_1 = the left neighbour of (a_i) in P_2 and for $\forall$ $a_i \in P_1$ s.t. (a_i) is not the end node of P_1, the right neighbour of (a_i) in P_1 = the right neighbour of (a_i) in P_2.

Let f be a 1-1 function that maps S into R. We know that f totally (linearly) orders the elements of S s.t. $\forall$ $a_i, a_j \in S$, $i \neq j$, a_i procedes a_j iff $f(a_i) < f(a_j)$. Let the linear ordering defined by f be O. If there exist intervals $I_1, I_2, \ldots, I_m$ s.t. $(f; I_2, \ldots, I_m)$ implies that Q has the L.O. property, then we say that the linear ordering O implies the L.O. property of Q. The following assertion is stronger than theorem 5, but the proof is essentially the same.

Theorem 6: Every distinct linear ordering of the elements $\in$ S, which implies the L.O. property of Q, corresponds to a distinct Hamiltonian path in $\overline{G}(Q)$ and vice versa when $\Omega(Q)$ is complete.

Proof: In the proof of the necessity part of theorem 5, we observe that the function f_h defines a linear ordering, say O, of the elements of S. We found a Hamiltonian path in $\overline{G}(Q)$ that corresponded to O.

Let $(f_{h'}; I_1', I_2', \ldots, I_m')$ satisfy the L.O. property of Q and $f_{h'}$ be different from f_h. $f_{h'}$ then gives a total ordering O' different from O. Applying the same arguments as in theorem 5, we get a Hamiltonian path h' corresponding to O'. h' is different from h.

In the proof of the only-if part of theorem 5, we defined a function f_h and intervals $I_1, I_2, \ldots, I_m$ corresponding to a Hamiltonian path h of

$\overline{G}(Q)$ s.t. $(f_h; I_1, I_2, \ldots, I_m)$ implied the L.O. property of Q. Any other Hamiltonian path h' would have resulted in a function $f_{h'}$, $f_{h'} \neq f_h$, and a set of intervals $I'_1, I'_2, \ldots, I'_m$ s.t. $(f_{h'}; I'_1, I'_2, \ldots I'_m)$ implied the L.O. property of Q. Since f_h and $f_{h'}$ are distinct, the linear orderings defined by them are distinct. QED.

Lemma 8: If $h = \langle \textcircled{a_1}, \textcircled{a_2}, \ldots, \textcircled{a_i}, \textcircled{a_{i+1}}, \ldots, \textcircled{a_n} \rangle$ is a Hamiltonian path in $\overline{G}(Q)$, then $h^R = \langle \textcircled{a_n}, \textcircled{a_{n-1}}, \ldots, \textcircled{a_{i+1}}, \textcircled{a_i}, \ldots, \textcircled{a_1} \rangle$ is a Hamiltonian path of $\overline{G}(Q)$.

Proof: Since there exists a Hamiltonian path in $\overline{G}(Q)$, by Lemma 5 $I = \bigcap_{q_i \in Q} q_i \neq \phi$. By the direction changing property of the nodes corresponding to I and by Lemma 4, we have that h^R is a Hamiltonian path of $\overline{G}(Q)$.

Corollary 3: Let $\Omega(Q)$ be complete. Then $\Omega(Q)$ is a L.O. graph iff there exist at least two Hamiltonian paths in $\overline{G}(Q)$.

Section 4: Union of Linearly Orderable Families.

In the sequel, Q_1 and Q_2 denote two distinct families of sets. $Q_1 \cap Q_2$ need not be empty. $\overline{G}(Q_1)$ and $\overline{G}(Q_2)$ represent the DSG of Q_1 and Q_2 respectively. S_1 denotes the set $\bigcup_{\substack{a_i \in q_j \\ q_j \in Q_1}} a_i$ and S_2 the set $\bigcup_{\substack{a_i \in q_j \\ q_j \in Q_2}} a_i$.

$\tilde{S}$ indicates $(S_1 \cup S_2)$.

Lemma 9: Let $\Omega(Q_1 \cup Q_2)$ be a L.O. graph and $\Omega(Q_1)$, $\Omega(Q_2)$ be complete. Let $I = S_1 \cap S_2$ and f be a function that defines a linear ordering of the elements $\in \tilde{S}$ implying that $(Q_1 \cup Q_2)$ has L.O. property. Then there does not exist a_p, a_i, a_j s.t. $a_p \in (\tilde{S}-I)$ and $\{a_i, a_j\} \subseteq I$ and $f(a_p)$ is between $f(a_i)$ and $f(a_j)$.

Proof: By contradiction. Suppose that there exist such a_p, a_i and a_j. Without loss of generality, we shall assume that $f(a_i) < f(a_p) < (a_j)$.

Since $\Omega(Q_1 \cup Q_2)$ is a L.O. graph, by Lemma 1 $\Omega(Q_1)$ and $\Omega(Q_2)$ are L.O. graphs. Then by Lemma 3, $I_1 = \bigcap_{q_i \in Q_1} q_i \neq \phi$ and $I_2 = \bigcap_{q_i \in Q_2} q_i \neq \phi$.

As a_i, $a_j \in S_1$, there exist sets q_c, $q_d \in Q_1$ s.t. $a_i \in q_c$ and $a_j \in q_d$. $I_1 \subseteq q_c$ and $I_1 \subseteq q_d$.

Case (i) $a_i \in I_1$. Then the interval I_d^1 corresponding to q_d contains $f(a_i)$, $f(a_j)$ and hence $f(a_p)$. Since the linear ordering defined by f implies the L.O. property of $Q_1 \cup Q_2$, a_p is not foreign to q_d.

Case (ii) $a_j \in I_1$. By the same arguments as in case (i), we have $a_p \in q_c$.

Case (iii) If a_i, $a_j \notin I_1$ then there exists an element $a_\ell \in I_1$ s.t. either $f(a_\ell) < f(a_i)$ or $f(a_i) < f(a_\ell)$. In either case, $a_p \in q_c$ or q_d. For, if a_p is foreign to both q_c and q_d, we will have a contradiction that the linear ordering defined by f does not imply the L.O. property of $(Q_1 \cup Q_2)$.

By cases (i), (ii) and (iii), we see that there exists a $q_i \in Q_1$ s.t. $a_p \in q_i$. Then, $a_p \in S_1$. By similar arguments, we can prove that $a_p \in S_2$. This means that $a_p \in S_1 \cap S_2 = I$. Contradiction. QED.

Lemma 10: Let $\Omega(Q_1 \cup Q_2)$ be a L.O. graph and $\Omega(Q_1)$, $\Omega(Q_2)$ be complete. Let $I = S_1 \cap S_2 \neq \phi$, S_1 or S_2. Let f define a linear ordering of the elements $\in \tilde{S}$ implying the L.O. property of $(Q_1 \cup Q_2)$. Let a_c and a_k be s.t. $f(a_c) = \underset{a_i \in I}{\mathrm{Min}} f(a_i)$ and $f(a_k) = \underset{a_i \in I}{\mathrm{Max}} f(a_i)$. Then

(i) for $\forall a_i \in (S_1 - I)$ either $f(a_i) < f(a_c)$ or $f(a_i) > f(a_k)$
i.e. there does not exist $a_p, a_q \in (S_1 - I)$ s.t. $f(a_p) < f(a_c)$ and $f(a_q) > f(a_k)$.

