Dalhousie University
CSCI 6057/4117 — Advanced Data Structures
Winter 2022
Assignment 2

Distributed Wednesday, January 26 2022.

Due at 23:59 Wednesday, Rebruary 9 2022.

Guidelines:

1. All assignments must be done individually. The solutions that you hand in must be your own work.
2. Submit a PDF file with your assignment solutions via Brightspace. We encourage you to typeset your solutions using LaTeX. However, you are free to use other software or submit scanned handwritten assignments as long as **they are legible**.
3. Working on the assignments of this course is a learning process and some questions are expected to be challenging. Start working on assignments early.
4. Whenever you are asked to prove something, give sufficient details in your proof. When it is straightforward to prove a lemma/property/observation, simply say so, but do not claim something to be straightforward when it is not.
5. We have the following late policy for course work: `http://web.cs.dal.ca/~mhe/csci6057/assignments.htm`

Questions:

1. [10 marks] In the `Locate` algorithm shown in class for the resizable array solution that requires a space overhead of $O(\sqrt{n})$, we saw the following formula:

$$\sum_{j=0}^{k-1} 2^{\lfloor j/2 \rfloor} = \begin{cases} 2 \times (2^{k/2} - 1) & \text{if } k \text{ is even;} \\ 3 \times 2^{(k-1)/2} - 2 & \text{otherwise.} \end{cases} \tag{1}$$

   (i) [5 marks] Prove by induction that this equality holds for any positive integer k.

   (ii) [5 marks] Suppose that you were not given this equality. Instead, derive this formula from scratch. Show your work.

2. [10 marks] Prove the $\Omega(n \lg n)$ lower bound on sorting under the binary decision tree model. That is, prove that, given a sequence of n elements, the worst-case running time of any comparison-based algorithm that sorts these elements is $\Omega(n \lg n)$.

   Hint: Stirling's approximation may be helpful. It is on page 57 of the CLRS book. You can also find it in the following course note that I wrote for CSCI 3110:

   `https://web.cs.dal.ca/~mhe/csci3110/handouts/lecture3.htm`

3. [15 marks] In class we learned that the space cost of van Emde Boas trees satisfies

$$S(u) = (\sqrt{u} + 1)S(\sqrt{u}) + \sqrt{u}$$

   (Note that the last additive term on the right-hand side is slightly simpler that what is given in class; you are expected to know that this would not affect our result of analysis when presented using order notation.)

   (i) [8 marks] Use the substitution method to prove that $S(u) = O(u)$.

   Hint: Review Section 4.3 of the CLRS book if you are not familiar with the substitution method. Pages 85-86 on "Subtleties" are particularly helpful.

   `http://site.ebrary.com/lib/dal/docDetail.action?docID=10397652`

   (ii) [7 marks] Prove that $S(u) = \Omega(u)$.

4. [15 marks] Suppose that we have the following variant of van Emde Boas trees: Let W be a widget whose universe size is u (i.e., W stores u bits). Then, instead of using $\sqrt{u}$ subwidgets each of size $\sqrt{u}$ to represent W as done in class, we use $u^{1/3}$ subwidgets, and the universe size of each subwidget is $u^{2/3}$. The size of the summary widget of W is then set to $u^{1/3}$ so that it still has exactly one bit for each subwidget.

   The above is done for every widget in the van Emde Boas tree when building it recursively. For simplicity, assume that $u^{1/3}$ and $u^{2/3}$ are both integers.

   (i) [6 marks] To perform `Member`, `Successor`, `Insert` and `Delete` operations over this variant, what changes need to make to the pseudocode shown in class?

   (ii) [9 marks] Analyze the running time of each operation over this variant.