Chapter 11

Lecture 11

Advanced Topics in Graph Algorithms

Spring Semester, 1994

Lecturer	:	Ron Shamir
Scribe	:	Vered Zilkha
Lecture 11	:	June 23

11.1 Preference and Indifference

Let V be a set of possibilities. A decision maker defines a preference relation on this set: for every pair of distinct members of V, he either prefers one over the other, or he feels indifferent about them.

We construct an oriented graph $H = (V, P)$, and an undirected graph $G = (V, E)$ as follows:

$$\forall x, y \in V$$

$xy \in P \Leftrightarrow x$ is preferred over y

$xy \in E \Leftrightarrow$ indifference is felt between x and y

By definition, $(V, P + P^{-1} + E)$ is complete.

We assume that the preference relation is acyclic. In fact we would want it to be transitive. Thus we require that P be a partial order.

One way of formalizing preference relations (in utility theory) is the notion of *Semiorder*.

Definition Given $\delta > 0$, a real-valued function $u : V \to R$ is called *a utility function of semiorder* if $\forall x, y \in V$, $xy \in P \Leftrightarrow u(x) \geq u(y) + \delta$

Remark: If P has a utility function of semiorder then P is a partial order (irreflexive,transitive).

Let us now characterize the partial orders that have a utility function of semiorder. We first state without proof a fundamental result from utility theory:

Theorem 11.1 (Scott - Suppes [1958]) *A binary relation (V, P) has a utility function of semiorder iff $\forall x, y, z, w \in V$*

(S1) P is irreflexive

(S2) If $xy \in P$, $zw \in P$ then either $xw \in P$ or $zy \in P$

(S3) If $xu \in P$, $yz \in P$ then either $xw \in P$ or $wz \in P$

Definition An *interval order* is a binary relation (V, P) such that $\forall i \in V$,we can match an interval I_i on the real line, such that $ij \in P \Leftrightarrow I_i < I_j$ (i.e $I_i \cap I_j = \emptyset$ and I_i is on the left side of I_j)

Claim 11.2 *A partial order (V, P) is an interval order iff it satisfies (S1),(S2).*

(The proof of this claim is left as an exercise)

The claim implies: If (V, P) is an oriented graph, (V, E) is an undirected graph, and $P + P^{-1} + E$ is the edge set of the complete graph on V then if (V, P) is a semiorder then (V, E) is an interval graph.

(The converse is not true. For example see figure 11.1).

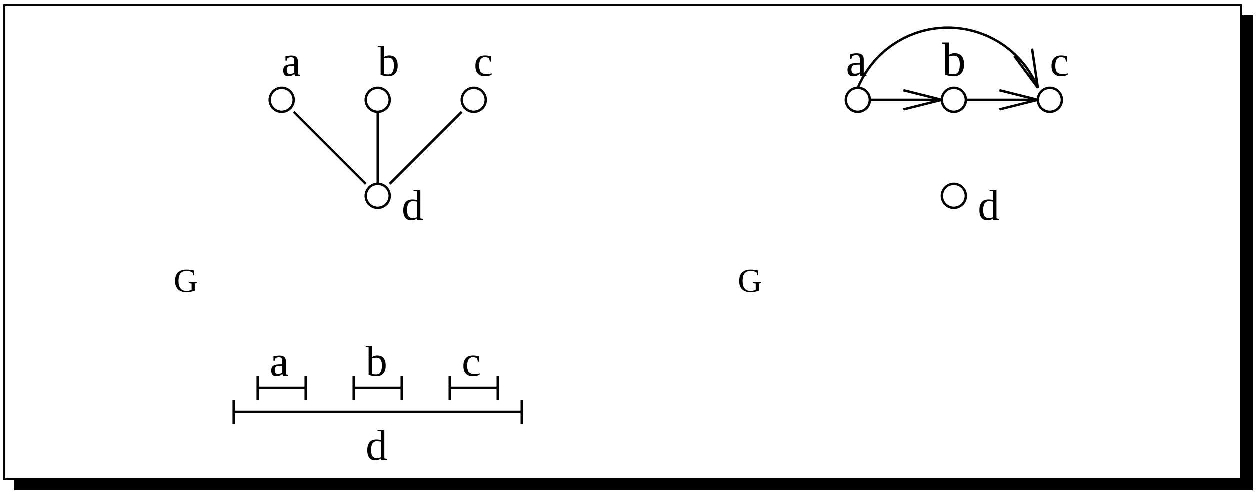

Figure 11.1: Example

Definition An interval graph $G = (V, E)$ is called a *proper* interval graph if it has a realization $\{I_v\}_{v \in V}$ such that no interval properly contains another.