(ii) $\forall a_i \in (S_1 - I), f(a_i) < f(a_c) \Leftrightarrow \forall a_j \in (S_2 - I), f(a_j) > f(a_k)$

<u>Proof</u>: We shall prove the lemma by contradiction.

<u>Part (i)</u> Suppose there exist $a_p, a_q \in (S_1 - I)$ s.t. $f(a_p) < f(a_c)$ and $f(a_q) > f(a_k)$. By Lemma 1 and Lemma 3 we have that $I_1 = \bigcap_{q_i \in Q_1^1} q_i \neq \phi$ and $I_2 = \bigcap_{\bar{q}_i \in Q_2^1} \bar{q}_i \neq \phi$. For $\forall \bar{q}_i \in Q_2$, $\bar{q}_i \cap (S_1 - I) = \phi$ and $\bar{q}_i \supseteq I_2$. Hence $I_2 \cap (S_1 - I) = \phi$. Similarly $I_1 \cap (S_2 - I) = \phi$. As $a_k \in S_2$, there exists a $\bar{q}_i \in Q_2$ s.t. $\bar{q}_i \supseteq \{a_k\} \cup I_2$. Since $a_p, a_q \notin \bar{q}_\ell$ for $\forall \bar{q}_\ell \in Q_2$ and the linear ordering defined by f implies the L.O. property of $(Q_1 \cup Q_2)$, we have that $f(a_p) < f(a_j) < f(a_q)$ for all $a_j \in I_2$. This further implies that for $\forall a_\ell \in S_2$, $f(a_\ell)$ is between $f(a_p)$ and $f(a_q)$. Thus we have:

$$\ldots < f(a_p) < \ldots \quad \underbrace{\ldots < f(a_c) < \ldots\ldots < f(a_k)}_{\text{images of elements } \in I \text{ (By Lemma 9)}} < \ldots\ldots \ldots < f(a_q) < \ldots$$

←This interval contains images of all elements $\in S_2$→

images of elements $\in$ I
(By Lemma 9)

Now there exist $q_i, q_j \in Q_1$ s.t. $q_i \supseteq \{a_p\} \cup I_1$ and $q_j \supseteq \{a_q\} \cup I_1$. Since $I_1 \cap (S_2 - I) = \phi$, the interval corresponding to q_i or the interval corresponding to q_j contains images of elements $\in (S_2 - I)$ which are foreign to all sets in Q_1. This means that the linear ordering defined by f does not imply L.O. property of $(Q_1 \cup Q_2)$. Contradiction.

Part (ii) $\Rightarrow$. We have $f(a_i) < f(a_c)$ for $\forall\, a_i \in (S_1-I)$. Assume to the contrary that there exists an $a_j \in (S_2-I)$ s.t. $f(a_j) < f(a_k)$. By Lemma 9, $f(a_j) < f(a_c)$. Since a_j is foreign to all sets containing elements $\in (S_1-I)$, we have $f(a_j) < \min_{a_r \in S_1} (f(a_r))$. There exist a set $\bar{q}_i \in Q_2$ s.t. $\bar{q}_i \supseteq \{a_j\} \cup I_2$. As $\bar{q}_i \cap (S_1-I) = \phi$, $f(a_\ell) < \min_{a_r \in S_1} (f(a_r))$ for $\forall\, a_\ell \in I_2$. Now consider the set $\bar{q}_p \in Q_2$ s.t. $\bar{q}_p \supseteq \{a_k\} \cup I_2$. The interval corresponding to $\bar{q}_p$ contains the images of elements $\in (S_1-I)$ which are foreign to all sets in Q_2. This leads to a contradiction.

$\Leftarrow$. The same arguments as above direct us to the conclusion that if for $\forall\, a_j \in (S_2-I)$, $f(a_j) > f(a_k)$, then for $\forall\, a_j \in (S_1-I)$, $f(a_j) < f(a_c)$. QED.

Let h_1 and h_2 be Hamiltonian paths in $\bar{G}(Q_1)$ and $\bar{G}(Q_2)$ respectively. Let $I = S_1 \cap S_2$. We see that h_1 induces a subpath in the set of nodes that correspond to I in $\bar{G}(Q_1)$. The starting and end nodes of this subpath are nodes that correspond to some elements in I and the subpath contains all the nodes which correspond to the elements $\in$ I. Let h_1^I denote this subpath. Similarly h_2^I is the subpath induced by h_2 in the set of nodes that correspond to I in $\bar{G}(Q_2)$. We say that the Hamiltonian paths h_1 and h_2 are consistent ($h_2 \sim h_2$) iff exactly one of the following holds:

(i) $S_1 \cap S_2 = I = \phi$

(ii) $h_1^I = h_2^I$ and the left neighbour of the starting node of h_1^I is empty in h_1 or h_2 and the right neighbour of the end node of h_1^I is empty in h_1 or h_2.

Example 9: $h_1 = \langle \textcircled{a_1}, \textcircled{a_2}, \textcircled{a_3}, \textcircled{a_4}, \textcircled{a_5} \rangle$

$h_2 = \langle \textcircled{a_3}, \textcircled{a_4}, \textcircled{a_5}, \textcircled{a_6}, \textcircled{a_7}, \textcircled{a_8} \rangle$

$S_1 = \{a_1, a_2, a_3, a_4, a_5\}$

$S_2 = \{a_3, a_4, a_5, a_6, a_7, a_8\}$

$I = S_1 \cap S_2 = \{a_3, a_4, a_5\}$

$h_1^I = \langle \textcircled{a_3}, \textcircled{a_4}, \textcircled{a_5} \rangle = h_2^I$

The starting node of $h_1^I = \textcircled{a_3}$ and the end node of $h_1^I = \textcircled{a_5}$. The left neighbour of $\textcircled{a_3}$ is empty in h_2 and the right neighbour of $\textcircled{a_5}$ is empty in h_1. Hence $h_1 \sim h_2$.

Theorem 7: Let $\Omega(Q_1)$ and $\Omega(Q_2)$ be complete. If $\Omega(Q_1 \cup Q_2)$ is a L.O. graph, then there exist Hamiltonian paths h_1 in $\overline{G}(Q_1)$ and h_2 in $\overline{G}(Q_2)$ s.t. h_1 and h_2 are consistent.

Proof:

By Lemma 1, $\Omega(Q_1)$ and $\Omega(Q_2)$ are L.O. graphs. Since $\Omega(Q_1)$ and $\Omega(Q_2)$ are also complete, by theorem 5 there exist Hamiltonian paths in $\overline{G}(Q_1)$ and in $\overline{G}(Q_2)$.

Case (i): $I = \phi$. Any Hamiltonian path in $\overline{G}(Q_1)$ is consistent with any Hamiltonian path in $\overline{G}(Q_2)$.

