Aids Allowed: ONLY your *own notes* taken during lectures and office hours, the lecture *slides and recordings* (for all sections), and the *Course Notes* (textbook).

Submission Instructions

- Submit your work directly on **MarkUs**—even if you are late!

- You may type your answers or hand-write them *legibly*, on paper or using a tablet and stylus.

- You may write your answers directly on the question paper, or on another piece of paper/document.

- You may submit your answers as a single file/document or as multiple files/documents. Each document may contain answers for only part of one question, an entire question, or multiple questions, but *please label each part of your answers* to make it clear what you are answering.

- There is no "required file", *but please give short names to your file(s)*, like "Q2.png" or "TT4.pdf".

- You **must** submit your answers in PDF or as photos (JPEG/JPG/GIF/PNG/HEIC/HEIF). **Other formats** (e.g., Word documents, LATEX source files, ZIP files) **are NOT accepted**—you must **export** or **compile** documents to PDF, **convert** images into a supported format, and upload each file **individually**.

For all questions in this test, write your proofs *formally*, including a header and a proof body with justifications for each deduction. Remember that we are looking for evidence that you understand the conventions for writing correct proofs, so pay attention to the *structure* of your answers, in addition to their content!

5. **[8 marks] Algorithm Analysis: Average-Case**

Consider the following algorithm.

```
1  def max_alpha(s: str) -> int:
2      ''' Return the largest index k such that the characters in s[0:k] are
3          sorted in alphabetical order.  Precondition: s is non-empty. '''
4      i = 1
5      while i < len(s) and s[i-1] <= s[i]:
6          i = i + 1
7      return i
```

For each $n \in \mathbb{N}$ with $n \geq 2$, let $\mathcal{I}_n$ be the set that contains all strings of length n with 2 a's and $(n-2)$ b's, in any order. (For example, $\mathcal{I}_4 = \{aabb, abab, abba, baab, baba, bbaa\}$.)

Note that $|\mathcal{I}_n| = \binom{n}{2} = \frac{n(n-1)}{2}$ because each element of $\mathcal{I}_n$ is made up of n individual characters, all of which are equal to b except for 2 of them, and there are exactly $\binom{n}{2}$ many different ways to choose the 2 positions that will contain a.

(a) **[2 marks]** Let k be the value returned by `max_alpha`(s), for some input $s \in \mathcal{I}_n$. Write an expression for the "exact" number of steps executed by `max_alpha`(s), *as a function of k.*

Show your work (explain how you count your steps and how you arrive at your answer).

*Reminder: this test contains **five (5)** separate questions, plus the Academic Integrity statement!*

(b) **[2 marks]** For each $n \in \mathbb{N}$ such that $\mathcal{I}_n$ is defined, and each possible return value k for `max_alpha`, give an expression for **the number of inputs** $s \in \mathcal{I}_n$ for which `max_alpha`(s) returns k. In other words, calculate $\left|\{s \in \mathcal{I}_n \mid \texttt{max\_alpha}(s) \text{ returns } k\}\right|$.

Show your work (explain how you obtain your expression, and how it relates to the algorithm).

(c) **[2 marks]** What is the exact average-case running time of `max_alpha` over the set of inputs $\mathcal{I}_4$? Give your answer in the form of a simplified, concrete rational number (like $17/5$).

Show your work (explain what you are calculating at each step).

(d) **[2 marks]** Perform an average-case analysis of `max_alpha`, for the input set $\mathcal{I}_n$ defined above, for an arbitrary $n \in \mathbb{N}$ such that $\mathcal{I}_n$ is defined. **You may leave your final answer in the form of a sum**, but it should be simplified as much as possible.

Show your work. In particular, explain what each part of your sum represents.