Theorem 11.3 (Roberts [1969]) *Let $G = (V, E)$ be an undirected graph. The following statements are equivalent:*

(1) There exists a real valued function $u : V \to R$ such that $\forall xy \in V$, $x \neq y$, $xy \in E \Leftrightarrow | u(x) - u(y) | < 1$

(2) There exists a semiorder (V, P) such that $\bar{E} = P + P^{-1}$

(3) $\bar{G}$ is a comparability graph and every transitive orientation of $\bar{G} = (V, \bar{E})$ is a semiorder

(4) G is an interval graph that doesn't contain an induced $K_{1,3}$

(5) G is a proper interval graph

(6) G is a unit *interval graph*

Proof:

$(1) \Rightarrow (6)$ Let u be a function that satisfies (1). We will match to each vertex $x \in V$ an open interval $I_x = (u(x) - \frac{1}{2}, u(x) + \frac{1}{2})$. Then $\forall x, y, x \neq y$, $I_x \cap I_y \neq \emptyset \Leftrightarrow | u(x) - u(y) | < 1 \Leftrightarrow xy \in E$. Therefore $\{I_x\}_{x \in V}$ is a unit-interval realization of G.

$(6) \Rightarrow (5)$ In a unit-interval realization of G, no interval can properly contain another, because they all have length 1. Therefore this realization will be proper.

$(5) \Rightarrow (4)$ Let $\{I_x\}_{x \in V}$ be a proper interval realization of G. Suppose G contains an induced subgraph $G_{\{y,z_1,z_2,z_3\}}$ isomorphic to $K_{1,3}$ where $\{z_1, z_2, z_3\}$ is a stable set and y is adjacent to each z_i. Without lost of generality, suppose that $I_{z_1} < I_{z_2} < I_{z_3}$. Then I_y must properly contain I_{z_2}, a contradiction. Thus, G can have no induced copy of $K_{1,3}$.

$(4) \Rightarrow (3)$ Since G is an interval graph, $\bar{G} = (V, \bar{E})$ is a comparability graph. Let F be a transitive orientation of $\bar{G}$. It is easy to show that F satisfies $(S1), (S2)$. Now we will see that it also satisfies $(S3)$. By Scott-Suppes Theorem this will imply F is a semiorder.

Assume to the contrary that $xy \in F$, $yz \in F$, $xw \notin F$, $wz \notin F$. By the transitivity of F it follows that: $wx \notin F$, $zw \notin F$, $wy \notin F$, $yw \notin F$. But $xz \in F$. Thus $G_{x,y,z,w}$ induces a $K_{1,3}$, a contradiction.

$(3) \Rightarrow (2)$ Immediate.

$(2) \Rightarrow (1)$ If (V, P) is a semiorder, then by definition there exist a real valued function $\tilde{u} : V \to R$ and a constant $\delta > 0$ such that $xy \in P \Leftrightarrow \tilde{u}(x) - \tilde{u}(y) \geq \delta$. Denoting $u(x) = \frac{\tilde{u}(x)}{\delta}$ we get $xy \in P \Leftrightarrow u(x) - u(y) \geq 1$, and because $P + P^{-1} = \bar{E}$ we get that $xy \in E \Leftrightarrow xy \notin P$ and $xy \notin P^{-1} \Leftrightarrow | u(x) - u(y) | < 1$.

■

Additional results on unit interval graphs:

Definition A (0,1)-valued matrix satisfies the *consecutive 1's property* for the columns if its rows can be permuted such that on each column, the 1's will appear consecutively.

