LE/EECS 4101 GS/EECS 5101
Advanced Data Structures

Quiz 1

Shahin Kamali

York University

February 15th, 2023

Write your name and student id here:

"The formula 'Two and two make five' is not without its attractions." *Fyodor Dostoevsky*

- **Do not open this booklet until instructed.**
- You are NOT allowed to use any printed/written material.
- Please **turn off your cell phones** and put them in your bags.
- Manage your time. We start the exam at 5:30 and end at 6:00. **You have 30 minutes**. **Don't waste too much time on a single question.**
- The exam is printed **double-sided**. There are **6 pages**, including this cover page and a blank page (use it if you need more space). You must submit ALL pages.

1. True/False and Multiple-Choice Questions (15 marks)

Provide your short answers in the provided boxes. There is no need to justify your answers.

1. True or False: If an algorithm has an amortized cost of 2023 per operation, then the running time of Alg for each individual operation is $O(1)$. False **Answer:** The conclusion is not always correct. E.g., $n-1$ operations could have a cost of 1 each and the last operation could have cost n. The amortized cost is $O(1)$ but the las operation is not $O(1)$.

2. True or False: True or false: Height of a height-balanced binary search tree with n keys is always $\Theta(\log n)$. True

3. It is possible to augment an AVL tree on n nodes so that *rank* and *select* operations can be performed in $O(\log n)$. True

4. We aim to investigate the competitive ratio of an algorithm Alg for a minimization problem. We have discovered the following input sequences:

   - A sequence S_1 for which the cost of Alg is 10 and the optimal offline cost is 5;
   - A sequence S_2 for which the cost of Alg is 15 and the optimal offline cost is 10;
   - A sequence S_3 for which the cost of Alg is 20 and the optimal offline cost is 15.

   Specify which of the following statements holds:

   The competitive ratio of A is exactly 2. ☐
   The competitive ratio of A is at most 2. ☐
   The competitive ratio of A is at least 2. ☐
   The competitive ratio of A is exactly 4/3. ☐
   The competitive ratio of A is at most 4/3. ☐
   The competitive ratio of A is at least 4/3. ☐
   None of the other statements hold. ☐

   Answer: We cannot infer that competitive ratio is "at most" v for some value c; this is because we have not looked at all sequences, and the c.r. is the max ratio over all sequences. We can, however, infer that the c.r. is at least 2, because we have found a sequence for which Alg is twice worse than OPT.

   In order to get the full mark, you must choose "competitive ratio of A is at least 2". You may or may not choose "competitive ratio of A is at least 4/3" but you must not choose any other option.

5. True or False: There is an optimal offline list-update algorithm that only uses free exchanges. False

   Answer: It is False; there is an optimal algorithm that only uses paid exchanges but the same is not true for free exchanges.

6. Let A be a smart cow algorithm that works by walking 1 unit to the **right**, then back to the origin, then 2 units to the left, then back to the origin, then 4 units to the right, then back to the origin, and then continues with this doubling strategy. Let B be a similar algorithm, except that it first goes 1 unit to the **left**, then back to the origin,

then 2 units to the right, then back to the origin and continues this doubling strategy.
True or False: the competitive ratios of both A and B are 9. **True**

Answer: A symmetric argument shows that both algorithms have a c.r. of 9.

7. True or False: there is a unique AVL tree for a dictionary with keys $\{1, 2, \ldots, n\}$. **False** **Answer:**
For example, if you have four keys $1, 2, 3, 4$, any tree with 2 or 3 as the root will be a valid AVL tree.

2. Short Answer Questions (21 marks)

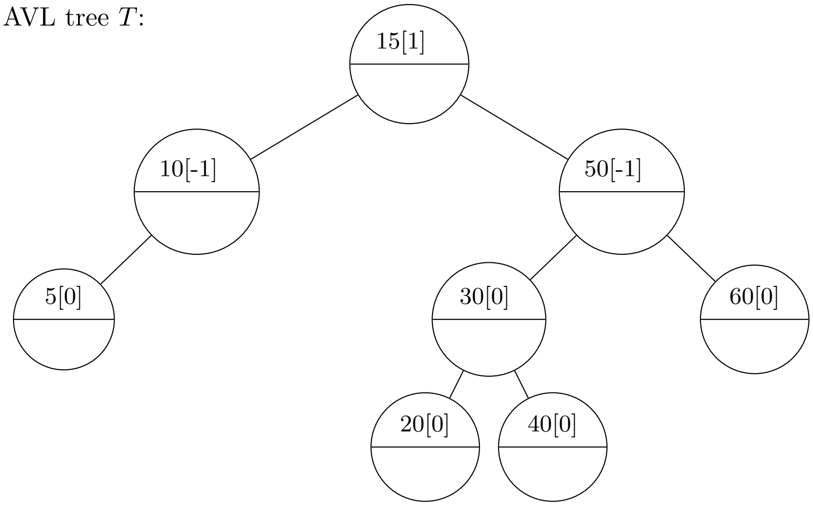
(a) Consider a variant of dynamic arrays in which when an array becomes full, instead of doubling the size of the array, we multiply the size of the array by 4. This way, array sizes will be powers of 4. For example, on the 17th operation (insertion), the size of the array is changed from 16 to 64.
What is the amortized cost of each operation? Briefly justify your answer.
Hint: $1 + 4 + 16 + 64 + \ldots + 4^k = \frac{4^{k+1}-1}{3}$

