Map Assembly - Continued

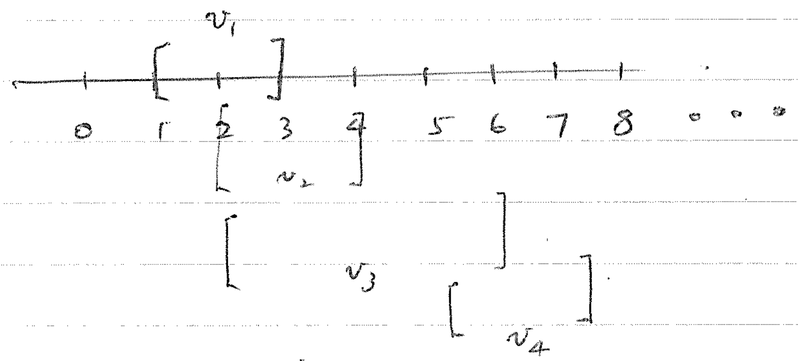
Lecture 8

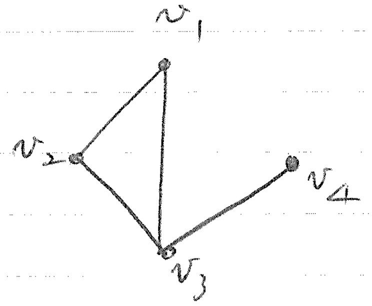
Interval Graphs

An interval graph $G=(V,E)$ is the graph where $V(G)$ is a set of closed intervals in real line and two vertices are adjacent iff their corresponding intervals have non-empty intersection.

For example:

$$V = \{[1,3],[2,4],[2,6],[5,7]\}$$





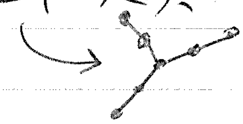
G

<u>Lemma 1</u> Every interval graph is triangulated, i.e., every cycle of length larger than 3 has a "chord".

<u>Proof</u>. Let C be a cycle of G with $|V(G)| \geq 4$ and

C has no chords. Let $C = (u_1, u_2, \cdots, u_k)$, $k \ge 4$. W.L.O.G We may $a_1 \le a_2 \le \cdots \le a_k$ where $u_i = [a_i, b_i]$, $i = 1, 2, \cdots, k$. Since $u_1 \sim_G u_k$, $a_k \le b_1$. Now, if $b_1 \le b_j$ for some $j = 2, \cdots, k-2$, then $a_j \le a_k \le b_1 \le b_j$ and therefore $u_j \sim_G u_k$. $\rightarrow\leftarrow$. Hence, $b_1 > b_j \ \forall j = 2, \cdots, k-2$. This implies $u_1 \sim u_j$. $\rightarrow\leftarrow$. ∎

(其逆不一定真)



<u>Definition</u> (Transitive orientation property)

An undirected graph $G = (V, E)$ has a transitive orientation property if G has an orientation (on E) such that the resulting digraph $D_G(V, A)$ satisfies $\forall a, b, c \in V$, $(a, b) \in A$ and $(b, c) \in A$ imply $(a, c) \in A$. (G is said to be transitively orientable.)

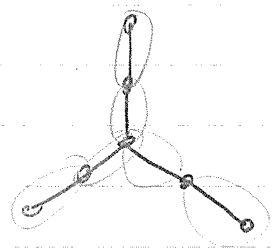
<u>Definition</u> (Comparability graph)

An undirected graph that is transitively orientable is called a comparability graph.

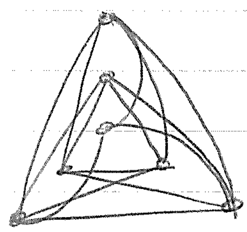
<u>Lemma 2</u> The complement of an interval graph is a comparability graph. (其逆不一定真)

Proof. Let $\bar{G}$ be the complement of an interval graph G. Then, $[i,j]$ and $[i',j']$ of $V(G)$ are adjacent in $\bar{G}$ if and only if $[i,j] \cap [i',j'] = \phi$. Now, define $[i,j] \longrightarrow [i',j']$ provided $j < i'$. Now, it is easy to check if $[i_1,j_1] \rightarrow [i_2,j_2]$ and $[i_2,j_2] \longrightarrow [i_3,j_3]$, then $[i_1,j_1] \longrightarrow [i_3,j_3]$ since $i_3 > j_2 > i_2 > j_1$. ∎

Example



G



$\tilde{G}$

Not a comparability graph

(?)

