Randomized Algorithms, WS 2002/03
Problem Set 4
Dec. 6, 2002

Lecturers:
Privatdozent Dr. Peter Sanders, Building 46, R.308, sanders@mpi-sb.mpg.de
Dr. Ernst Althaus, Building 46, R.304, althaus@mpi-sb.mpg.de
Home page: `http://www.mpi-sb.mpg.de/~sanders/courses/randalg/index.html`

Try to solve (at least) two of the following three exercises
This homework is due on Friday, Jan. 10.

Problem 4-1 (Set Coloring)
Consider a base set $S = \{x_1, \ldots, x_n\}$ and subsets $S_1, \ldots, S_k$ of size r with $k \leq 2^{r-2}$. The problem is to color each element of S red or green such that all S_i contain both a red and a green element. Develop a randomized Las Vegas style algorithm to solve this problem and bound its expected running time. (A (pseudo) application of this could be called party balancing: Suppose S in a coalition bargaining between two parties there are a set of political jobs to fill (minister of x, state secretary of $y, \ldots$). The S_i stand for a number of committees whose memberships are coupled to the jobs. Now jobs must be assigned to the parties such that no committee has solely members of a single party.)

Problem 4-2 (Quick Sort, again)
Show that the Quick-Sort algorithm of the lecture runs in time $O(n \log n)$ with high probability ($\geq 1 - 1/n$).

Hint: The leaves of the recursion tree correspond to the elements to be sorted. Show that for each element the corresponding leave is at height $O(\log n)$ with high probability.

Problem 4-3 (Sharp Concentration for the Coupon Collectors Problem)
Show that with high probability ($\geq 1 - 1/n$), $O(n \log n)$ balls suffice to leave no empty bins. We propose three methods. We ask you to choose one and explain why you do not use the other two. As far as we know one method yields an easy somewhat loose bound, another a bound with better constant factors but is more difficult computationally, whereas the remaining one does not seem to work at all. Here are the methods in random order.

- Let T_j denote the time needed to fill the next bin when $j-1$ bins are already filled. Exploit that the T_j are independent geometrically distributed random variables. Develop a Chernoff type tailbound for $\sum_j T_j$.
- Use the method of bounded differences.
- Bound the probability that a fixed bin j remains empty. Use this bound to bound the probability that some bin remains empty.