Algorithms in Comparative Genomics

Exercise sheet 7, 09.12.2021

Exercise 1 (Bounds for the SCJ distance) (6 pts)

Theoretical bounds for the SCJ distance with respect to the DCJ and the inversion distances are

$$\mathrm{d}_{\mathrm{DCJ}}(\mathbb{A}_{\triangleright}^{f}, \mathbb{B}_{\triangleright}^{f}) \leq \mathrm{d}_{\mathrm{SCJ}}(\mathbb{A}_{\triangleright}^{f}, \mathbb{B}_{\triangleright}^{f}) \leq 4\,\mathrm{d}_{\mathrm{DCJ}}(\mathbb{A}_{\triangleright}^{f}, \mathbb{B}_{\triangleright}^{f})$$

$$2\,\mathrm{d}_{\mathrm{INV}}(\mathbb{A}_{\triangleright}^{f}, \mathbb{B}_{\triangleright}^{f}) \leq \mathrm{d}_{\mathrm{SCJ}}(\mathbb{A}_{\triangleright}^{f}, \mathbb{B}_{\triangleright}^{f}) \leq 4\,\mathrm{d}_{\mathrm{INV}}(\mathbb{A}_{\triangleright}^{f}, \mathbb{B}_{\triangleright}^{f})$$

For each one of these four bounds:

  Show that it is tight by giving an example of pairs of mutually distinct genomes that fulfill it.

Exercise 2 (Singular DCJ-indel model) (12 pts)

Consider the following singular genomes:

$$\mathbb{A} \;=\; [\,a_1\ 4\ 1\,]\quad [\,a_2\ 6\ a_3\ 3\ a_4\ 2\ a_5\ 5\,]\quad [\,a_6\ a_7\ 7\ 8\ a_8\ \overline{9}\,]\quad [\,a_9\,]\quad\text{and}$$

$$\mathbb{B} \;=\; [\,b_1\ 1\ b_2\ 2\ b_3\ 3\ b_4\ 4\ b_5\,]\quad [\,5\ b_6\ 6\,]\quad [\,b_7\ 7\ 8\ b_8\ 9\,]\quad (\,b_9\ b_{10}\,)$$

1. Give the sets of genes $\mathcal{G}_\star$, $\mathcal{A}$ and $\mathcal{B}$.

2. Construct the relational graph $RG(\mathbb{A}, \mathbb{B})$.

3. For each component C of the relational graph $RG(\mathbb{A}, \mathbb{B})$:

   (a) Give the type of C (cycle, singleton, $\mathbb{A}\mathbb{B}$-path, $\mathbb{A}\mathbb{A}$-path or $\mathbb{B}\mathbb{B}$-path).
   (b) Give the number of runs $\Lambda(C)$;
   (c) Give the run-type of C (ε, $\mathcal{A}$, $\mathcal{B}$, $\mathcal{AB}$ or $\mathcal{BA}$);
   (d) Compute the minimum number of indels $\lambda(C)$ that are necessary for sorting C separately.

4. Find all chains of deducting recombinations.

5. Compute the DCJ-indel distance $\mathrm{d}_{\mathrm{DCJ}}^{\mathrm{ID}}(\mathbb{A}, \mathbb{B})$.