Case (ii): $I \neq \phi$. Let f be a function that defines a linear ordering of the elements $\in \tilde{S}$ implying the L.O. property of $Q_1 \cup Q_2$. We have the following situations:

<u>(1)</u> $I = S_2$, i.e. $S_2 \subseteq S_1$ and $\tilde{S} = S_1$

Define $f_1(a_i) = f(a_i) \quad \forall\ a_i \in S_1$

$f_2(a_i) = f(a_i) \quad \forall\ a_i \in S_2$

Clearly, f_1 defines a linear ordering, say O_1, of the elements $\in S_1$ which implies the L.O. property of Q_1. So does f_2 w.r.t. Q_2. Let the linear ordering defined by f_2 be O_2. Since $f_2(a_i) = f_1(a_i)$ for $\forall\ a_i \in S_2$, we have that, for $\forall\ a_\ell$, $a_k \in S_2$, a_ℓ precedes a_k in O_2 iff a_ℓ precedes a_k in O_1. By Lemma 9, there does not exist an $a_p \in (S_1-S_2)$ and a_c, $a_d \in S_2$ s.t. $f_1(a_p)$ is in between $f_1(a_c)$ and $f_1(a_d)$. Let h_1 and h_2 be the Hamiltonian paths in $\overline{G}(Q_1)$ and $\overline{G}(Q_2)$ corresponding to O_1 and O_2 respectively (see proof of theorem 5). Then $h_2 = h_2^I = h_1^I$. The left neighbour of the starting node of h_2 and the right neighbour of the end node of h_2 are both empty. Then $h_1 \sim h_2$.

<u>(2)</u> $I = S_1$, i.e. $S_1 \subseteq S_2$. The proof is similar to that of (1), above.

<u>(3)</u> $I \neq \phi$ or S_1 or S_2. Let $I = \{a_1, a_2, \ldots, a_k\}$. Without loss of generality, we can let f be s.t. $f(a_1) < f(a_2) < \ldots < f(a_k)$. By Lemma 10, we can assume without any loss in generality that for $\forall\ a_i \in (S_1-I)$, $f(a_i) < f(a_1)$ and $\forall\ a_i \in (S_2-I)$, $f(a_i) > f(a_k)$.

Now, define f_1, f_2:

$f_1(a_i) = f(a_i) \quad \forall\ a_i \in S_1$

$f_2(a_i) = f(a_i) \quad \forall\ a_i \in S_2$

Clearly, f_1 and f_2 define linear orderings, say O_1 and O_2, of the elements belonging to S_1 and S_2 respectively s.t. the L.O. property of Q_1 and Q_2 are implied. Let h_1 and h_2 be the Hamiltonian paths corresponding to O_1 and O_2 in $\overline{G}(Q_1)$ and $\overline{G}(Q_2)$ respectively (see the proof of theorem 5). Then $h_1^I = \langle (a_1), (a_2), \cdots (a_k) \rangle = h_2^I$. (a_1) is the starting node of h_1^I and its left neighbour is empty in h_2. The right neighbour of (a_k), the end node of h_1^I, is empty in h_1. We thus have $h_1 \sim h_2$. QED

Let $Q = \{Q_1, Q_2, \ldots, Q_m\}$ be a set of distinct families of sets with $Q_i \cap Q_j$ not necessarily empty. For $1 \leq i \leq m$, let $\Omega(Q_i)$ be complete and S_i denote the set $\bigcup_{\substack{a_j \in q \\ q \in Q_i}} a_j$. $\tilde{S}$ indicates $\bigcup_{i=1,2,\ldots,m} S_i =$ $\{a_1, a_2, \ldots, a_n\}$. Let $h_1, h_2, \ldots, h_m$ be pair-wise consistent Hamiltonian paths in $\overline{G}(Q_1), \overline{G}(Q_2), \ldots, \overline{G}(Q_m)$ respectively where $\overline{G}(Q_i)$ is the DSG of Q_i. Let $O_1, O_2, \ldots, O_m$ be the linear orderings corresponding to $h_1, h_2, \ldots, h_m$. Define a directed graph $\tilde{G}(Q) = [\tilde{S}', \tilde{R}]$. $\tilde{S}'$ is the set of nodes of $\tilde{G}(Q)$ and is $\{(a_1), (a_2), \ldots, (a_i), \ldots, (a_n)\}$, where node (a_i) corresponds to element $a_i \in \tilde{S}$ and vice versa. $\tilde{R}$ is an irreflexive relation defined on $\tilde{S}$ s.t. for $\forall\, a_i, a_j \in \tilde{S}$, $i \neq j$, $a_i \tilde{R} a_j$ iff there exists a linear ordering O_k, $1 \leq k \leq m$, in which a_i precedes a_j. $(a_i) \tilde{R} (a_j)$ iff $a_i \tilde{R} a_j$. We use the same symbol $\tilde{R}$ for a relation between two elements of $\tilde{S}$ and two elements of $\tilde{S}'$ since there is no confusion. $\tilde{R}$ is the connectivity relation of $\tilde{G}(Q)$ and $(a_i) \tilde{R} (a_j) \Leftrightarrow$ $\langle (a_i), (a_j) \rangle$ is an edge of $\tilde{G}(Q)$. $\tilde{G}(Q)$ is called a Partial Order (P.O.) graph of Q corresponding to $Q_1, Q_2, \ldots, Q_m$, defined by O_1, $O_2, \ldots, O_m$ or $h_1, \ldots, h_m$.

An undirected path or simply a path in a directed graph G is a sequence of distinct nodes $(a_1), (a_2), \ldots, (a_k)$ s.t. for $i = 1, 2, \ldots, (k-1)$,

(a_i , a_{i+1}) are edges of G. Note that we ignore the direction of the edges in G. A connected-directed graph is a directed graph in which there is a path between every pair of distinct nodes. A component G' of a directed graph G is a subgraph of G s.t. G' is a connected-directed graph and is not properly contained in any other connected-directed subgraph of G.

A directed-cycle C in a directed graph G is a sequence of distinct nodes a_1, a_2, ..., a_k s.t. for i = 1, 2, ..., (k-1), $\langle a_i, a_{i+1} \rangle$ and $\langle a_k, a_1 \rangle$ are edges of G. The length of a cycle is the number of nodes in the cycle.

Example 10: Consider the graph G in Fig. 12. It has two components. If R is the connectivity relation of G, then the components are $[\{a_1, a_2, a_5, a_6\}, R]$ and $[\{a_3, a_4\}, R]$. a_1, a_2, a_5 is a directed cycle of G and is of length 3.

Let G be an undirected graph. $G_1, G_2, \ldots, G_m$ be a set of complete subgraphs of G(i.e. $G_1, G_2, \ldots G_m$ are subgraphs of G and are complete) s.t. every node and edge of G is in at least one of them. Then G is said to be covered by $G_1, G_2, \ldots G_m$. This is exemplified in example 11. We are now ready to prove a result concerning the L.O. property of an arbitrary family of sets.

Theorem 8: Let $G_1, G_2, \ldots G_m$ be a set of complete subgraphs of $\Omega(Q)$ that cover $\Omega(Q)$. Let $Q_i \subseteq Q$ be s.t. $G_i = \Omega(Q_i)$ for $1 \le i \le m$. $\Omega(Q)$ is a L.O. graph iff there exists at least one P.O. graph $\tilde{G}(Q)$ corresponding to Q_1, $Q_2, \ldots, Q_m$ and any $\tilde{G}(Q)$ is directed-cycle free.