Definition A **layout** of a graph $G = (V, E)$ is a one-to-one function (a permutation) $L : V \to \{1, \ldots, \mid V \mid\}$. We will denote $\underline{\alpha_L}$ - the linear order $L^{-1}(1) < L^{-1}(2) < \ldots < L^{-1}(n)$.

Let $G = (V, E)$ be an interval graph. Then other equivalent conditions to G being unit interval are:

- The clique matrix of G satisfies the consecutive 1's property for rows and for columns.
- Let $A^*(G)$ be the adjacency matrix of G when we add 1's on the diagonal. Then $A^*(G)$ satisfies the consecutive 1's property for columns. (Roberts 1968)
- G is a triangulated graph and it doesn't contain as an induced subgraph any of the graphs appearing in figure 11.2.

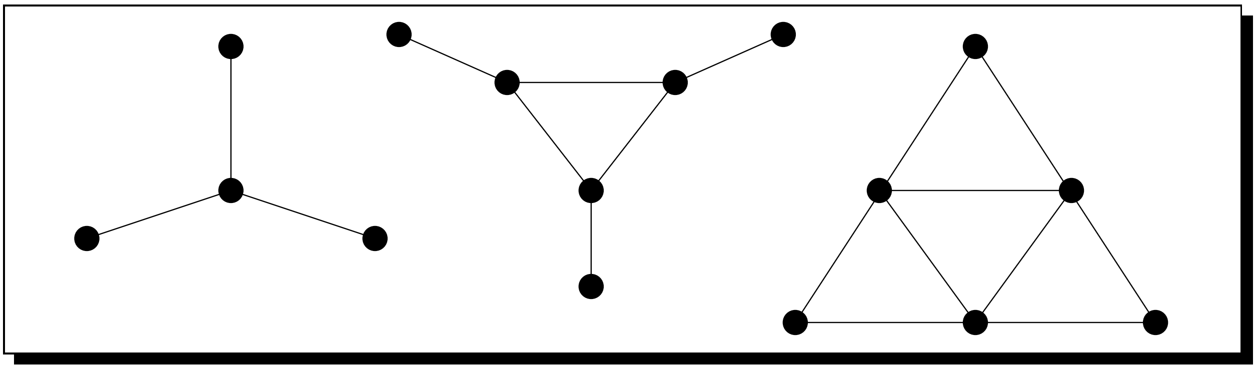

Figure 11.2: Wegner 1967

- There exists a layout L of G such that for every vertex v, $L(Adj(v)) = \{L(v) - k_v, L(v) - k_v + 1, \ldots, L(v) - 1, L(v) + 1, \ldots, L(v) + l_v\}$ (Deng, Hell, Huang 1992)

Some hard optimization problems involving unit interval graphs are:

- Adding a minimum number of edges to a graph to get a unit interval graph is NP-complete. (Golumbic, Kaplan, Shamir 1993)
- Removing a minimum number of edges to get a unit interval graph is NP-complete. (Golumbic, Kaplan, Shamir 1993)

11.2 Recognizing interval graphs

Theorem 11.4 *$G = (V, E)$ is an interval graph iff it has a layout L that satisfies:*

$$[L(u) < L(v) < L(w), uw \in E] \Rightarrow vw \in E \tag{11.1}$$

(The proof of this theorem is left as an exercise).

Theorem 11.5 *Let $G = (V, E)$ be a co-comparability graph, and let F be a transitive orientation of $\bar{G}$. We denote $d(u)$ the out-degree of u in (V, F). If L is a layout of G such that*

$$d(u) > d(v) \Rightarrow L(u) < L(v) \tag{11.2}$$

then G is an interval graph iff L satisfies condition 11.1

Proof:

$\Rightarrow$ If condition 11.1 holds then according to Theorem 11.4 G is an interval graph.

$\Leftarrow$ Assuming, to the contrary that L satisfies condition 11.2 but not condition 11.1, we will prove that G contains an induced C_4:

(a) if $uv \in F$ then by transitivity of F, $d(u) > d(v)$, and thus $L(u) < L(v)$.

