CSC148H Week 7

Ilir Dema, Michael Miljanovic

Summer 2021

Motivating Trees

- A **data structure** is a way of organizing data
- Lists are linear data structures
  - They are linear in the sense that data is ordered (i.e. first piece of data is followed by second is followed by third ...)
- Underlying lists make sense for many applications:
  - Function calls in programs (stack)
  - A lineup in a bank (queue)
  - Event-handling based on timestamps (priority queue)

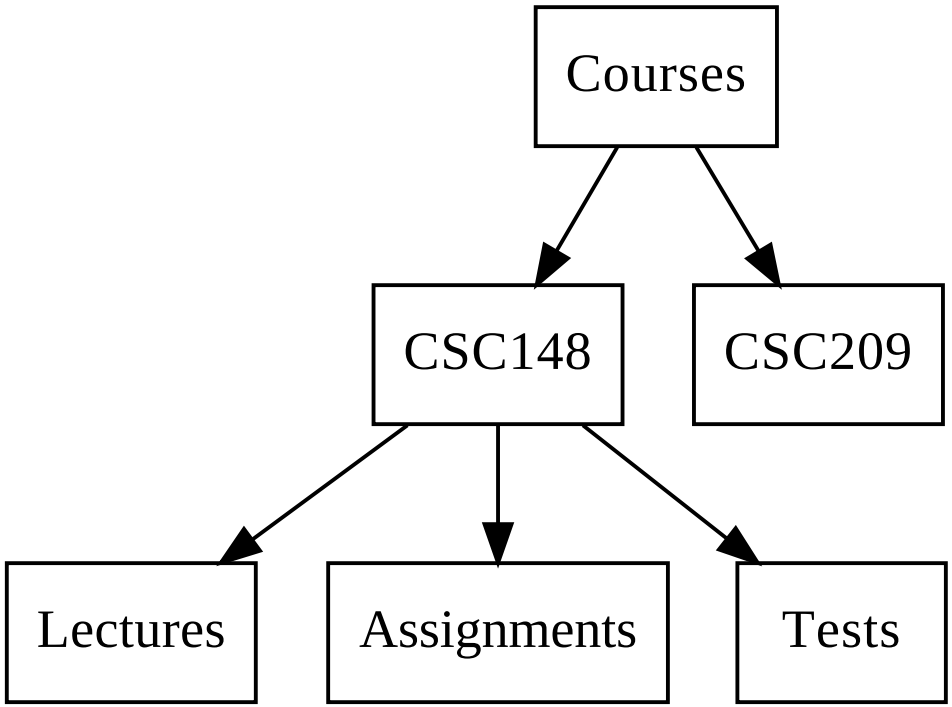
Motivating Trees...

- It doesn't make sense to organize certain types of data into a linear structure
- Consider directories in a file system. They have a natural hierarchical structure that is difficult to represent linearly
- If we want to use a list, we might try storing the root directory at the first position and its subdirectories and files to its right
- But how would we know when the files of a subdirectory end and we are back up one level?
- Other examples:
  - Structure of an HTML document
  - Structure of a Python program

Tree Definition

- A **tree** has a set of nodes (often with values or labels), and directed edges that connect nodes
- One specified node is the **root**
- Every node besides the root has exactly one parent

Sample Tree



- How many nodes are there? What are they?
- How many edges are there? What are they?
- What is the root?

