CS2223
HW#1

DUE: Thursday, November 4

1 (15 points) One algorithm to estimate the size, *n*, of a set of labelled objects is to select members of the set randomly, **with replacement**, until any object is selected a second time. If *k* distinct objects are drawn before the first duplicate, then we estimate *n* to be $2k^2/\pi$. Test this by writing and executing a program for the following algorithm:

Pick a fixed $n \geq 1$.
$S \leftarrow \varnothing$
$a \leftarrow$ random integer in the range [1,...,*n*]
repeat
    $S \leftarrow S \cup \{a\}$
    $a \leftarrow$ random integer in the range [1,...,*n*]
until $a \in S$
return $2|S|^2/\pi$

(*a*) What operations are performed on *S*?
(*b*) Describe a reasonable data structure to implement *S*.
(*c*) What is the worst-case time to perform each of the operations of part (*a*) for your data structure of part (*b*)? You should use Θ-notation.
(*d*) Time your program and try to determine the rate of growth of its execution time as a function of *n*. That is, for each of several values of *n*, execute and time your program. Try to express your program's execution time as a function of *n*.

Describe your implementation (the machine and compiler you are using). Submit a listing of your program, along with evidence that it executes correctly. The evidence should include your selected values of *n* and your estimates of *n*. Is your program's estimate of *n* reasonable?

2. (2 points) Do **Problem 2.6** from our text. In our text, "log" denotes the natural logarithm (to the base *e*=2.7182818...).