Universität Siegen
Lehrstuhl Theoretische Informatik
Markus Lohrey

Exercise 3

Task 1
Sort the array $[3, 19, 8, 4, 13, 7, 29, 1]$ using Quicksort (with median-out-of-three).

Solution
quicksort$(1, 8)$: $p = 1$ (index of median of $A[\ell]$, $A[(\ell + r) \text{ div } 2]$, $A[r]$), partition$(1, 8, 1)$: swap$(1, 8)$ (pivot), swap$(1, 1)$, swap$(2, 8)$ (pivot), $[1, 3, 8, 4, 13, 7, 29, 19]$, $m = 2$

- quicksort$(1, 1)$
- quicksort$(3, 8)$: $p = 5$, partition$(3, 8, 5)$: swap$(5, 8)$ (pivot), swap$(3, 3)$, swap$(4, 4)$, swap$(5, 6)$, swap$(6, 8)$ (pivot), $[1, 3, 8, 4, 7, 13, 29, 19]$, $m = 6$
  - quicksort$(3, 5)$: $p = 5$, partition$(3, 5, 5)$: swap$(5, 5)$ (pivot), swap$(3, 4)$, swap$(4, 5)$ (pivot), $[1, 3, 4, 7, 8, 13, 29, 19]$, $m = 4$
    - quicksort$(3, 3)$
    - quicksort$(5, 5)$
  - quicksort$(7, 8)$: $p = 8$, partition$(7, 8, 8)$: swap$(8, 8)$ (pivot), swap$(7, 8)$ (pivot), $[1, 3, 4, 7, 8, 13, 19, 29]$, $m = 7$
    - quicksort$(7, 6)$
    - quicksort$(8, 8)$

Task 2 (Slides 53 and 58)
Show that for the n-th harmonic number H_n the following inequalities hold:

$$\ln(n + 1) \leq H_n \leq \ln(n) + 1.$$

Hint: $\ln(n) = \int_1^n \frac{1}{x} \, \mathrm{d}x$.

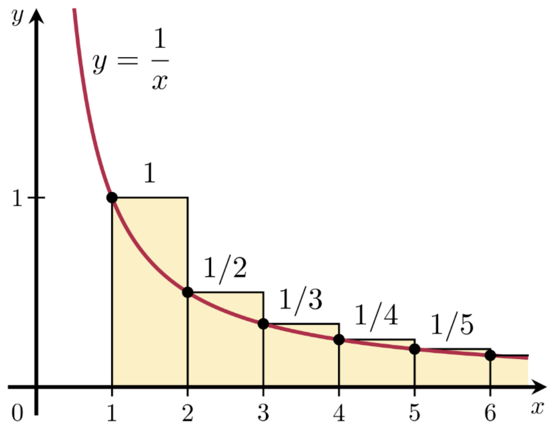
Solution
Since $\frac{1}{x}$ is monotonically decreasing, we have

$$\ln(n + 1) = \int_1^{n+1} \frac{1}{x} \, \mathrm{d}x = \sum_{k=1}^{n} \int_k^{k+1} \frac{1}{x} \, \mathrm{d}x \leq \sum_{k=1}^{n} \frac{1}{k} = H_n$$

and

$$H_n - 1 = \sum_{k=2}^{n} \frac{1}{k} \leq \sum_{k=2}^{n} \int_{k-1}^{k} \frac{1}{x} \, \mathrm{d}x = \int_1^n \frac{1}{x} \, \mathrm{d}x = \ln(n).$$

This picture illustrates the first inequality:



Source: Wikipedia

The second inequality is a similar picture, with the only difference that the bars are strictly left of the red curve.

Task 3
Sort the array

$$[7, 3, 8, 1, 5, 2, 4, 6]$$

using Standard Heapsort and then sort it using Bottom-up Heapsort. How many comparisons do you need in each case?

Solution
build-heap(8) [10 comparisons]

- reheap(4,8)
  - swap(4,8)
  - reheap(8,8)
- reheap (3,8)
- reheap (2,8)
  - swap(2,4)
  - reheap (4,8)
- reheap (1,8)
  - swap(1,3)
  - reheap (3,8)

Array after build heap: $[8, 6, 7, 3, 5, 2, 4, 1]$.

Standard Heapsort [17+10=27 comparisons]:

- build-heap(8)
- swap(1,8); reheap(1,7)
  - swap(1,3); reheap(3,7)
    - swap(3,7); reheap(7,7)
- swap(1,7); reheap(1,6)
  - swap(1,2); reheap(2,6)
    - swap(2,5); reheap(5,6)
- swap(1,6); reheap(1,5)
  - swap(1,2); reheap(2,5)
    - swap(2,4); reheap(4,5)
- swap(1,5); reheap(1,4)
  - swap(1,3); reheap(3,4)
- swap(1,4); reheap(1,3)
  - swap(1,2); reheap(2,3)
- swap(1,3); reheap(1,2)
  - swap(1,2); reheap(2,2)
- swap(1,2); reheap(1,1)

Bottom-up Heapsort [14+10=24 comparisons]:
Since there is no pseudocode here, we informally define

- sink-path$(1, i)$ to be the function, which computes the sink path of A[1] in the array $A[1, \ldots, i]$ $(i > 1)$,
- comp(i, j) to be the function, which compares $A[i]$ and $A[j]$,
- cyclic$(i_1, \ldots, i_k)$ to be the functions, which performs a cyclic rotation of the elements $A[i_1], \ldots, A[i_k]$. Clearly cyclic(i, j) = swap(i, j).