Proof: The if-part of the theorem: We first show that there exists a P.O. graph $\tilde{G}(Q)$ of Q. Let $\overline{G}(Q_i)$ be the directed semantic graph of G_i which is $\Omega(Q_i)$. If $\overline{G}(Q_i)$ does not have a Hamiltonian path, then $\Omega(Q_i)$ is not a L.O. graph. By Lemma 1, this implies that $\Omega(Q)$ is not a L.O. graph which is a contradiction. Hence every one of $\overline{G}(Q_i)$, $i = 1, 2, \ldots, m$, has a Hamiltonian path. If $\overline{G}(Q_i)$ and $\overline{G}(Q_j)$, $i \neq j$, does not have Hamiltonian paths that are consistent, then by theorem 7, $\Omega(Q_i \cup Q_j)$ is not a L.O. graph. Again by Lemma 1, Q is not linearly orderable which is not true. Hence there exist Hamiltonian paths $h_1, h_2, \ldots, h_m$ in $\overline{G}(Q_1), \overline{G}(Q_2), \ldots, \overline{G}(Q_m)$ s.t. they are pair-wise consistent, i.e. a $\tilde{G}(Q)$ exists.

Suppose to the contrary that there exists a $\tilde{G}(Q)$ containing directed cycles. Let $C = \textcircled{a_1}, \textcircled{a_2}, \ldots, \textcircled{a_i}, \textcircled{a_{i+1}}, \ldots, \textcircled{a_k}$ be a directed cycle of minimum length in $\tilde{G}(Q)$.

If the length of C is 2, then $\langle \textcircled{a_1}, \textcircled{a_2} \rangle$ and $\langle \textcircled{a_2}, \textcircled{a_1} \rangle$ are edges of $\tilde{G}(Q)$. This implies that there exist linear orderings O_i, O_j among the linear orderings used to construct $\tilde{G}(Q)$, s.t. $i \neq j$ and a_1 precedes a_2 in O_i and a_2 precedes a_1 in O_j. Then Hamiltonian paths h_i and h_j which define O_i and O_j are not consistent. Contradiction.

Then let the length of $C \geq 3$. Let $S_\ell = \bigcup_{\substack{a_j \in q_i \\ q_i \in Q_\ell}} a_j$ for $1 \leq \ell \leq m$ and

$\tilde{S} = \bigcup_{\ell=1,2,\ldots,m} S_\ell$. Note that for $\forall\, a_i, a_j \in S_\ell$, $i \neq j$, $1 \leq \ell \leq m$, either $\langle \textcircled{a_i}, \textcircled{a_j} \rangle$ or $\langle \textcircled{a_j}, \textcircled{a_i} \rangle$ is an edge of $\tilde{G}(Q)$. Since C is of minimum length a_p and a_q do not simultaneously belong to S_ℓ for $1 \leq p, q \leq k$, $q \neq p+1$ or $p-1$ and $1 \leq \ell \leq m$. (see figure 13.) Further for $1 \leq i \leq k$, $1 \leq \ell \leq m$,

$a_i, a_{i+1} \in S_\ell \Rightarrow a_j, a_{j+1} \notin S_\ell$

where $j \neq i$, and $i+1$, $j+1$ are modulo k. Hence w.ℓ.g. we can assume that $a_i, a_{i+1} \in S_i$ for $i = 1, 2, \ldots, (k-1)$ and $a_k, a_1 \in S_k$.

Since $\Omega(Q)$ is a L.O. graph, there exists a function f that defines a linear ordering of the elements $\in \tilde{S}$ implying the L.O. property of Q. Consider $\Omega(Q_i)$. Since $a_1, a_2, \ldots, a_{i-1}, a_{i+2}, \ldots, a_k$ are foreign to all sets in Q_i and $a_i, a_{i+1} \in S_i$, $f(a_\ell)$ is not between $f(a_i)$ and $f(a_{i+1})$ for $\ell = 1, 2, \ldots, i-1, i+2, \ldots, k$. This can be seen by similar arguments as in Lemma 9.

Applying the above contention to $\Omega(Q_1), \Omega(Q_2), \ldots, \Omega(Q_{k-1})$ and $\Omega(Q_k)$, we get a contradiction that $f(a_3), f(a_4), \ldots, f(a_{k-1})$ are between $f(a_1)$ and $f(a_k)$ and $f(a_3), f(a_4), \ldots, f(a_{k-1})$ are not between $f(a_k)$ and $f(a_1)$.

To prove the sufficiency of the conditions, we shall show how to construct for any graph $\Omega(Q)$ satisfying the conditions, a function f and a set of intervals implying the L.O. property of Q.

Let $\tilde{G}_1\ \tilde{G}_2, \ldots, \tilde{G}_p$ be the components of $\tilde{G}(Q)$. Define function f:

(i) for $\forall$ (a_i), (a_j) $\in \tilde{G}_k$, $1 \leq k \leq p$, $f(a_i) < f(a_j)$ iff there exists a directed path from (a_i) to (a_j), i.e. $\langle (a_i), (a_j) \rangle$ is an edge of $\tilde{G}_k$ or there exist $(a_{d_1}), (a_{d_2}), \ldots, (a_{d_k}), (a_{d_{k+1}}), \ldots, (a_{d_p}) \in \tilde{G}_k$ s.t. $\langle (a_i), (a_{d_1}) \rangle, \langle (a_{d_1}), (a_{d_2}) \rangle, \ldots, \langle (a_{d_p}), (a_j) \rangle$ are edges of $\tilde{G}_k$.

(ii) for $\forall$ $a_i \in \tilde{G}_k$ and $\forall$ $a_j \in \tilde{G}_\ell$, $k < \ell$, $f(a_i) < f(a_j)$. Since $\tilde{G}(Q)$ is directed cycle free, such a function exists. For each $q_i \in Q$, define $I_i = [\min_{a_j \in q_i} (f(a_j)), \max_{a_j \in q_i} (f(a_j))]$. Interval I_i corresponds to q_i.

To see that I_i does not contain images of any foreign elements w.r.t. q_i, we suppose to the contrary that it does and show that it leads to a contradiction. Let there exist $a_b, a_d \in q_i, a_{c_1}, a_{c_2}, \ldots, a_{c_j} \in (\tilde{S}-q_i)$ s.t. for $1 \leq k \leq j$, $f(a_{c_k})$ is between $f(a_b)$ and $f(a_d)$. Further let a_b, a_d be s.t. there does not exist $a_p \in q_i$ and $f(a_p)$ is between $f(a_b)$ and $f(a_d)$. W.ℓ.g., we can assume that $f(a_b) < f(a_{c_1}) <\ldots< f(a_{c_j}) < f(a_d)$. Then, $\langle \textcircled{a_b}, \textcircled{a_{c_1}} \rangle$ is an edge of $\tilde{G}(Q)$ and $\textcircled{a_{c_1}}$ is the right neighbour of $\textcircled{a_b}$ in some Hamiltonian path h_t, used to define $\tilde{G}(Q)$.

q_i belongs to at least one complete subgraph, say G_ℓ, that was chosen to cover $\Omega(Q)$. Let Q_ℓ be s.t. $\Omega(Q_\ell) = G_\ell$ and h_ℓ be the Hamiltonian path in $\bar{G}(Q_\ell)$ that was used in the definition of $\tilde{G}(Q)$. If $\textcircled{a_d}$ precedes $\textcircled{a_b}$ in h_ℓ, then $\langle \textcircled{a_d}, \textcircled{a_b} \rangle$ will be an edge of $\tilde{G}(Q)$ and hence $\textcircled{a_b}, \textcircled{a_{c_1}}, \ldots, \textcircled{a_{c_k}}, \ldots, \textcircled{a_{c_j}}, \textcircled{a_d}$ will be a cycle of $\tilde{G}(Q)$ which is not possible. Hence, let $\textcircled{a_b}$ precede $\textcircled{a_d}$ in h_ℓ. Since a_{c_1} is foreign to q_i, by Lemma 6 $\textcircled{a_{c_1}}$ is not between $\textcircled{a_b}$ and $\textcircled{a_d}$ in h_ℓ. Hence $h_\ell \neq h_t$.