(b) if $L(u) < L(v)$ and $uv \notin E$ then $uv \in F$ (else $vu \in F$ and from (a) we get $L(v) < L(u)$).

From (b) and the transitivity of F it follows that:

$$[L(u) < L(v) < L(w), uv \notin E, vw \notin E] \Rightarrow uw \notin E \tag{11.3}$$

$$[L(u) < L(v) < L(w), uw \in E, uv \notin E] \Rightarrow vw \in E \tag{11.4}$$

Now, suppose L does not satisfy 11.1. Thus there exist u, v, w such that

$$L(u) < L(v) < L(w), uw \in E, vw \notin E \tag{11.5}$$

From (b) it follows that $vw \in F$, and from condition 11.3 it follows that $uv \in E$.

Since $L(u) < L(v)$, it follows that $d(u) \geq d(v)$. $vw \in F, uw \notin F$, thus there exists y such that $uy \in F$, $vy \notin F$. Moreover, $yv \notin F$ (else $uy, yv \in F \Rightarrow uv \in F$, a contraction), thus $vy \in E$. consider that

If $yw \in F$ then $uy, yw \in F \Rightarrow uw \in F$, a contradiction.

If $wy \in F$ then $vw, wy \in F \Rightarrow vy \in F$, a contradiction.

Thus $yw \in E$, and $G_{\{u,v,w,y\}}$ induces a C_4.

■

Theorem 11.6 *Let $G = (V, E)$ be a graph with a layout L, and let us define functions $l, h : V \to \{0, \ldots, |V|\}$:*

$$l(u) = \min_{v \in N[u]} L(v)$$

$$h(u) = \begin{cases} \max\limits_{\substack{v \notin N[u] \\ L(v) < L(u)}} L(v) & \{v \mid L(v) < L(u), v \notin N(u)\} \neq \emptyset \\ 0 & \text{otherwise} \end{cases}$$

then L satisfies condition 11.1 iff $\forall w \in V, h(w) < l(w)$

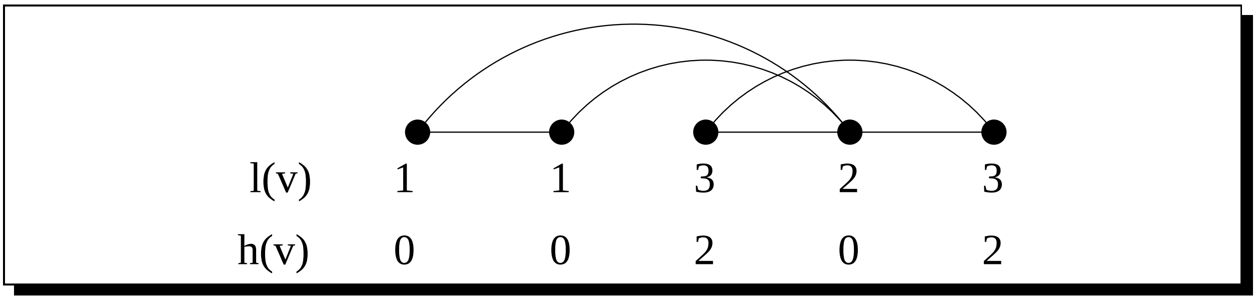

Figure 11.3: Example of function l, h

Proof:

- $\Rightarrow$ If L does not satisfy condition 11.1 then there exist u, v, w such that $L(u) < L(v) < L(w)$, and $uw \in E$, but $vw \notin E$. Thus $l(w) \leq L(u) < L(v) \leq h(w)$.

- $\Leftarrow$ If there exists $w \in V$ such that $l(w) < h(w)$ then taking $u = L^{-1}(l(w))$, $v = L^{-1}(h(w))$ gives us vertices u, v, w such that $L(u) < L(v) < L(w)$, and $uw \in E$, but $vw \notin E$, thus contradicting condition 11.1.

■

11.2.1 A quadratic algorithm 1

We first describe an elementary $O(n^2)$ algorithm for recognizing interval graph. This algorithm is due to Ramalingam and Rangan [1990].

Let $G = (V, E)$ be an undirected graph.

