Exercise 4.3

If two routes separate that means that at a common kth bit they chose different ways. This again means that they must have different goals, differing on the kth bit. Since bit correction is going from left to right, these routes will never rejoin, since they will differ on this kth bit until their goals are reached.

Problem 4.14

Let us look at the sorting tree of some execution of the *RandQS* algorithm on some set S. Say that at some node the algorithm chooses a *good pivot* provided that it selects an element from its given subset $S' \subset S$ that is results in two subsets (elements of higher resp. lower order) with size no larger than $3/4|S'|$. We claim that the chance of this selection by the algorithm is $\geq 1/2$. To see this note that the number of elements in the *lower* subset L follows a uniform distribution on $\{0, \cdots, n-1\}$, where $n = |S'|$. The number of elements in the *higher* subset H depends entirely on $|L|$ in that $|L| + |H| = n - 1$. The chance that a bad pivot is chosen is therefore

$$\begin{aligned} P(|L| > 3/4n \vee |H| > 3/4n) &= 2P(|L| > 3/4n) \\ &= 2P(|L| \geq \lceil 3/4n \rceil) \\ &= 2\left(\frac{n - 1 - \lceil 3/4n \rceil + 1}{n}\right) \\ &\leq 2\frac{1}{4} = 1/2 \end{aligned}$$

Thus the chance of a good pivot chosen is $\geq 1/2$. Fix an element $x \in S$ and let h denote the length of the path from the root of the tree to x. For each step on this path if a good pivot is chosen the number of remaining elements in the batch containing x is reduced by at least $3/4$. Solving

$$(3/4)^k n \leq 1 \iff k + \log_{3/4}(n) \leq 0 \iff k \geq \log_{4/3}(n)$$

we see that if more than $m = \log_{4/3}(n) = \ln(n)/\ln(4/3)$ good pivots is chosen the height of x has been attained. Now we calculate the chance that less than m good pivots are chosen in $8m$ steps. The choices of pivots are independent and each is good with probability at least $1/2$. By Chernoff with $\delta = 3/4$ and $\mu = 1/2m$ the probability of having too few good pivots is less than

$$e^{-\frac{m}{2}\left(\frac{3}{4}\right)^2/2} \leq e^{-\frac{m}{8}} = e^{-\frac{\ln(n)}{\ln(4/3)}} = n^{-\frac{1}{\ln(4/3)}} \leq n^{-3}$$

Thus for any $x \in S$ the probability that x has height greater than $\frac{8}{\ln(4/3)}\ln(n) = h_m$ is less than n^{-3}. By union bounding the chance that the whole tree has height more than h_m is less than n^{-2}. The sum of the heights of all elements equal the number of comparisons of the algorithm. Therefore the number of comparisons is at most $\frac{8}{\ln(4/3)} n \ln(2)$ with probability $1 - n^{-2}$ proving that *RandQS* is $O(n \ln(n))$ with high probability.