The right neighbour of $\textcircled{a_b}$ is not empty in both h_ℓ and h_t and the right neighbour of $\textcircled{a_b}$ in $h_t = \textcircled{a_{c_1}}$ which is not the right neighbour of $\textcircled{a_b}$ in h_ℓ since $\textcircled{a_b}$ precedes $\textcircled{a_d}$ in h_ℓ and $\textcircled{a_{c_1}}$ is not between $\textcircled{a_b}$ and $\textcircled{a_d}$ in h_ℓ. This leads us to the contradiction that h_ℓ and h_t are not consistent. QED.

<u>Example 11</u>:

Let $q_1 = \{a_1, a_2, a_3\}$

$q_2 = \{a_2, a_3, a_4, a_5\}$

$q_3 = \{a_2, a_3, a_4\}$

$q_4 = \{a_4, a_5, a_6\}$

$q_5 = \{a_4, a_5, a_6, b_1\}$

$q_6 = \{b_1, b_2\}$

$q_7 = \{a_7, a_8\}$

and $q_8 = \{a_7, a_9\}$

$Q = \{q_1, q_2, q_3, q_4, q_5, q_6, q_7, q_8\}$.

$\Omega(Q)$ is given in figure 14. Let R be the connectivity relation of $\Omega(Q)$

Let $G_1 = [\{ⓠ_1, ⓠ_2, ⓠ_3\}, R]$

$G_2 = [\{ⓠ_2, ⓠ_3, ⓠ_4, ⓠ_5\}, R]$

$G_3 = [\{ⓠ_5, ⓠ_6\}, R]$

and $G_4 = [\{ⓠ_7, ⓠ_8\}, R]$

G_1, G_2, G_3, G_4 are complete subgraphs of $\Omega(Q)$ that cover $\Omega(Q)$. We have $Q_1 = \{q_1, q_2, q_3\}$, $Q_2 = \{q_2, q_3, q_4, q_5\}$, $Q_3 = \{q_5, q_6\}$ and $Q_4 = \{q_7, q_8\}$. $\overline{G}(Q_1)$, $\overline{G}(Q_2)$, $\overline{G}(Q_3)$ and $\overline{G}(Q_4)$ are given in figures 15, 16, 17 and 18 respectively.

$h_1 = \langle ⓐ_1, ⓐ_2, ⓐ_3, ⓐ_4, ⓐ_5 \rangle$

$h_2 = \langle ⓐ_2, ⓐ_3, ⓐ_4, ⓐ_5, ⓐ_6, ⓑ_1 \rangle$

$h_3 = \langle a_4, a_5, a_6, b_1, b_2 \rangle$

$h_4 = \langle a_8, a_7, a_9 \rangle$

h_1, h_2, h_3, h_4 are pair wise consistent Hamiltonian paths in $\bar{G}(Q_1)$, $\bar{G}(Q_2)$, $\bar{G}(Q_3)$ and $\bar{G}(Q_4)$ respectively. Let h_1, h_2, h_3 and h_4 define $\tilde{G}(Q)$. See figure 19. For the sake of clarity, we have not shown in fig. 19 all the edges of $\tilde{G}(Q)$ which is directed-cycle free. Hence Q is linearly orderable. Let $\tilde{R}$ be the connectivity relation of $\tilde{G}(Q)$. We note that there are two components of $\tilde{G}(Q)$, namely $\tilde{G}_1$ and $\tilde{G}_2$ where $\tilde{G}_1 = [\{a_1, a_2, a_3, a_4, a_5, a_5, b_1, b_2\}, \tilde{R}]$ and $\tilde{G}_2 = [\{a_7, a_8, a_9\}, \tilde{R}]$. Let f be: $f(a_1) = 1$, $f(a_2) = 2$, $f(a_3) = 3$, $f(a_4) = 4$, $f(a_5) = 5$, $f(a_6) = 6$, $f(b_1) = 7$, $f(b_2) = 8$, $f(a_8) = 9$, $f(a_7) = 10$ and $f(a_9) = 11$. The intervals corresponding to q_i, $1 \leq i \leq 8$ are shown in figure 20.

<u>Lemma 11</u>: Let G_1 and G_2 be complete subgraphs of $\Omega(Q)$ s.t. G_1 and G_2 cover $\Omega(Q)$. Let $Q_1 \subseteq Q$, $Q_2 \subseteq Q$ be s.t. $\Omega(Q_1) = G_1$ and $\Omega(Q_2) = G_2$. Let h_1 and h_2 be consistent Hamiltonian paths in $\bar{G}(Q_1)$ and $\bar{G}(Q_2)$. Then $\tilde{G}(Q)$ defined by h_1 and h_2 is directed-cycle free.

<u>Proof</u>: Suppose that $\tilde{G}(Q)$ is not directed-cycle free. Let $C = a_1, a_2, \ldots a_k$ be a cycle of minimum length in $\tilde{G}(Q)$. First, we observe that the length of $C \leq 3$. This can be seen by arguments similar to the ones in Theorem 8. If the length of C is 2, then $\langle a_1, a_2 \rangle$, $\langle a_2, a_1 \rangle$ are edges of $\tilde{G}(Q)$. This implies that a_1 precedes a_2 in $h_2(h_1)$ and a_2 precedes a_1 in $h_1(h_2)$. We then have a contradiction that h_1 and h_2 are not consistent. If the length of C is 3, then in $h_1(h_2)$

(a_1) precedes (a_2) and (a_3) and (a_2) precedes (a_3). In $h_2(h_1)$, (a_3) precedes (a_1). This also leads to the contradiction that h_1 is not consistent with h_2. QED

Theorem 9: If G_1 and G_2 are complete subgraphs of $\Omega(Q)$ s.t. G_1 and G_2 cover $\Omega(Q)$, then $\Omega(Q)$ is a L.O. graph iff there exist consistent Hamiltonian paths in $\overline{G}(Q_1)$ and $\overline{G}(Q_2)$ where $Q_1 \subseteq Q$, $Q_2 \subseteq Q$ and $\Omega(Q_1) = G_1$, $\Omega(Q_2) = G_2$.

Proof: The necessity is by theorem 7 and the sufficiency by theorem 8 and lemma 11.

Section 5: An Application of the L.O. Property to Information Retrieval.

We shall consider a particular form of file organization in the field of data management. A record in a data structure is a set of elements where each element is a two tuple, called a field. The first member of the tuple is called an attribute and the second its value. A file is a collection of records. Normally a field (or a combination of fields) of a given file uniquely identifies each record of the file. Such a field is called a primary field and the attribute of a primary field a primary key. Fields other than the primary field(s) are called data fields.