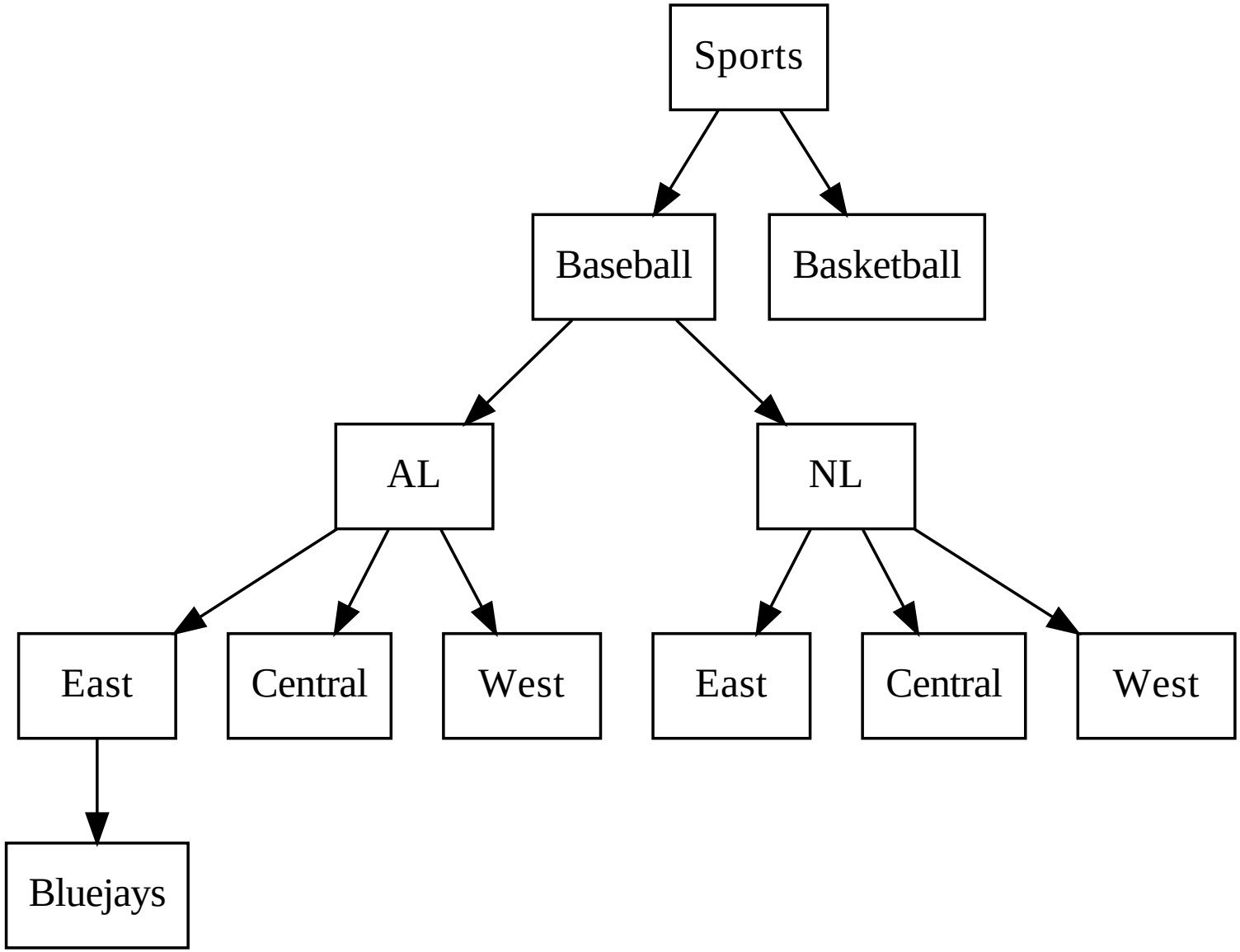
Tree Terminology

- Parent: a node is the parent of all nodes to which it has outgoing edges
- Siblings: set of nodes that share a common parent
- Leaf: node that has no children (i.e. no outgoing edges)
- Internal node: nonleaf node
- Path: sequence of nodes $n_1, n_2, \ldots, n_k$, where there is an edge from n_1 to n_2, n_2 to n_3, etc.
- Descendant: node n is a descendant of some node p if there is a path from p to n
- Subtree: a subtree of tree T is a tree whose root node r is a node in T, and which consists of all the descendants of r and the edges among them

Tree Terminology...

- Branching factor: maximum number of children of any node
- Length of a path: number of edges on the path
- Tree height: longest path from root to one of the leaves; count nodes
- Node's height is 1 plus maximum path length of tree rooted at that node
- Level (Depth): the level (or depth) of node n is the length of the path from the root node to n. The level of the root is typically defined to be 0 or 1

Another Sample Tree



What is the height of the tree? Branching factor? Depth of Baseball? Length of path from Sports to AL?

Worksheet 1

Worksheet 1, writing tree methods