Department of Computer Science

Prof. Dr. Nikolaus Augsten
Jakob-Haringer-Str. 2
5020 Salzburg, Austria
Phone: +43 662 8044 6347
E-Mail: nikolaus.augsten@plus.ac.at

PARIS LODRON UNIVERSITÄT SALZBURG

Similarity Search in Large Databases
Winter Semester 2025/26

Exam
28.01.2026

Name: _____ **Student ID:** _____

Hints

- Check whether you received all pages of the exam (5 pages).
- Write your name or your student ID on each sheet of the exam and hand in all pages.
- All answers are expected to be written on the exam sheets.
- Clearly highlight and enumerate additional pages that are used for longer answers. Match your text with the according exercise.
- Only use pencils that are permanent and non-red colored.
- Use the notation and techniques discussed in the lecture.
- Exercises with more than one solution are not graded.
- You are allowed to use one A4 sheet with your personal notes (single-sided, handwritten or printed).
- Exam duration: **45 minutes**

Signature _____

Grading

Filled by the examiner

Exercise	1	2	3	Bonus	Sum
Total points	10	10	10	4	30 + 4
Points reached					

Exercise 1 - *Algorithm for Left-Most Leaf Descendants.* 10 Points

Execute the recursive algorithm for computing the left-most leaf descendants of tree T in Figure 1. What is the input, what is the output, what are the steps that you execute?

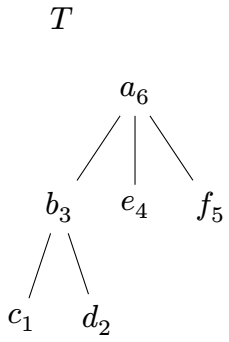


Figure 1: Example tree T.

Exercise 2 - *pq-Gram Distance*. 10 Points

Compute the pq-gram distance between the trees T_1 and T_2 in Figure 2.

(1) Compute the distance for the parameters $q = 1$ and $p = 1$.

(2) Compute the distance for the parameters $q = 2$ and $p = 2$.

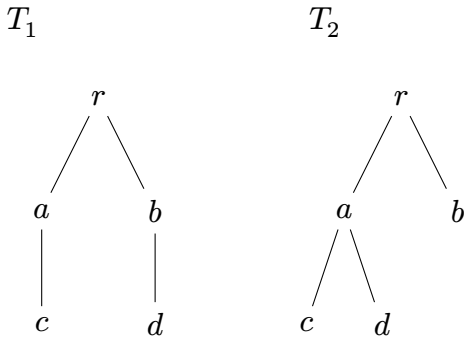


Figure 2: Two example trees T_1 and T_2.

Exercise 3 - *TED Lower Bounds*. 10 Points

Consider the trees T_1 and T_2 in Figure 3.

(1) Compute the lower distance bound between T_1 and T_2 based on *binary branches*.

(2) Compute the lower distance bound between T_1 and T_2 based on *traversal strings*.

(3) Based on (1) and (2), what can we say about the tree edit distance between T_1 and T_2?

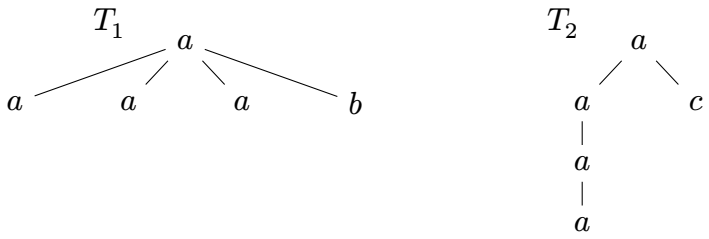


Figure 3: Two example trees T_1 and T_2.

Exercise 4 - ***Bonus Question:*** *Traversal Strings.* 4 Points

Prove the following statement:

For two labeled, ordered trees T_1, T_2 with preorder and postorder strings $pre(.)$ and $post(.)$, **unique labels**, and string edit distance δ_s, we have

$$\max\{\delta_s(pre(T_1), pre(T_2)), \delta_s(post(T_1), post(T_2))\} = 0 \Leftrightarrow T_1 = T_2$$

Hint: $\Leftarrow$ is obvious and needs no proof. For $\Rightarrow$, it is enough to show that the preorder and postorder strings of a tree uniquely determine the tree.