File organization is the arrangement of the records of a file F on a storage medium S so that a family of querries Q can be answered. A storage medium S is called "linear" if the storage locations of S can be arranged linearly and the access time between any two storage locations is an increasing function of the distance between them. Tapes and tracks of a disk are some examples of linear storage media. We restrict our attention to file organizations with linear storage media.

When the querries belonging to Q are based on a primary key, a simple hash coding scheme constitutes a good form of file organization. When the querries are based on data fields, one resorts to inverted (and/or a multilist) file organization. If the questions are related to only one field f_i, then all records in which the attribute of f_i takes a particular value can be stored in consecutive storage locations so that minimum retrieval time is guaranteed for each querry. However if the querries relate to more than one field (primary or data), then to achieve minimum retrieval time records may have to be stored redundantly, i.e. a record is stored more than once.

Suppose that a querry family Q is such that there exists a 1-1 function f which maps the records belonging to the file F into storage locations of a linear storage medium satisfying (i) for each querry $q_i \in Q$, there exists a sequence S_i of consecutive storage locations containing all records pertinent to q_i and (ii) S_i does not contain any record not pertinent to q_i. We, then, say that the family of querries Q has the Consecutive Retrieval property (C.R. property) [4], which in other words means that Q is linearly orderable. A file organization having this property is called a C.R. organization. Note that a C.R. organization precludes redundant storage of records. By knowing the first and the last pertinent records of a querry in a C.R. organization, all relevant records of all querries can be retrieved in minimum time.

Our results regarding the L.O. property of a family of sets (i.e. the C.R. property of a family of querries where a querry is a set of reply records) do not require any restrictions on the family of sets and are most general in nature. Till this memo, there have been no results to check if an arbitrary family of querries Q has the C.R. property and to construct a C.R. organization if Q is consecutively retrievable.

ACKNOWLEDGEMENTS:

The author would like to thank Prof. L. A. Zadeh and R. M. Karp for their comments. He is also grateful to Dr. S. S. Reddi for many valuable suggestions.

References.

[1] P.C. Gilmore and A.J. Hoffman, "A characterization of comparability graphs and of interval graphs". Canadian J. Math. 16 (1964), 539-548.

[2] A. Pnueli, A. Lampel and S. Even, "Transitive Orientation of graphs and identification of permutation graphs". Canadian J. Math. 23 (1971), 160-175.

[3] F. Harary, "Graph Theory", (Addition-Wesley, Reading, Massachusetts, 1969).

[4] S.P. Ghosh, "File organization: The consecutive retrieval property". Comm. of the ACM, Vol. 15, No. 9 (1972), 802-808.