Theorem 3 (Gilmore and Hoffman, 1964) Let G be an undirected graph. The following statements are equivalent.

(1) G is an interval graph.

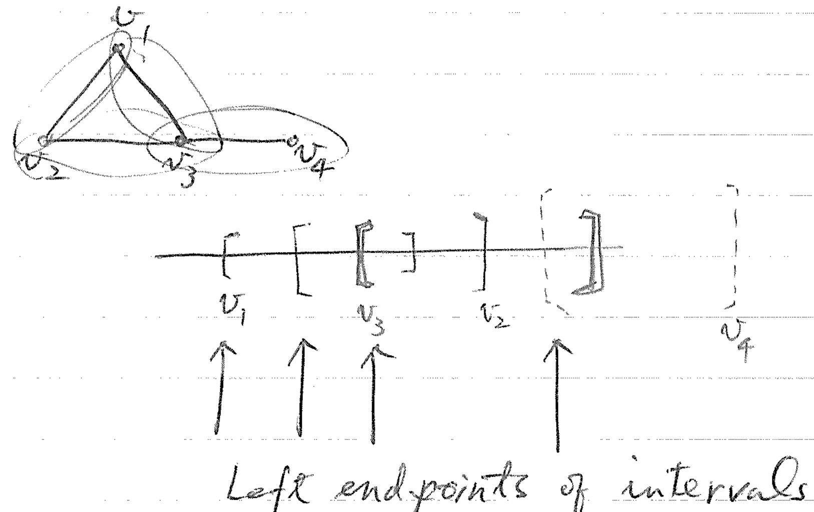
(2) G is a triangulated graph and its complement is a comparability graph.

(3) The maximal cliques of G can be linearly ordered s.t. $\forall\ x \in V(G)$, the maximal cliques containing x occur consecutively

Proof. (1) ⇒ (2) (By Lemma 1, 2.)

(2) ⇒ (3) complicate.

(3) ⇒ (1) $\forall x \in V(G)$, let $I(x)$ denote the set of all maximal cliques of G that contains x. The sets $I(x)$ form the intervals of the interval graph.



(*) Theorem 3 reduces the problem of recognition of an interval graph to the problems of recognition of triangulated and comparability graphs (Fulkerson and Gross, 1965; Pnueli et al., 1971).

Mapping with Restriction Fragment Fingerprints

The simplest case of mapping with restriction fragment fingerprints is Single Complete Digest mapping (SCD mapping). (Olson et al. 1986)

In this case the fingerprint of a clone is a multiset of the sizes of its restriction fragments in a digest by a restriction enzyme. An SCD map is a placement of clones and restriction fragments consistent with the given SCD data (Gillett et al. 1995)

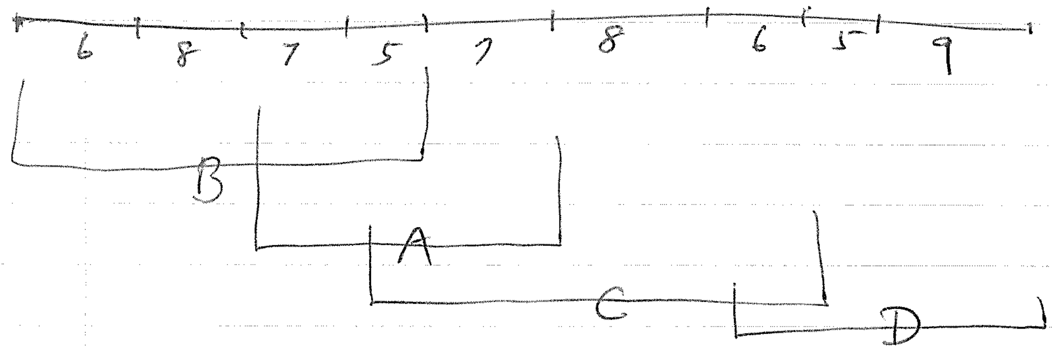
SCD Mapping Problem

Find a most compact map (i.e., a map with the minimum number of restriction fragments) that is consistent with SCD data.