1. Find a transitive orientation of $\bar{G}$, and call it F.

2. Compute a layout of G according to a decreasing order of out-degrees of vertices in (V, F).

3. For each vertex, compute $l(u), h(u)$.

G is an interval graph iff $\forall w \in V,\ h(w) < l(w)$.

correctness:

1. If G is an interval graph then $\bar{G}$ has a transitive orientation. From theorem 11.5, the layout L computed in step 2 of the algorithm satisfies condition 11.1. From theorem 11.6 the condition $h < l$ holds.

2. If G is not an interval graph then no layout L satisfies condition 11.1, and from theorem 11.6 the condition $h < l$ fails.

complexity:

1. Finding a transitive orientation of $\bar{G}$: $O(n^2)$ (Spinrad [1985]), $O(n + \bar{m} \log n)$ (McConnell - Spinrad 1994). (We don't need to check that the orientation is transitive)

2. Sorting by out-degrees of vertices: $O(\bar{m} + n \log n)$. [Where $\bar{m} = | E(\bar{G}) |$].

3. Computing $l(u)$ for all vertices: $O(\bar{m})$.

   Computing $h(u)$ for all vertices: $O(m)$.

We can see that the complexity does not exceed $O(n^2)$.

Remarks:

* We can produce a realization of an interval graph, using the layout L.(this follows from the proof of theorem 11.4, which is constructive.

* The proof implies that every transitive orientation of $\bar{G}$ has a linear extension L (one or more) that satisfies condition 11.1.

11.2.2 A Linear Algorithm

To check whether a graph $G = (V, E)$ is an interval graph we can use the following algorithm:

1. Check whether G is triangulated

2. Produce all the maximal cliques in G, and then produce the clique matrix M of G

3. Check whether the matrix M satisfies the consecutive 1's property

The first two stages can be implemented in a linear time: $O(\mid V \mid + \mid E \mid)$. We will see now that the third stage can be implemented in a linear time, by an algorithm of Booth and Leuker [1975].

The data structure being used is called PQ - trees: a PQ-tree is a tree with a root, that has vertices of two kinds: P and Q. The sons of a P-vertex are an unordered set. The sons of a Q-vertex are an ordered set. The leaves of the tree are labeled by 1-1 correspondence with the members of the set X. The *frontier* of the tree is the permutation of X, obtained by reading the labels on the leaves from left to right.

Two PQ-trees, T, T' are *equivalent*, denoted $T \equiv T'$, if one can be obtained from the other by applying a sequence of the following transformation rules:

- Arbitrarily permute the children of a P-vertex.
- Reverse the order of children of a Q-vertex.

Any frontier obtainable from a PQ-tree equivalent with T is said to be *consistent* with T. the set of permutations on X consistent with T is denoted by:

