COMP4804 Midterm Exam — Duration 1h20m **Winter 2006**

Answer all questions in your exam booklet.

Chernoff: For a binomial(p,n) random variable B, $\Pr\{B \geq (1+\epsilon)np\} \leq e^{-\epsilon^2 np/3}$ and $\Pr\{B \leq (1-\epsilon)np\} \leq e^{-\epsilon^2 np/2}$.

1. **[3 marks]** Suppose we store n distinct elements $x_1, \ldots, x_n$ using an array of m lists $L_1, \ldots, L_m$ in the following way. For each x_i we have a *hash value* $h(x_i) \in \{1, \ldots, m\}$ and we store x_i in the list $L_{h(x_i)}$. The hash funcion h is chosen in such a way that, for $x \neq y$,

$$\Pr\{h(x) = h(y)\} \leq 1/m \ .$$

   (a) Consider a value $x \notin \{x_1, \ldots, x_n\}$. Use indicator variables to find the best upper bound possible on the expected size of the list $L_{h(x)}$. That is, prove that $\mathrm{E}[|L_{h(x)}|] \leq$ *blah* for the appropriate value *blah*

   (b) Consider a value $x_i \in \{x_1, \ldots, x_n\}$. Give the best upper bound you can on $\mathrm{E}[|L_{h(x_i)}|]$.

   (c) For a particular list L_j and a particular number $d > 0$, give the best upper bound you can on $\Pr\{|L_j| \geq d\}$.

2. **[3 marks]** Suppose you have a Monte-Carlo algorithm $\mathcal{A}$ for testing whether a graph has some property $\mathcal{P}$. When we run $\mathcal{A}$, it runs in $O(n+m)$ time and produces the correct answer (yes or no) with probability 5/9.

   (a) If we run $\mathcal{A}$ k times, what is the expected number of times $\mathcal{A}$ produces the correct answer?

   (b) Give a tight upper bound on the probability that $\mathcal{A}$ produces the correct answer fewer than $k/2$ times.

   (c) Describe how, using $\mathcal{A}$ as a subroutine, we can get a Monte-Carlo algorithm that runs in $O(k(n+m))$ time and produces the correct answer with probability at least $1 - e^{-\Omega(k)}$.

3. **[2 marks]** Consider an algorithm that works with a list L and runs in m rounds. During round i, the algorithm either appends one element to L (at a cost of $C_i = O(1)$) or deletes some number, $0 \leq k_i \leq |L|$, of elements from L (at a cost of $C_i = O(1 + k_i)$).

   (a) Define a non-negative potential function $\Phi(L)$ and use it show that the amortized cost of the ith round is $O(1)$.

4. **[2 marks]** Let b_1, b_2, m_1 and m_2 be any real numbers such that $m_1 \neq m_2$. Pick a random integer $r \in \{1, \ldots, k\}$.

   (a) Give an upper bound on $\Pr\{rm_1 + b_1 = rm_2 + b_2\}$.

   (b) Give an upper bound on $\Pr\{r^2 m_1 + rb_1 = r^2 m_2 + b_2\}$.