UNIVERSITY OF TORONTO – CSC 373 – JUNE 28, 2007
TEST 1

Student Number: ____________________

Email Address: ____________________

Last (Family) Name(s): ____________________

First (Given) Name(s): ____________________

- This test is worth 15% of your final mark.
- Answer each question directly on the test paper, in the space provided. Use the reverse side of the pages for rough work. If you need more space for one of your solutions, use the reverse side of a page and *indicate clearly the part of your work that should be marked.*
- **20% rule:** If you are unable to answer (a part of) a question, you will get 20% of the marks for the (part of the) question if you write "I don't know" and *nothing else* for that part/question.

The Master Theorem

If for some $a, b, d > 0$:

$$T(n) = aT(n/b) + \mathcal{O}(n^d)$$

then

$$T(n) = \begin{cases} \mathcal{O}(n^d) & \text{if } a < b^d \\ \mathcal{O}(n^d \log(n)) & \text{if } a = b^d \\ \mathcal{O}(n^{\log_b(a)}) & \text{if } a > b^d \end{cases}$$

Question 1 (out of 10)	
Question 2 (out of 12)	
Total (out of 22)	

Question 1 [10]

Given a list of n distinct numbers, the **median** defined here to be the $\lceil n/2 \rceil$-th smallest value.

You are interested in analyzing some hard-to-obtain data from two separate databases. Each database contains n numerical values —so there are $2n$ values in total—and you may assume that no two values are the same. You'd like to determine the median of this set of $2n$ values (i.e., the n-th smallest value).

Suppose that the only way you can access these values is through *queries* to the databases. In a single query, you can specify a value k to one of the two databases, and the chosen database will return the k-th smallest value that it contains. Since queries are expensive, you would like to compute the median using as few queries as possible.

Give an algorithm that finds the median value using at most $\mathcal{O}(\log(n))$ queries. You can use the following fact:

Fact: Let $k = \lceil n/2 \rceil$. (Note that then $A(k)$ and $B(k)$ are the medians of the two databases, respectively.) Suppose that $A(k) > B(k)$, then the median of A and B is the same as the median of $A[1 \ \dots \ k]$ and $B[(\lfloor n/2 \rfloor + 1) \ \dots \ n]$.

Question 1 continues here

Question 2 [12]

Given a set of n items $\{1, \ldots, n\}$, where each item i has a nonnegative integer weight $w(i)$ and a nonnegative integer value $v(i)$ (for $1 \leq i \leq n$), and a nonnegative integer bound W. The Knapsack problem is to find a subset S of the items with maximum total value, whose total weight is $\leq W$. That is, $S \subseteq \{1, \ldots, n\}$ so that

$$\sum_{i \in S} w(i) \leq W$$

and $\sum_{i \in S} v(i)$ is maximum.

a [**3**] Define an array that you will use to solve this problem. Explain the meaning of each entry in the array.

b [**3**] Give the initialization and recurrence for computing the array.

c **[3]** Give a program that computes the array.

d **[3]** Give a program that compute an optimal set S for the Knapsack problem using the array defined in part **a**.