Answer: The total cost will be m (for inserting new items) plus $\sum_{j=0}^{\lfloor \log_4(m-1) \rfloor} 4^j$ (for copying old items to new arrays). It will be at most $m + \frac{4^{\log_4(m-1)+1}-1}{3} = m + 4(m-1)/3 - 1/3 = 7m/3 - 1/3$. So, the amortized cost per operation is roughly $7/3 \in \Theta(1)$.

(b) Consider the following randomized algorithm for the ski-rental problem: With a chance of $\frac{b-1}{b}$, buy the equipment at the beginning, and with a chance of $\frac{1}{b}$, always rent and never buy. Recall that the cost of buying is b and the cost of renting is 1 per day; let x denote the number of skiing days.
Write down the competitive ratio of the algorithm as a function of b and x and indicate whether the competitive ratio of the algorithm is bounded by a constant. Show your work.

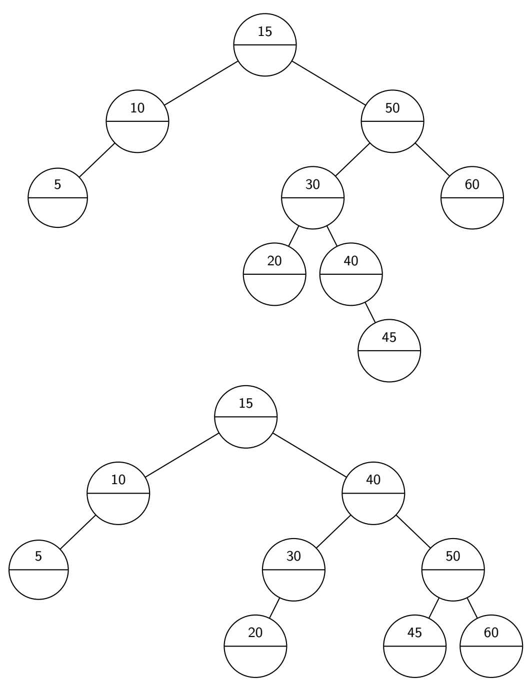
Answer: For any input defined by x (no. of skiing days), the expected cost of the algorith is $\frac{b-1}{b} \times b + \frac{1}{b}x = b - 1 + x/b$. The cost of OPT is $\min\{b, x\}$. The competitive ratio is then the max value of $\frac{b-1+x/b}{\min\{b,x\}}$, which is maximized when x goes to infinity; in this case, the competitive ratio converges to $\frac{b-1+x/b}{b} = 1 - 1/b + x/b^2$, which grows with x (the competitive ratio is unbounded).

(c) Consider the following AVL tree T:



1. Write in the missing balance factors in the figure above.

2. Perform operation *insert*(45) on T.
   It suffices to show the final tree (no need to show balance factors).



(d) Apply the Burrows-Wheeler transform on the following string; show your work and the output[1]

$$kakabeka\$$$

Assume \$ precedes all characters when you sort rotations.

Answer:

Rotations:	Sorted Rotations:	Transform:
kakabeka$	$kakabeka	akkkabea$
akabeka$k	a$kakabek	
kabeka$ka	abeka$kak	
abeka$kak	akabeka$k	
beka$kaka	beka$kaka	
eka$kakab	eka$kakab	
ka$kakabe	ka$kakabe	
a$kakabek	kabeka$ka	
$kakabeka	kakabeka$	

(e) Assume an initial list $\$ \to X \to Y \to Z \to W$. A compressing scheme that uses Move-To-Front has encoded the following numbers for a text T. Show what the actual text is. The numbers are 2 0 3 1 0.

Answer:

$\$ \to X \to Y \to Z \to W \to$ decode Y at 2
$Y \to \$ \to X \to Z \to W \to$ decode Y at 0
$Y \to \$ \to X \to Z \to W \to$ decode Z at 3
$Z \to Y \to \$ \to X \to W \to$ decode Y at 1
$Y \to Z \to \$ \to X \to W \to$ decode Y at 0

[1] Kakabeka is a beautiful waterfall in northwest Ontario.

Use this blank page for your draft work.