day 21 4/1/98 "I knew there should have been another step in here."

1 Recurrence relations

Running time of merge sort:

$T(n) = 2T(\frac{n}{2}) + n$ with $T(1) = 0$

Assume we want $T(n) = O(n)$

Then $T(n) = 2O(\frac{n}{2}) + n = O(n)$? Wrong! Bad!

Instead, try: $T(n) = cn$ then $T(n) = 2c\frac{n}{2} + n = (c+1)n$ which doesn't work.

try: $T(n) = cn \log n$ then $T(n) = 2c\frac{n}{2} \log \frac{n}{2} + n = cn \log \frac{n}{2} + n = cn \log n - cn + n$ which is solvable for $c = 1$.

2 All Prefix scan

The Random-Mate algorithm essentially has every node in the list guess it's parity. The algorithm (covered last meeting) works in 2 phases. In the first phase, nodes are eliminated while in the second phase, you back up the tree filling in the values of each node as in Wiley's algorithm.

2.1 Timing analyses

Probability that a node is removed from the list at each stage in the forward algorithm is $\frac{3}{4}$. So $P_t = (\frac{3}{4})^t =$ probability that a node is still in the list after t rounds. If we want $P_t = \frac{1}{n^2}$ then $t = 2 \log_{\frac{4}{3}} n$.

The probability that no node is left in the list is: $P_{left} \leq \sum_{i=1}^{n-1} P_t \leq nP_t \leq \frac{1}{n}$ if $t = 2 \log_{\frac{4}{3}} n$

2.2 Work analyses

Because of the random nature of this algorithm it is difficult to group enough nodes on each processor s.t. every processor will always have work to do. A prescan can be done over the elements to reallocate their location at each step.

3 Another random list ranking algorithm

1. Make $\frac{n}{\log n}$ queues of size $\log n$
2. Set the sex of all nodes to M.
3. Reset sex of top in each queue to a random sex.
4. If top element is F and points to M, the splice out F

The algorithm works on a doubly linked list.

3.1 Timing analyses

Timing is $O(\log n)$ with high probability

Let $P(head) = \frac{1}{4}$ and $P(tail) = \frac{3}{4}$

Let $S_t^p = \#head$ in t trials.

$P(S_t^p < \log n)$ can be bounded by Chernoff Bounds

3.1.1 Chernoff Bound

$P(S_t^p < (1-\beta)pt) < e^{-\beta^2 \frac{pt}{2}}$ for $0 \le \beta \le 1$

In particular, $P(S_{16\log n}^{\frac{1}{4}} < (\frac{1}{4})4\log n) < e^{-\frac{3^2}{4^2}\frac{4\log n}{2}} = e^{-\frac{9}{8}\log n} = \frac{1}{n^{\frac{9}{8}}} < \frac{1}{n}$

$=$ probability that a processor is not done.

The probability that all processors is done is $\le \sum_{i=1}^{n\log n} \frac{1}{n} = \frac{n}{\log n}\frac{1}{n} = \frac{1}{\log n}$.