$$CONSISTENT(T) = \{FRONTIER(T') \mid T \equiv T'\}$$

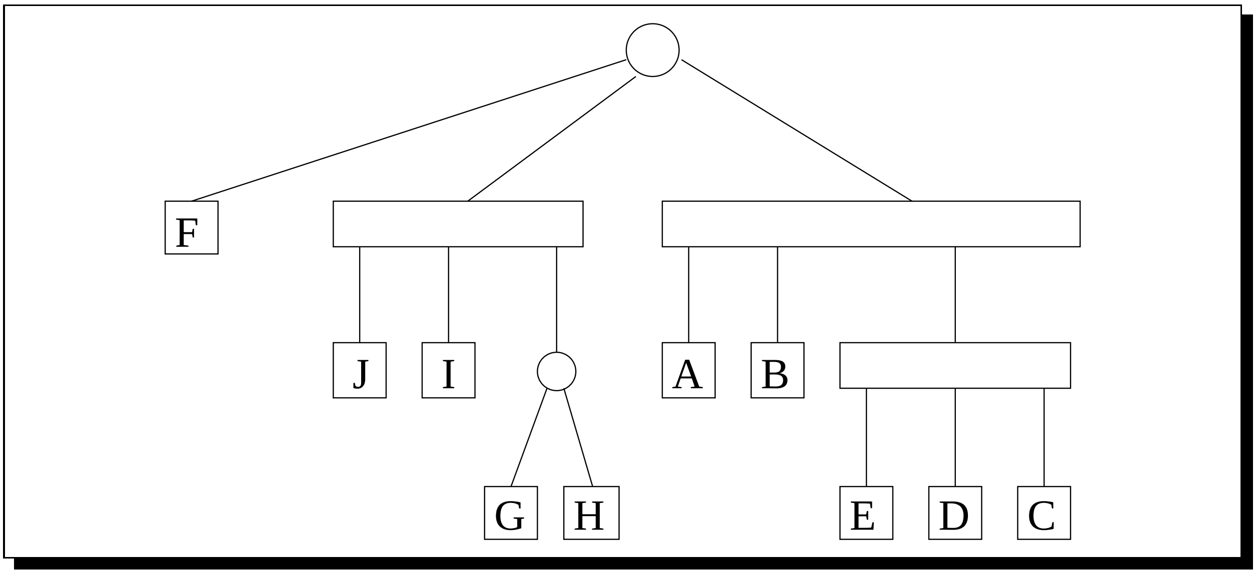

Figure 11.4: a PQ-tree. The FRONTIER here is [F J I G H A B E D C]

The *null tree* T_0 has no nodes and $CONSISTENT(T_0) = \emptyset$. The *universal tree* has one internal P-vertex, the root, and a leaf for every member of X.

The relation between the consecutive arrangement problem and PQ-trees is as follows:

Let φ be a collection of subsets of a set X, and let $\Pi(\varphi)$ be the collection of all permutations π of X such that the members of each $I \in \varphi$ occur consecutively in π.

Theorem 11.7 (Booth - Leuker [1976])

1. *For every collection of subsets φ of X there exists a PQ-tree T such that* $\Pi(\varphi) = CONSISTENT(T)$.
2. *For every PQ-tree there exists a collection of subsets φ such that* $\Pi(\varphi) = CONSISTENT(T)$.

The role of the Q-vertices is to restrict the number of permutations by making some of the brother relationships rigid.

The main approach of the algorithm is as follows:

Begin with the collection of all possible permutations. Introduce the subjects one at a time, reducing the tree to those permutations that also satisfy the constraints of the next subset. Below is an implementation of the algorithm:

Procedure CONSECUTIVE(X, φ):
begin
  $T \leftarrow$ universal tree;
  for each $I \in \varphi$ **do**
  $T \leftarrow$ REDUCE(T,I);
return T;
end

The routine REDUCE attempts to apply on the existing tree, one of 11 templates, according to the tree and the set I, and replaces a part of the tree accordingly. The templates are applied from the bottom to the top of the tree. We omit the details here. figure 11.5 gives an example of the execution of the algorithm.

Theorem 11.8 (Booth - Leuker [1976]) *The collection of permutations $\Pi(\varphi)$ can be calculated (when implemented with a PQ-tree) in time $O(\mid \varphi \mid + \mid X \mid + \sum_{I\in\varphi} \mid I \mid)$.*

Corollary 11.9 *A $(0,1)$-valued matrix $A_{m\times n}$ with f 1's can be checked for the consecutive 1's property in $O(m + n + f)$ steps.*

Corollary 11.10 *Interval graphs can be recognized in linear time.*

Definition A (0,1)-valued matrix on satisfies the *circular 1's property* in the columns if its rows can be rearranged such that on each column, the 1's will appear cyclicly consecutive.

Remarks:

- A (0,1)-valued matrix satisfies the circular 1's property iff it satisfies the circular 0's property.