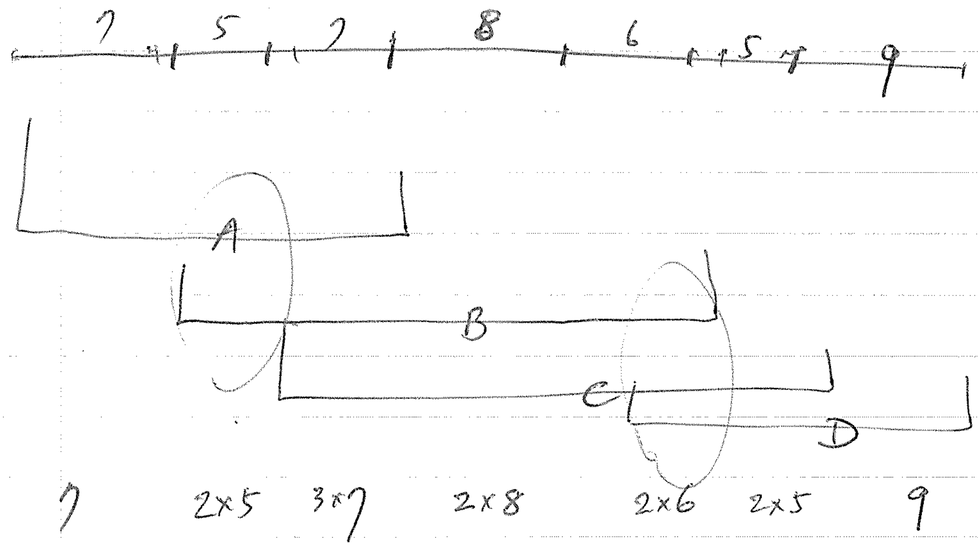
$$A = \{5, 7, 7\},\ B = \{5, 6, 7, 8\},\ C = \{5, 6, 7, 8\},\ D = \{5, 6, 7\}.$$

Solutions

①



②



Solution ② is better which has 7 restriction fragments.

(*) The problem of finding the most compact map is NP-hard.

(**) 在实际的应用上，"Fingerprints of clones"通常可以用统计的结果来估计 clones 出现的顺序(ordering)。

(⁂)

Jiang and Karp (1998) formulated SCD mapping with known clone ordering as a constrained path cover problem on a special multistage graph.

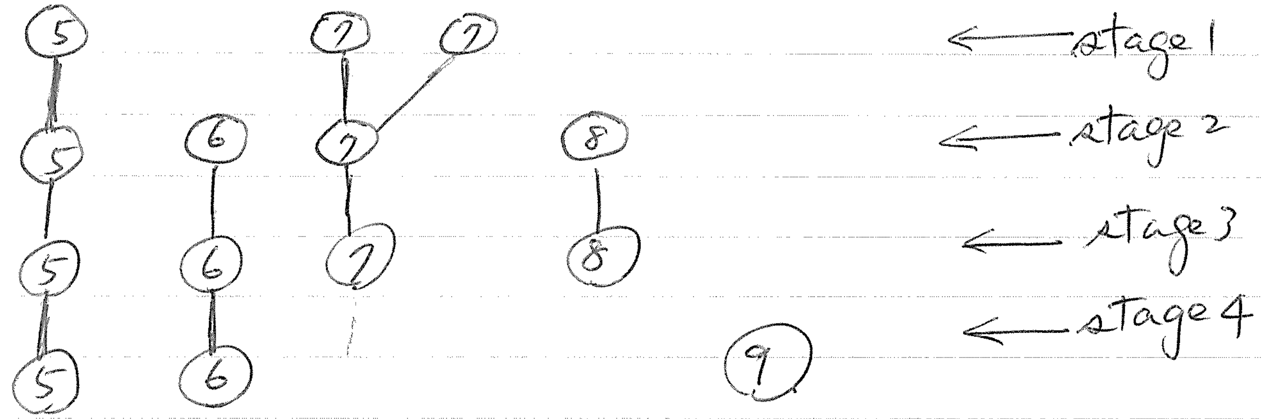
Let $S = \{S_1, S_2, \cdots, S_n\}$ be an instance of SCD mapping, where S_i is a multiset representing the fingerprint of the i-th clone in the clone ordering by the left endpoints.

