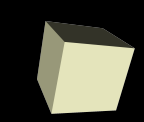
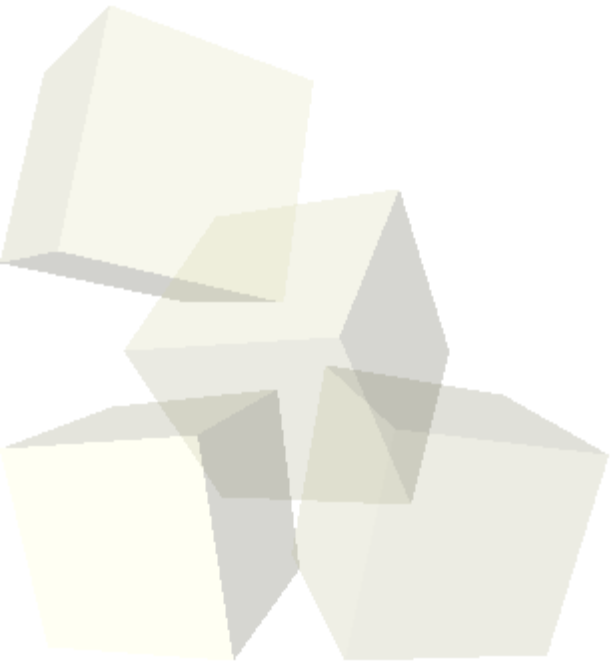


Augmented Data Structures

1-19-2005



Opening Discussion

- Do you have any questions about the reading?
- Do you have any questions about the test?
- Do you have any questions about your assignment? It will be due by midnight on Tuesday through electronic submission.

Augmenting Data Structures

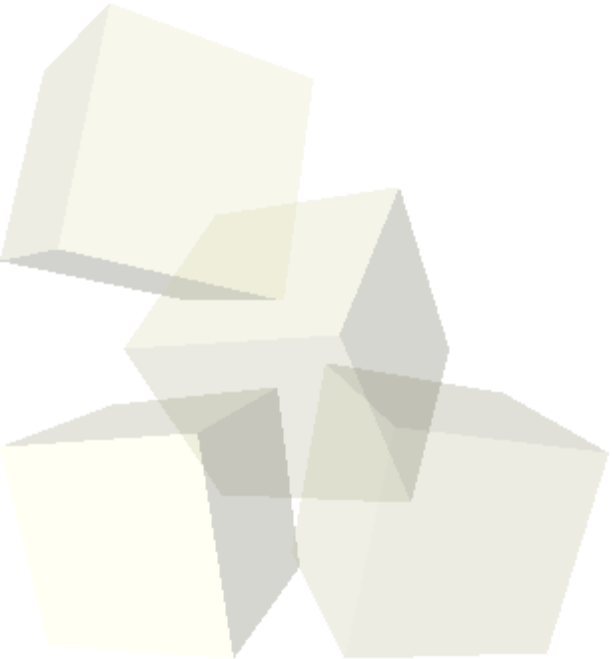
- The idea of augmenting data structures is fairly simple. We want to add data to the elements of our data structure that help us to quickly get to some type of information.
- Technically you can augment a data structure with whatever information you want, but certain things don't allow for the data structure to be efficiently implemented. In order to be practical, the augmented data must be updatable with the same order as the normal operations of the data structure.

Augmenting BSTs

- For BSTs, we can come up with a simple rule for what types of data we can augment the structure with.
- Any augmentation where the value of a node can be calculated from the values of the children will preserve the speed requirements.
- This holds because only O(log n) nodes are going to have one of their descendants changed.
- Most of the values that we typically talk about in relation to trees can be calculated this way which means we can augment the tree with them.

Order Statistic Trees

- Your first assignment has you writing an order statistic tree. These trees are augmented with the size of the node. This allows you to find a particular index in the tree in O(log n) time.
- Size is easily computed as the sum of the two children plus 1 so it fits our efficiency requirement.



Interval Trees

- CLR also discusses an interval tree which is a tree that is augmented to allow you to quickly look for intervals that overlap with a certain time.
- We sort the tree by the starting time of the interval and augment the nodes with the maximum end time below any given node.
- Given this you can always find an overlapping interval or say there are none in O(log n) time. If there are multiple overlapping intervals, finding all of them could take you to O(n), but that can't be avoided if the number of overlapping intervals is O(n).

Other Augmentations

- Augmentations that involve sums, minimums, and maximums can easily be put into trees.
- Often the most useful augmentations occur when we have data that is at least somewhat orthogonal to the data that the tree is sorted on. We see this in the order statistic tree where the end time is not strongly correlated to the start time.
- Augmentation could also be used for geometric computations where the augmented data includes bounds on a and y when the data is sorted in the tree by x.

Reminders

- Your first assignment is due on Tuesday. There is no reading for Tuesday though I might talk about divide and conquer a bit because our class will get a late start on Thursday as I agreed to give the math seminar.