Computational Structures in Data Science

UC Berkeley EECS
Lecturer Michael Ball

Lecture #20: Data Structures:

Trees

April 6, 2020

https://cs88.org

Updates

- Please checkout the polls on Piazza!
  - Are you taking CS88 P/NP?
  - When would you like the final?
    - We will have an alternate time for time zones
  - What do you want for a clobbering policy?

- As a reminder: Private piazza posts are best, since all the staff see them.

Why?

- **Trees represent lots of natural structures**
  - A boss who has employees report to them
  - Courses which belong to departments, and departments which colleges in a University
  - Anything with a hierarchy, really.
    - » A family tree
    - » Biological taxonomies (Kingdom, Phylum….)
- **Trees give us really cool approaches for “divide and conquer”**
  - Used in every computer to speed up searching for files
- **Another recursive data structure!**
  - We can keep practicing recursion and working with classes
- **Trees are a simplified form of a *graph*, a tool which can help us model just about anything.**

Searching Trees: Two Strategies

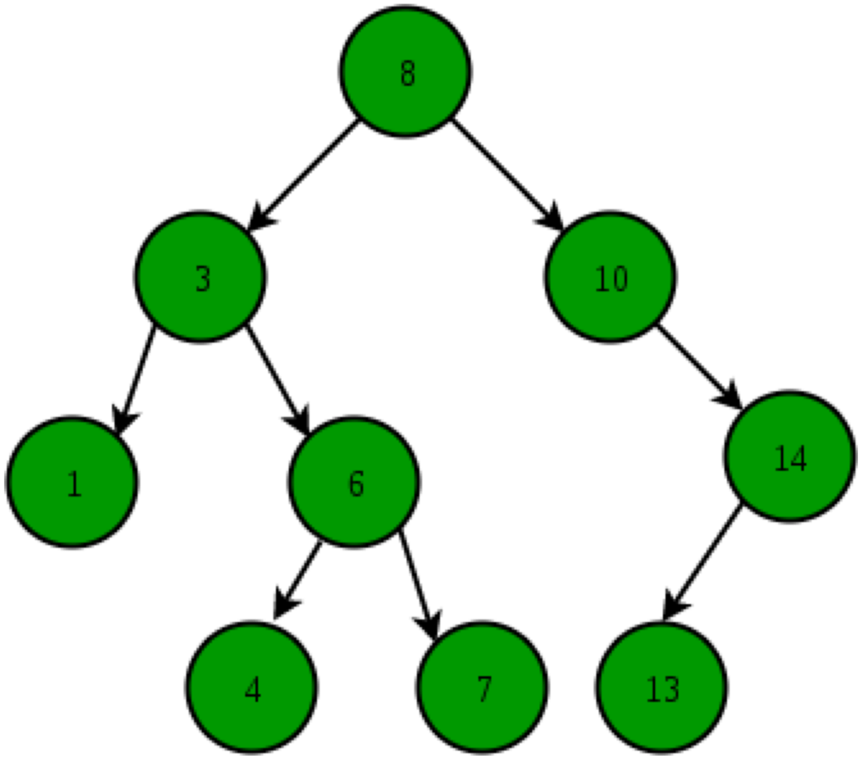
- The searching we have been doing today is called “Depth First Search”, or DFS.
- Recursion makes the algorithm very nice.
  - We always make a recursive call on the first branch
  - We continue recursing until there are no more branches
  - Then the function executes and we go back “up” a level and check out the next branch.
  - We sometimes say “popping up the stack”. The *stack* is the “stack of function calls” the computer uses to keep track of how things work, and you’ll learn about this in CS61B,

Searching by level

- What if I want to check out all the values of my branches before making a recursive call?
- We call this “Breadth First Search”
  - Start broad before going deeper.
- What if we said, you just can’t use recursion. (Sometimes, CS instructors do weird things like that…)
- This is used in practice for lots of cool things:
  - Shortest path between two items (more of a graph and not a tree, usually). Google Maps uses it for routing and the algorithms that power the internet use it.

Binary Search Trees

Notice how the tree is “ordered” such that the left is always less than the right side.


8
3
10
1
6
14
4
7
13