# The Weighted Median Problem

**Solution to Problem 10-2 in CLR**

**(a)** This is trivial.

**(b)** Sort the $\{x_i\}_{i=1}^n$. Suppose that this leaves them in the order $x_{\pi(1)}, x_{\pi(2)}, \ldots, x_{\pi(n)}$ (so $\pi$ is a permutation of the numbers $\{1, 2, \ldots, n\}$). Now, traverse the numbers, left-to-right, and sum the $w_{\pi(i)}$ as you proceed. The smallest $k$ for which $\sum_{i=1}^k w_{\pi(i)} \geq 1/2$ is the weighted median. Clearly, this procedure takes $O(n \log n)$ time using appropriate sorting algorithm.

**(c)** Using the SELECT algorithm (with $i = \lfloor \frac{n+1}{2} \rfloor$ as statistic order), we can find the unweighted median of $n$ numbers in linear time. We will now, however, solve a more general problem. For a collection $\{x_1, x_2, \ldots, x_n\}$ and positive weights $\{w_1, w_2, \ldots, w_n\}$, a $\delta$-median is an element $x_m$ such that

$$\sum_{x_i < x_m} w_i < \delta \quad \text{and} \quad \sum_{x_i > x_m} w_i \leq W - \delta$$

where $\sum_{i=1}^n w_i = W$. Observe that this generalizes the notion of weighted median (with the private case of $\delta = 1/2$). We present a linear time algorithm to compute the weighted median (notice that the $\delta$-median can always be found in $O(n \log n)$ time by sorting).

WMEDIAN $(\{x_i\}_{i=1}^n, \delta)$

1. If $n < 10$ then sort the $\{x_i\}$ and return the weighted median.
2. Find the median $x_m$ of the set $\{x_1, x_2, \ldots, x_n\}$.
3. Partition (using PARTITION) the set around $x_m$, creating two sets $A_0$, consisting of all elements of $\{x_i\}$ which are smaller than or equal to $x_m$ and $A_1$, consisting of all elements which exceed $x_m$. Let $S_0$ be the sum of the weights of all elements in $A_0$.
4. If $\delta \leq S_0$, then return WMEDIAN$(A_0, \delta)$. Otherwise return WMEDIAN$(A_1, \delta - S_0)$.

Correctness of the above algorithm follows in the same fashion as the correctness for SELECT. Observe that the total running time is

$$T(n) \leq T(\lceil n/2 \rceil) + O(n)$$

which is, according to the Master Theorem $O(n)$, as desired.