University of Waterloo
CS 466 — Advanced Algorithm
Spring 2013
Problem Set 10
Siwei Yang - 20258568

1. [14 marks: Find second largest value]

   Take a random sample, R, of r of the values; $r = \frac{n}{\lg n}$ (Assume r is a power of 2). By repeatedly pairing the elements that have won every comparison, find the maximum of the sample (in r-1 comparisons). Initialize m_1 as the sample maximum and m_2 as the element that was the maximum of half the sample but lost to the maximum in the last comparison.

   Scan through the remaining elements (i.e. those not in R), comparing each to m_2 and if it is larger than m_2 compare it with m_1. Update m_1 to be the maximum seen so far and m_2 to be the second largest (perhaps excluding some of the values from the R). When finished the scan, m_1 is the maximum of the set; under some condition, you may know m_2 is the second largest value in the set. If you cannot guarantee m_2 is the second largest, find the maximum of m_2 and the random sample elements that lost directly to the original m_1.

   - [2 marks]What is the maximum number of comparisons this method could use?

     The maximum number of comparisons happens when all remaining elements won m_2 and lose to m_1:

     $$\begin{aligned} C(n) &= (r-1) + 2 * (n-r) + \lg r \\ &= 2 * n - r - 1 + \lg r \\ &= 2 * n - \frac{n}{\lg n} + \lg n - \lg\lg n - 1 \end{aligned}$$

   - [4 marks]What condition, detectable in this algorithm, would guarantee that m_2 is the second largest after finishing the scan of all elements? What is the probability that this condition will hold?

If there exists an element in the remaining elements that wins both m_1 and m_2, then the m_2 is guaranteed the second largest after finishing the scan.

And this situation can be translated to that **the maximum element exists in the remaining element**, which has a probability of:

$$1 - \frac{1}{\lg n} \tag{1}$$

- [2 marks]If the condition above is not satisfied, how many more comparisons are required?

  If the condition is not satisfied, then we need to find the maximum between all elements lost to m_1 but haven't lost to m_2. That takes comparions:

  $$\lg n - \lg \lg n \tag{2}$$

- [4 marks]What is the expected number of times m_2 is replaced (i.e. the number of times a second comparison is done) in scanning the $n - \frac{n}{\lg n}$ non-sample elements?

  Given that not element in the remaining part is the global maximum, the expected number of times m_2 is replaced becomes the expected number of times an element is larger than all preceeding it with a twist on the start. Since m_2 is the maximum of $\frac{r}{2}$ elements, and there are $n-r$ remaining elements, we need to consider a list of number with size $n - \frac{r}{2}$ where the first $\frac{r}{2}$ elements has a maximum of m_2.

  Then we know the expected number of times an element is larger than all preceeding it is:

  $$\sum_{i=1}^{n-\frac{r}{2}} \frac{1}{i} = H(n - \frac{r}{2}) \tag{3}$$

  But the first $\frac{r}{2}$ elements does not contribute to the replacement numbers, thus the accounted number is:

  $$\sum_{i=\frac{r}{2}+1}^{n-\frac{r}{2}} \frac{1}{i} = H(n - \frac{r}{2}) - H(\frac{r}{2}) \tag{4}$$

- [4 marks]Given the probability of the condition being satisfied and the expected number of comparisons used in that case, and the probability of it not being satisfied and the number of comparisons used in the latter case, what is the expected number of comparisons used by the method.