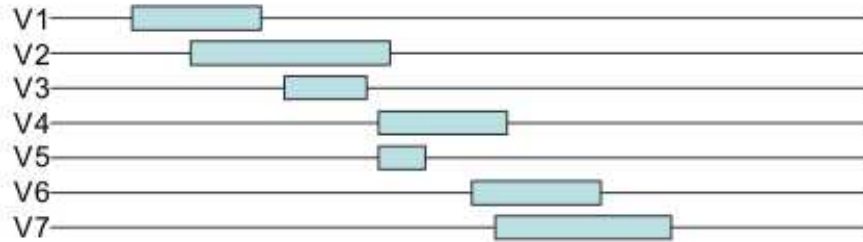


Introduction to Bioinformatics (autumn 2005)

Excercise 5

Group	time	place
Riikka Kaven	Tuesday 15.11 at 12.15–14.00	BK106

1. Construct the interval graph corresponding to the following intervals:



2. Let G be a graph. If G is an interval graph then each vertex corresponds to a one-dimensional interval; in this case the vertices can be listed in the order of the starting points of the intervals (like V1, V2, ... above). Consider a fixed order for the vertices. Describe how to test whether the graph is an interval graph under this assumption (i.e. whether vertices in this order can be mapped into intervals such that the edges define the intersections). Does this give a rise to a practical algorithm for determining whether a given graph is an interval graph or not?

3. (Problem 8.6 in J & P) Find the shortest common superstring for eight 3-mers:

{AGT, AAA, ACT, AAC, CTT, GTA, TTT, TAA}

and solve the following two problems:

- a) Construct the graph with 8 vertices corresponding to these 3-mers (Hamiltonian path approach) and find a Hamiltonian path (7 edges) which visits each vertex exactly once. Does this path visit every edge of the graph? Write the superstring corresponding to this Hamiltonian path.
 - b) Construct the graph with 8 edges corresponding to these 3-mers (Eulerian path approach) and find an Eulerian path (8 edges) which visits each edge exactly once. Does this path visit every vertex of the graph exactly once? Write the superstring corresponding to this Eulerian path.
4. (Problem 8.9) Use the Eulerian path approach to solve the SBH problem for the following spectrum:

$$\mathcal{S} = \{ATG, GGG, GGT, GTA, GTG, TAT, TGG\}.$$

Label edges and vertices of the graph, and give all possible sequences s such that $Spectrum(s, 3) = \mathcal{S}$.

5. (Problem 8.10) The SBH problem is to reconstruct a DNA sequence from its ℓ -mer composition.

- Suppose that instead of a single target DNA fragment, we have two target DNA fragments and we simultaneously analyze both of them with a universal DNA array. Give a precise formulation of the resulting problem (something like the formulation of the SBH problem).
- Give an approach to the above problem which resembles the Hamiltonian Path approach to SBH.
- Give an approach to the above problem which resembles the Eulerian Path approach to SBH.

(No algorithms are required here, just definition, and mapping to the two graph problems above.)