Algorithms in Genome Research
Winter 2016/2017

Exercises

Number 08, Discussion: 2017 January 27

1. Given two genomes (represented as sequences of gene families):

$$A = (1, 4, 7, 6, 5, 4, 5, 1, 4, 3, 2)$$
$$B = (1, 2, 3, 4, 3, 4, 5, 7)$$

   (a) Find all maximal common intervals of A and B of size at least 2.

   (b) Find all maximal median gene clusters of A and B of size at least 3 with a symmetric set distance of at most 1.

2. A common interval C is called a *nested common interval* of two genomes if either $|C| = 2$, or if $|C| > 2$ and it contains a nested common interval of size $|C| - 1$.

   A nested common interval of size ℓ is *maximal* if it is not contained in a nested common interval of size $\ell + 1$.

   (a) Find all maximal nested common intervals in the following two genomes:

$$A = (4, 6, 5, 7, 3, 1, 2)$$
$$B = (1, 2, 3, 4, 5, 6, 7)$$

   (b) Develop an algorithm to find all maximal nested common intervals in two permutations.

3. Given two genomes (represented as indeterminate strings):

$$A = (\{9\}, \{3, 7, 10\}, \{8, 13\}, \{4, 6\}, \{11\}, \{1, 9\}, \{2\})$$
$$B = (\{5\}, \{7\}, \{10\}, \{4, 8, 12\}, \{6\}, \{3\}, \{2, 10\}, \{1\}, \{12\})$$

   Find all weak common intervals of A and B of size at least 2.