Data Structures and Algorithms (CS 130A) **Subhash Suri**

Homework Assignment 1

Handed Out: Jan 15 **Due:** *Jan 29*

Please typeset your answers. Handwritten answers are extremely difficult for TAs to comprehend.

1. (20 pts) Prove the following identities.
   (a) $\sum_{i=1}^{n}(2i-1) = n^2$.
   (b) $\sum_{i=1}^{n} i^3 = (\sum_{i=1}^{n} i)^2$.

2. (20 pts) Given an array A of n positive numbers, describe an $O(n)$ time algorithm for each of the following problems. *A correct solution should have three parts: a clear and concise description of the algorithm; a proof of correctness; and time complexity analysis.* Find the indices i and j, with $j \geq i$, such that
   (a) $A[j] + A[i]$ has the maximum value,
   (b) $A[j] - A[i]$ has the maximum value,
   (c) $A[j] \times A[i]$ has the maximum value,
   (d) $A[j]/A[i]$ has the maximum value.

3. (20 pts) What is the expected number of empty locations in a hash table of size m when n keys are hashed into it? Derive an exact formula.

4. (40 pts) If we hash a set S of n keys into a table of size n with a universal hash function h, what is the *expected maximum number* of keys that collide? In other words, what is the maximum number of keys that are expected to hash to the same location? We break down this computation into a sequence of easier steps, as follows.

   Let A_j be the event that at least one slot in the hash table has $\geq j$ keys. We compute the largest j for which $Prob[A_j] \leq 1/2$; that j is our answer. Calculating A_j directly is not straightforward, so we proceed as follows. *In all cases, explain your reasoning.*

   (a) Let A_j^1 be the event that the table slot 1 gets $\geq j$ keys under h. Supposing you know $Prob[A_j^1]$, give an upper bound on $Prob[A_j]$.
   (b) Let B be the event that a *fixed subset* $C \subset S$ of size $|C| = j$ hashes into slot 1. That is, each key of C maps to slot 1 under h. Calculate the probability $Prob[B]$.
   (c) Use $Prob[B]$ to get an upper bound on the probability $Prob[A_j^1]$.
   (d) Compute the largest value of j for which $Prob[A_j^1] \leq \frac{1}{2n}$. Explain how in combination with (a), this j is the expected maximum number of collisions.