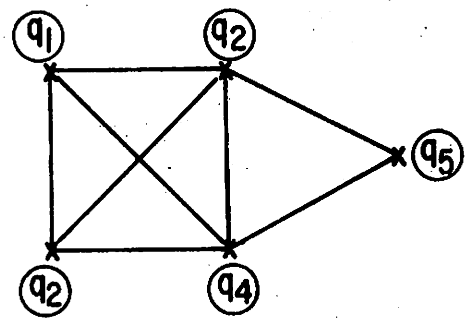


Figure 1

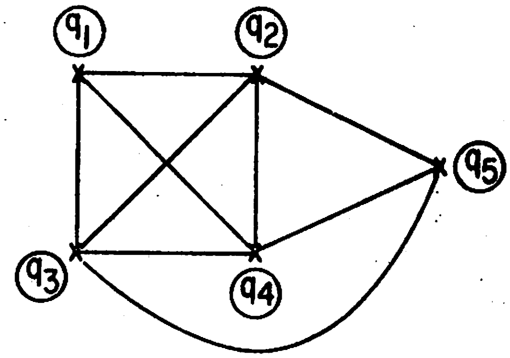


Figure 2

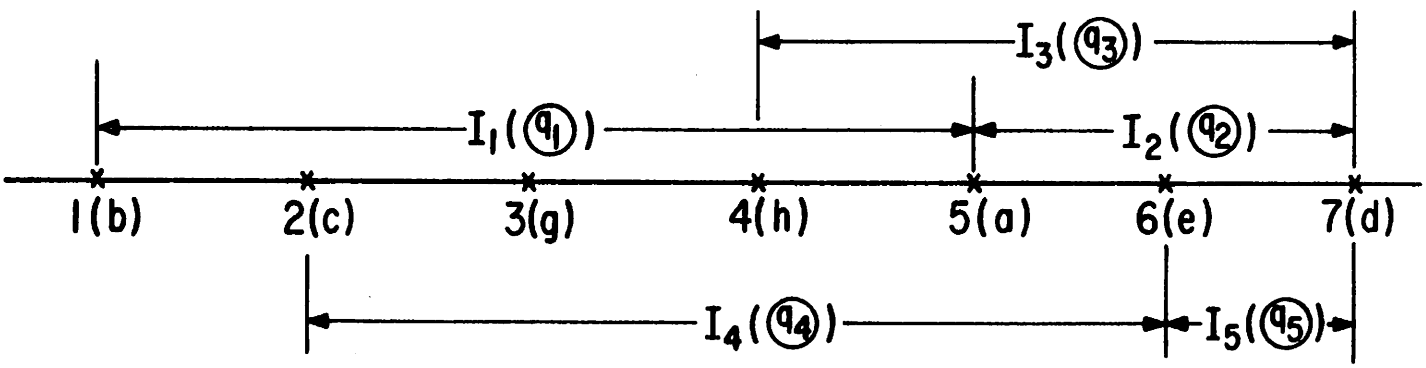


Figure 3

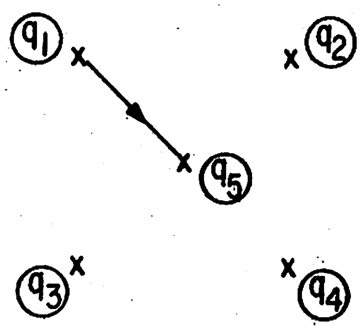


Figure 4

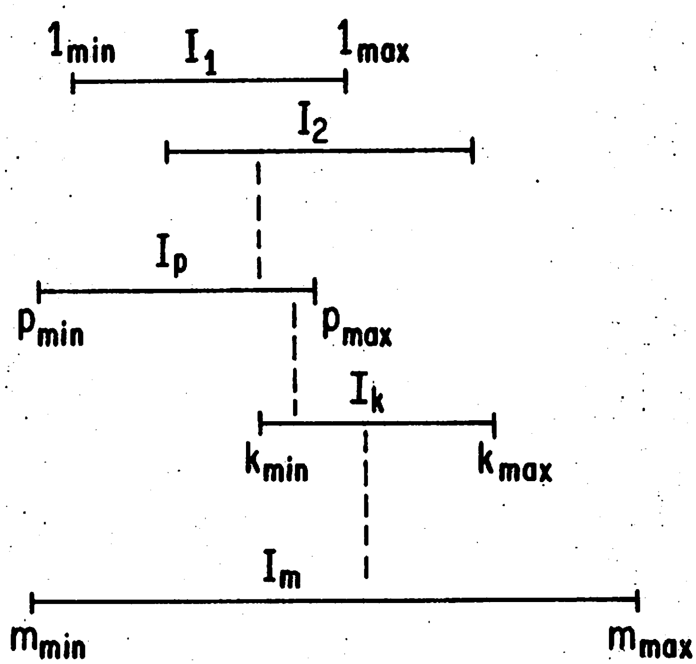


Figure 5

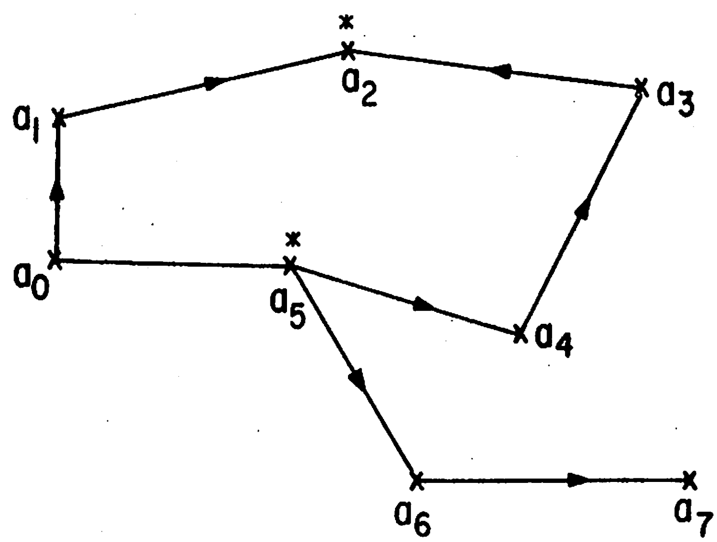


Figure 6

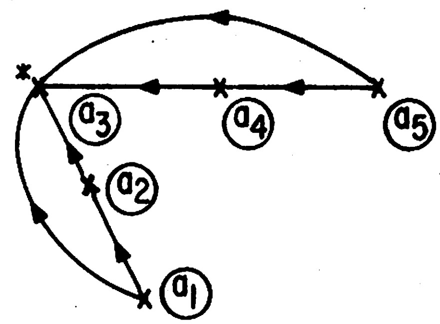


Figure 7

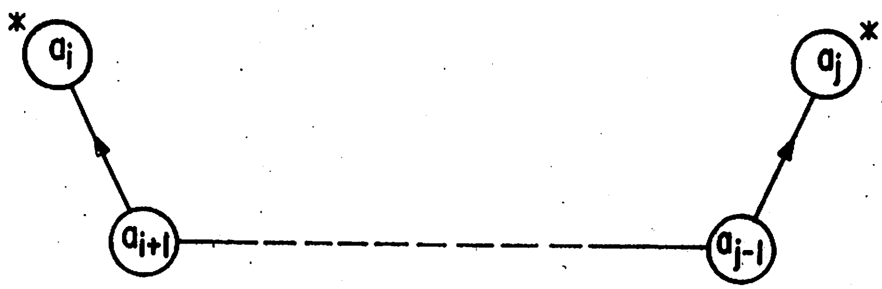


Figure 8

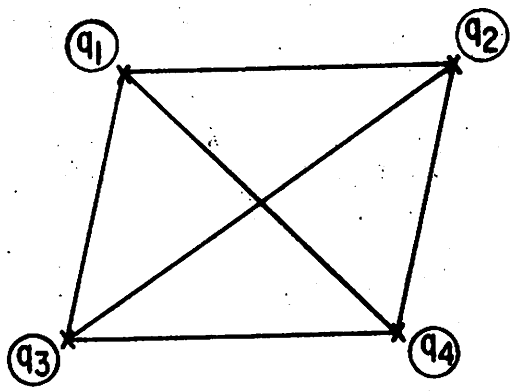


Figure 9

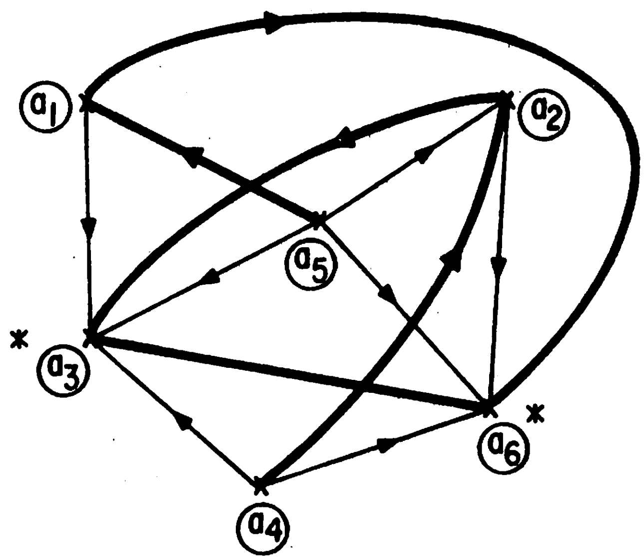


Figure 10

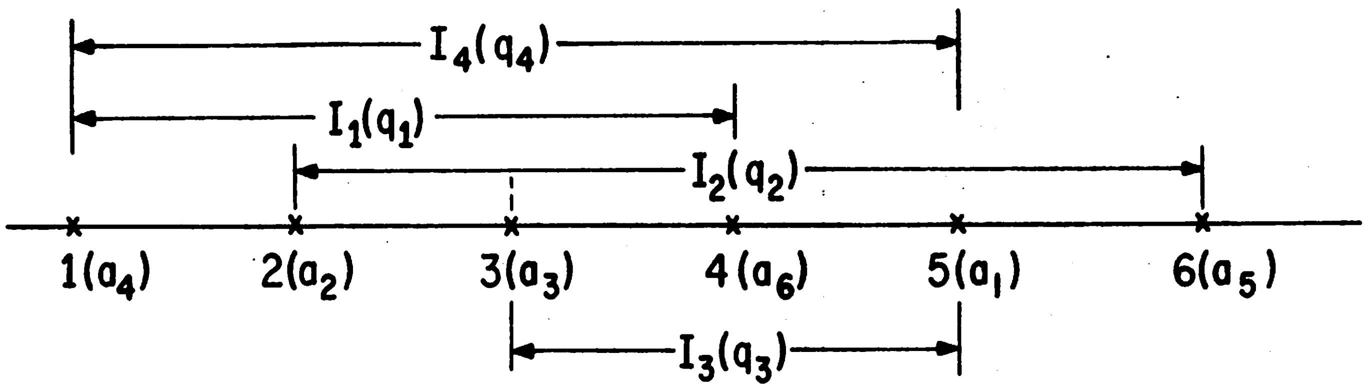


Figure 11

