95.4804W2004 Mid-Term Exam: Duration 1h20m

Write your name and student number on and answer all questions in the answer booklet provided.

1 Passing Students

In a certain fourth-year course at a certain university, the expected number of students that successfully complete the course is $n/2$, where n is the number of students in first year. Some years the professor is harder than others and students sometimes collaborate, so the event "Student A successfully completes the course" is not necessarily independent of the event "Student B successfully completes the course."

1. Give an upper bound on the probability that more than $3n/4$ students successfully complete the course.

2. Give an upper bound on the probability that less than $n/4$ students successfully complete the course.

2 Searching an Unsorted Array

Consider the problem of searching an unordered array $A[1], \ldots, A[n]$ for an element x. The only comparison operation you can do is to test if $x = A[i]$ for any $1 \le i \le n$. (Since the array is unordered this is the only useful kind of comparison anyway.)

1. Show that there exists a randomized algorithm that uses only 1 random bit and performs an expected number of comparisons that is at most $(n+1)/2$.

3 Finding large Non-Adjacent Sets in Parallel

Let $L = l_0, \ldots, l_{n+1}$ be a list of items. We pick a subset of items in L using the following algorithm: For each element l_i $(1 \le i \le n)$, we toss a fair coin. We say that l_i is included in the subset if l_i's coin toss came up heads and both its neighbours' (l_{i-1} and l_{i+1}) coin tosses came up tails.

1. What is the probability that l_i is included in the independent set. Using this, compute the expected size of the independent set.

2. Let E_i be the event "l_i is included in the independent set." Are E_i and E_{i+1} independent?

3. Are E_i and E_{i+2} independent?

4. Are E_i and E_{i+3} independent?

5. Use Chernoff's bounds to show that, with very high probability, this algorithm produces an independent set of size at least $(1-\epsilon)n/24$.

4 Lazy Deletion Revisited

Recall the lazy deletion data structure in assignment 3. (When we delete an item we simply mark it as deleted, when more than $n/2$ items are marked we rebuild a whole new data structure; insertion and searching take $O(\log n)$ time, rebuilding takes $O(n)$ time.)

1. Suppose that, when we rebuild the data structure we only remove half the marked items (so afterwards the data structure has size $3n/4$ and contains $n/4$ marked items). Show that the amortized cost of deletion is still $O(\log n)$.