Definition (labeled multistage graph)

A labeled multistage graph G (called a clone-fragment graph) consists of n stages, with the i-th stage containing $|S_i|$ vertices. At stage i, G has a vertex for each element x of S_i (including duplicates), with label x. Two vertices are adjacent if they are at adjacent stages and have identical label.

Example.

$S_1 = \{5, 7, 7\}$, $S_2 = \{5, 6, 7, 8\}$, $S_3 = \{5, 6, 7, 8\}$, $S_4 = \{5, 6, 9$



$$\begin{bmatrix} 5 & 0 & 7 & 7 & 0 & 0 \\ 5 & 6 & 7 & 0 & 8 & 0 \\ 5 & 6 & 7 & 0 & 8 & 0 \\ 5 & 6 & 0 & 0 & 0 & 9 \end{bmatrix}$$

Definition (Path cover)

A path cover is a collection of paths such that every vertex is contained in exactly one path.

(*) Any map for S corresponds to a path cover of G of the same cardinality.

From (2), we see that the map gives a path cover with "7" paths and they are "7", "2×5", 3×7, 2×8, 2×6, 2×5, and 9.

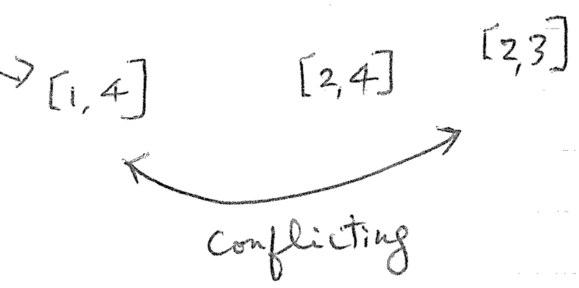
(**) We may have a better path cover, but it corresponds to no map! For example

(4×5), 3×6, 3×7, 7, (2×8), 9

(See page (8).)

<u>Definition</u> (Conflicting paths)

Let $[i,j]$ denote the path from the i-th stage to the j-th stage. Two paths $[i,j]$ and $[i',j']$ are conflicting if $i < i' < j' < j$. A path cover is conflict-free if it has no conflicting paths.

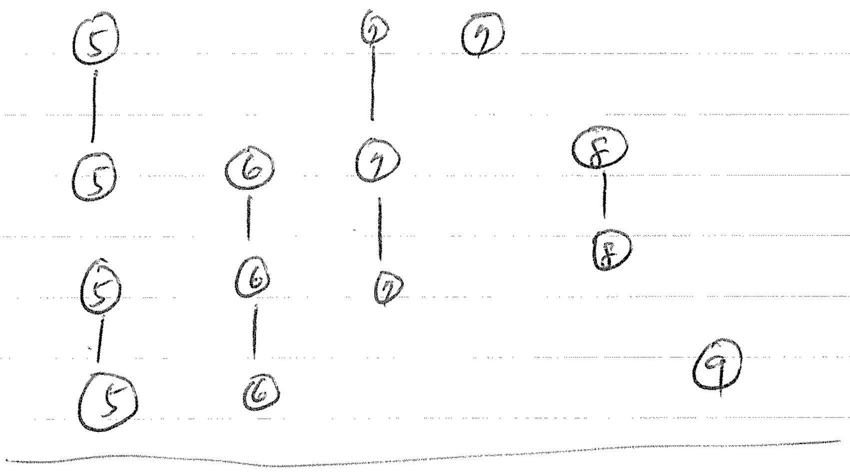


- The cardinality of a map of S is the # of fragments.

<u>Definition</u> (consistent path cover)

A path cover of G is consistent if it corresponds to a map of S with the same cardinality.

Example, Solution (2) corresponds to the following path cover:



"7 paths"

<u>Proposition</u> A path cover (A set of intervals) is consistent if and only if it is conflict-free.

<u>proof.</u> Derived from the fact of interval graphs. (?) ▨

(*) § Jiang and Karp (1998) obtained a 2-approximation algorithm for SCD mapping with a given ordering of clones.