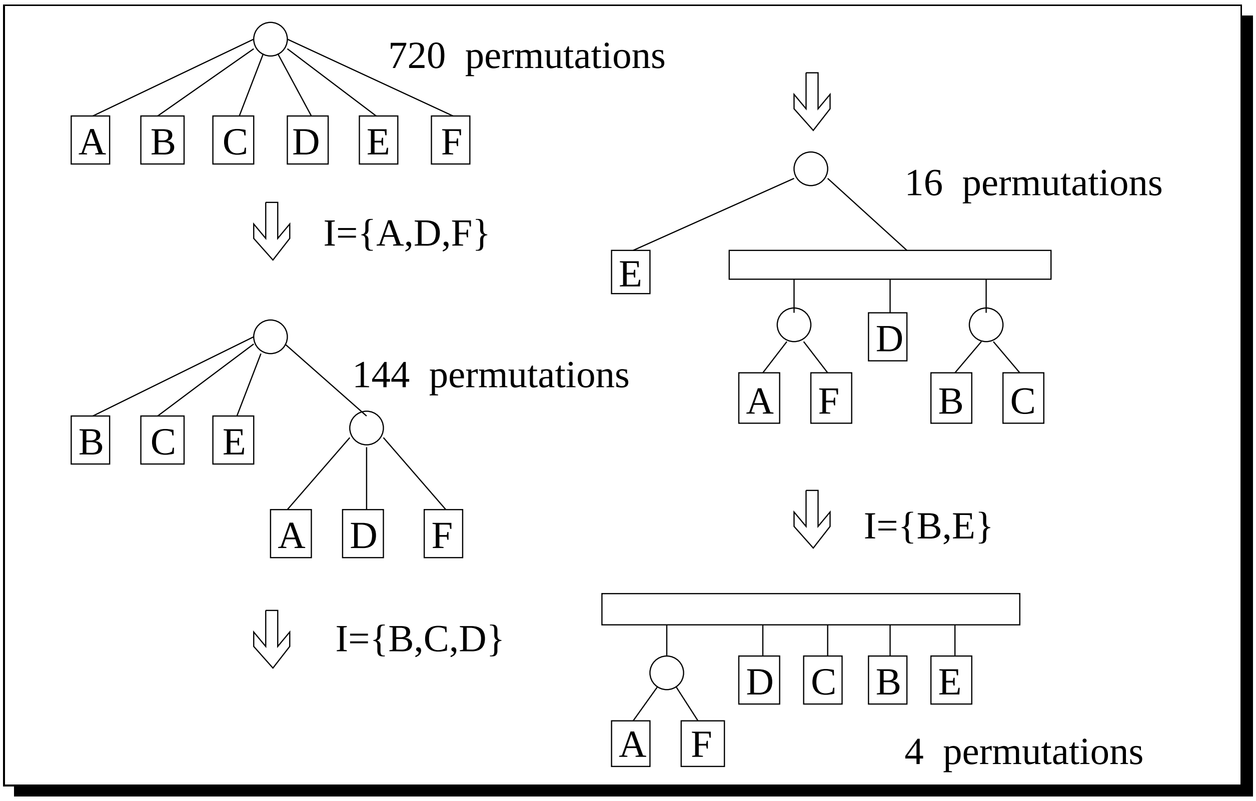

Figure 11.5: Example of execution of CONSECUTIVE

- If a matrix satisfies the circular 1's property for the columns, then after switching 0's with 1's in a column, the property still holds.

Theorem 11.11 (Tucker [1971) *] Given a (0,1)-valued matrix M, switch 0's with 1's in every column that has 1 in the k-th row. Denote the new matrix M'.*

M satisfies the circular 1's property iff M' satisfies the consecutive 1's property.

Proof:

$\Rightarrow$ If M satisfies the circular 1's property then M' satisfies the circular 1's property. Move the k-th row to the first row by rotating the rows cyclicly. In this row there will be 0's only in M'. Hence, M' satisfies the consecutive 1's property.

$\Leftarrow$ If M' satisfies the consecutive 1's property then it satisfies the circular 1's property, and therefore M satisfies the circular 1's property.

■

Theorem 11.12 (Booth [1975] **)** *A (0,1)-valued matrix $A_{m \times n}$, with f 1's can be checked for the circular 1's property in time $O(m+n+f)$, assuming that the matrix is given as a list of the 1's for each row.*

Proof: Let k be the row with the smallest number of 1's, and let M' be the matrix obtained from M by switching 0's with 1's in every column that has 1 in the k-th row.