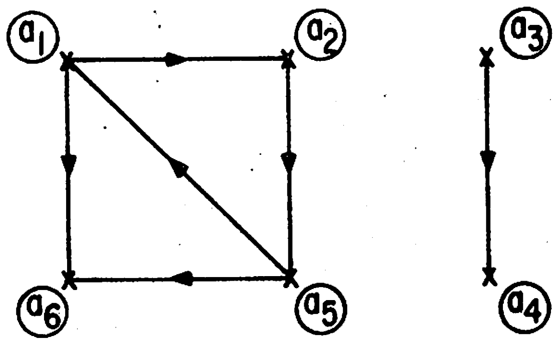


Figure 12

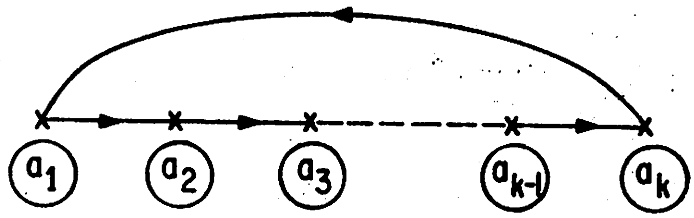


Figure 13

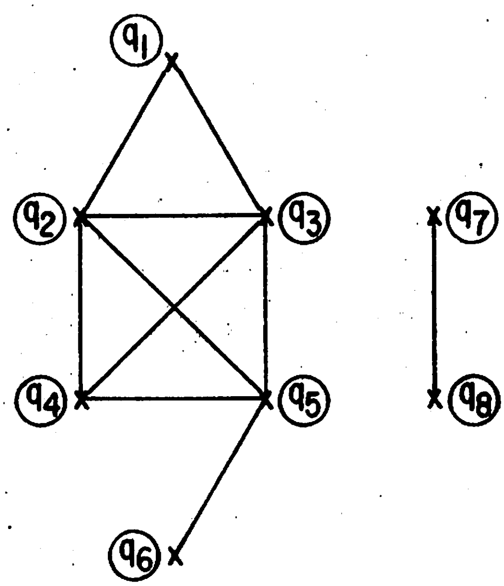


Figure 14

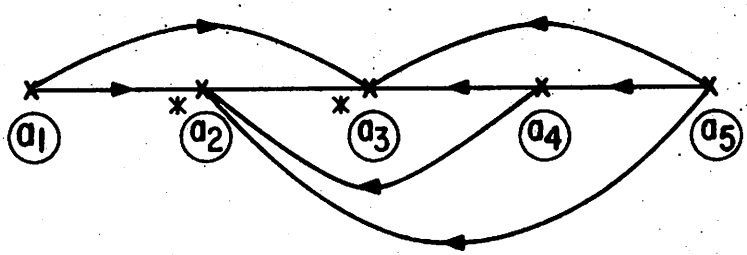


Figure 15

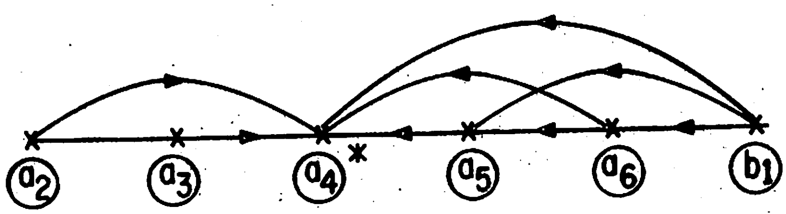


Figure 16

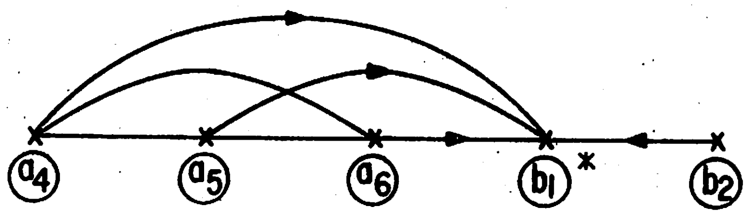


Figure 17

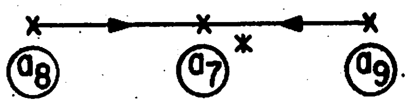


Figure 18

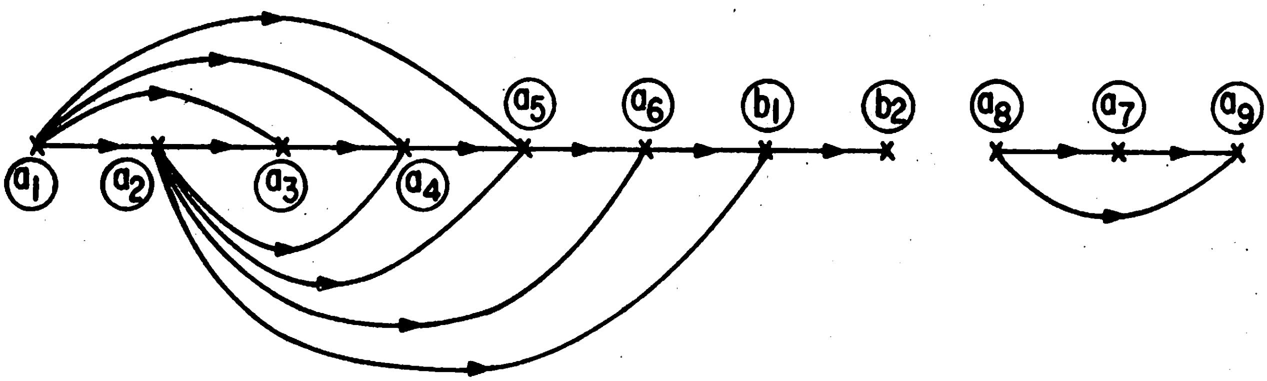
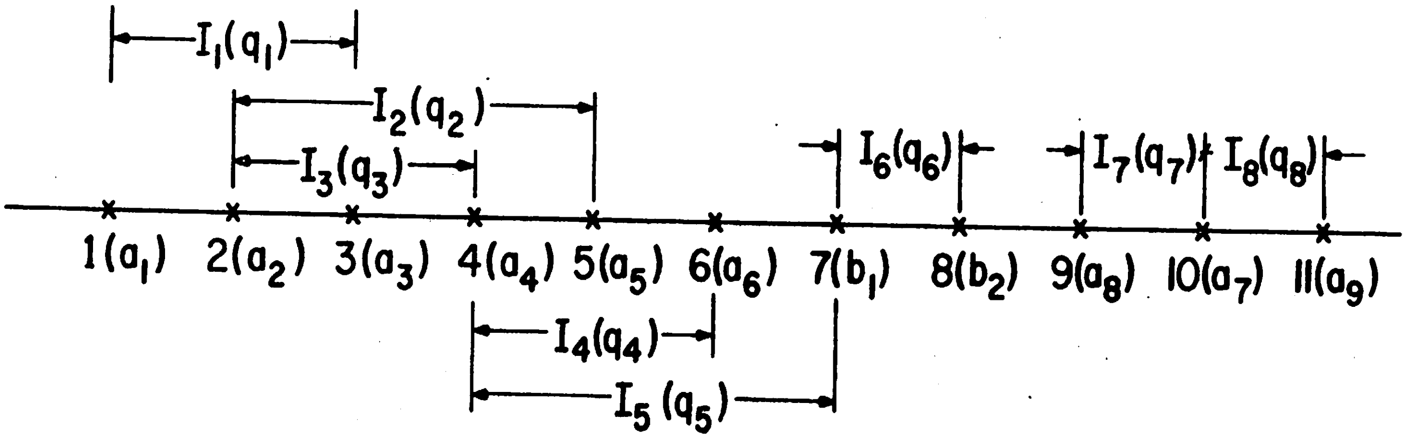


Figure 19



I1(q1)
I2(q2)
I3(q3)
I4(q4)
I5 (q5)
I6(q6)
I7(q7)
I8(q8)
1(a1)
2(a2)
3(a3)
4(a4)
5(a5)
6(a6)
7(b1)
8(b2)
9(a8)
10(a7)
11(a9)

Figure 20