CMSC 341 Data Structures
Disjoint Set Review

October 26, 2010

1. Define $lg^*(N)$. What is the value of $lg^*(1024)$?
2. Define the Union-by-Weight heuristic.
3. Define the Path Compression Heuristic.
4. When both Union-by-Weight and Path Compression are used on disjoint sets with a total of N elements, a sequence of M union-find operations can be done in $O(M\, lg^*N)$ time. It is sometimes said that under these conditions, union-find is done in constant time per operation. What does this mean? Why is it true?
5. In an uptree with root x, let $R(x)$ be the length of the longest path and let N be the number of nodes (including x). Assuming the uptree was created by means of multiple union operations using the Union-by-Weight heuristic. Prove $R(x) \leq lgN$.
6. Perform the following Union-by-Weight operations on a set of 10 elements (0 - 9, each initially in their own set). Draw the forest of trees that result. U(1,5); U(3, 7); U(1, 4); U(5, 7); U(0, 8); U(6, 9); U(3, 9)
7. Although uptrees are used to conceptualize disjoint sets, disjoint sets are generally implemented in an array. Explain how this is possible.
8. Prove that if Union-by-Weight is used for all unions, the length of the deepest node is no more than $lg(N)$.

9. Given the following forest of up-trees

   (a) show the array which represents them

   (b) show the result of find(6), using Path compression