With these functions, the algorithm works as follows (basically we save 3 comparisons at the beginning of the algorithm, where the sink path has length 2):

- build-heap(8)
- swap(1,8); sink-path(1,7)
  - comp(1,7); cyclic(1,3,7)
- swap(1,7); sink-path(1,6)
  - comp(1,5); cyclic(1,2,5)
- swap(1,6); sink-path(1,5)
  - comp(1,4); cyclic(1,2,4)
- swap(1,5); sink-path(1,4)
  - comp(1,3); swap(1,3)
- swap(1,4); sink-path(1,3)
  - comp(1,2); swap(1,2)
- swap(1,3); sink-path(1,2)
  - comp(1,2); swap(1,2)
- swap(1,2)

Task 4
Show Jensen's inequality (slide 8).

Solution
Let $f\colon D \to \mathbb{R}$ with $D \subseteq \mathbb{R}$. The function f is convex if for all $x, y \in \mathbb{R}$ and all $0 \le \lambda \le 1$

$$f(\lambda x + (1-\lambda)y) \le \lambda f(x) + (1-\lambda) f(y)$$

Let $n \ge 2$, $x_1, \ldots, x_n \in D$, $\lambda_1, \ldots, \lambda_n \ge 0$ and $\lambda_1 + \cdots + \lambda_n = 1$. We prove that

$$f\left(\sum_{i=1}^{n} \lambda_i \cdot x_i\right) \le \sum_{i=1}^{n} \lambda_i \cdot f(x_i)$$

In case $n = 2$, since $\lambda_1 + \lambda_2 = 1$, we have $\lambda_2 = 1 - \lambda_1$. So we obtain

$$\begin{aligned} f(\lambda_1 \cdot x_1 + \lambda_2 \cdot x_2) &= f(\lambda_1 \cdot x_1 + (1-\lambda_1) \cdot x_2) \\ &\le \lambda_1 f(x_1) + (1-\lambda_1) f(x_2) \\ &= \lambda_1 f(x_1) + \lambda_2 f(x_2). \end{aligned}$$

Let $n > 2$. We assume that the statement holds for n and show it for $n+1$. We assume that $\lambda_{n+1} > 0$ (the case $\lambda_{n+1} = 0$ is trivial) and $\lambda_{n+1} \neq 1$ (otherwise all other λ_i would be 0). Then we can write

$$\sum_{i=1}^{n+1} \lambda_i x_i = \lambda_{n+1} x_{n+1} + (1-\lambda_{n+1}) \sum_{i=1}^{n} \frac{\lambda_i}{1-\lambda_{n+1}} x_i.$$

This allows us to use the fact that f is convex:

$$\begin{aligned} f\left(\sum_{i=1}^{n+1} \lambda_i \cdot x_i\right) &= f\left(\lambda_{n+1} x_{n+1} + (1-\lambda_{n+1}) \sum_{i=1}^{n} \frac{\lambda_i}{1-\lambda_{n+1}} \cdot x_i\right) \\ &\leq \lambda_{n+1} f(x_{n+1}) + (1-\lambda_{n+1}) f\left(\sum_{i=1}^{n} \frac{\lambda_i}{1-\lambda_{n+1}} x_i\right) \\ &\leq \lambda_{n+1} f(x_{n+1}) + (1-\lambda_{n+1}) \sum_{i=1}^{n} \frac{\lambda_i}{1-\lambda_{n+1}} f(x_i) \\ &= \sum_{i=1}^{n+1} \lambda_i f(x_i) \end{aligned}$$

To show the statement for concave functions, only replace $\leq$ by $\geq$.

Task 5 (More harmonic numbers)
Show the following 2 statements by induction.

(a) $\sum_{k=1}^{n} H_k = (n+1)H_n - n$

(b) $\sum_{k=1}^{n} H_k^2 = (n+1)H_n^2 - (2n+1)H_n + 2n$

Solution
The base case is trivial for both statements, hence we proceed with the induction steps.

(a) We have

$$\sum_{k=1}^{n+1} H_k = \sum_{k=1}^{n} H_k + H_{n+1} \stackrel{\text{IH}}{=} (n+1)H_n - n + H_{n+1} = (n+1)H_n + 1 + H_{n+1} - (n+1)$$

and we can rewrite $(n+1)H_n + 1$ as $(n+1)H_n + \frac{n+1}{n+1} = (n+1)H_{n+1}$. This yields

$$\sum_{k=1}^{n+1} H_k = (n+1)H_{n+1} + H_{n+1} - (n+1) = (n+2)H_{n+1} - (n+1).$$

(b) This time we start with the right-hand side.

$$
\begin{aligned}
&(n+2)H_{n+1}^2 - (2n+3)H_{n+1} + (2n+2)\\
=&(n+1)H_{n+1}^2 - (2n+1)H_{n+1} + 2n + H_{n+1}^2 - 2H_{n+1} + 2\\
=&(n+1)H_n^2 + 2H_n + \frac{1}{n+1} - (2n+1)H_n - \frac{2n+1}{n+1} + 2n + H_{n+1}^2 - 2H_{n+1} + \frac{2n+2}{n+1}\\
=&(n+1)H_n^2 - (2n+1)H_n + 2n + H_{n+1}^2 + 2H_n + \frac{2}{n+1} - 2H_{n+1}\\
=&(n+1)H_n^2 - (2n+1)H_n + 2n + H_{n+1}^2\\
\overset{\text{IH}}{=}&\sum_{k=1}^{n} H_k^2 + H_{n+1}^2 = \sum_{k=1}^{n+1} H_k^2
\end{aligned}
$$

The second = makes use of the equality $H_{n+1}^2 = H_n^2 + 2\frac{H_n}{n+1} + \frac{1}{(n+1)^2}$.