M' has at most $2f$ 1's, because in every row the number of 1's is at most doubled (the number of 0's that become 1's is not larger than the original number of 1's in the row). Then apply the algorithm for consecutive 1's on M'. According to the previous theorem, M' satisfies the consecutive 1's property iff M satisfies the circular 1's property. ■

11.3 More about the consecutive 1's property

Theorem 11.13 (Booth [1975] **)** *Given a (0,1)-valued matrix $M_{r \times c}$ and an integer k, deciding whether or not there exists an $r \times k$ submatrix of M satisfying the consecutive 1's property (or circular 1's property) is NP-complete.*

Proof: It is easy to show that the problem is NP. To prove that the problem is NP complete we can show a reduction from the Hamiltonian-path problem: Given an undirected graph $G = (V, E)$, we produce the adjacency-matrix A of size $\mid V \mid \times \mid E \mid$ (The rows represent vertices and the columns represent edges). Then it is easy to see that A has a submatrix of size $\mid V \mid \times \mid V \mid$ that satisfies the consecutive 1's property $\Leftrightarrow$ There exists a Hamiltonian path in G. For circular 1's property, reduce from Hamiltonian cycle instead. ■

Theorem 11.14 (Kou [1977]) *Given a (0,1)-valued matrix, finding a permutation of the rows that minimizes the total number of consecutive blocks of 1's appearing in the columns is NP-complete.*

Theorem 11.15 (Kou [1977]) *Given a (0,1)-valued matrix, finding a permutation of the rows that minimizes the number of times a row must be split into two pieces to obtain consecutive 1's in the columns is NP-complete.*

Proof: First, we will introduce an equivalent problem: Given a finite alphabet Γ and a collection of subsets $s_1, \ldots s_k$ such that $s_i \subseteq \Gamma$, we want to find a chain $\sigma \in \Gamma^*$ of length $\leq B$, such that for each s_i there exists a consecutive interval of length $\mid s_i \mid$ in G which contains all letters in s_i.

It is obvious that the problem is NP. to show that it is NP-complete we will show a reduction from the Hamiltonian-path problem: given a graph $G = (V, E)$ we will define $\Gamma = V + E$, $B = 2 \mid E \mid +1$, $s_i = \{< v_i > \cup \bigcup_{x \in Adj(v_i)} (v_i, x)\}$ $\forall i = 1 \ldots n$.

- $\Rightarrow$ If there exists a Hamiltonian path, without lost of generality $v_1 \ldots v_n$, then we can produce the chain:

  $$v_1, \underbrace{(v_1, \cdot), (v_1, \cdot), \ldots, (v_1, v_2)}_{\text{all edges of } v_1}, v_2, \underbrace{(v_2, \cdot), \ldots, (v_2, v_3)}_{\text{the other edges of } v_2}, v_3, \ldots, v_n, \underbrace{(v_n, \cdot), \ldots, (v_n, \cdot)}_{\text{the other edges of } v_n} .$$

  The length of the chain: $\mid V \mid +2 \mid E \mid -(\mid V \mid -1) = B$.

- $\Leftarrow$ If there exists a chain with length $\leq B$, since for every two sets s_i, s_j, $s_i \not\subseteq s_j$, $\mid s_i \cap s_j \mid \leq 1$, it follows that the length of the chain is at least:

  $$\sum_{i=1}^{n} \mid s_i \mid -(n-1) = n + 2 \mid E \mid -(n-1) = B$$

  Therefore: if there exists a chain of length B, it matches a Hamiltonian cycle (there is an edge between each pair of adjacent s_i, s_j).

∎

Remarks

- The problem is NP-complete even if $\forall i$, $\mid s_i \mid = 4$ (the same reduction from a 3-regular graph).
- the problem is polynomial if $\forall i$, $\mid s_i \mid = 2$. (exercise)
- Finding a polynomial approximation is an open problem.

Theorem 11.16 (Goldberg, Golumbic, Kaplan, Shamir [1993)] *Given a* $(0, 1)$*-valued matrix, deciding there exists a permutation of the rows that gives at most k blocks of 1's in every column is NP-complete, for every $k \geq 2$.*