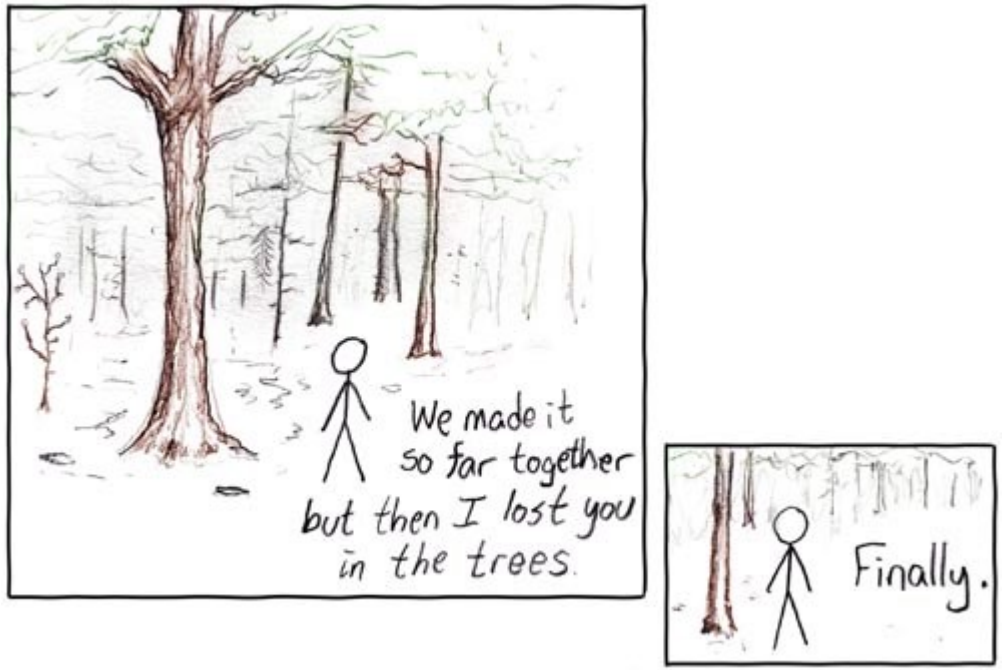
Lecture 04

Trees

The Tree Data Structure

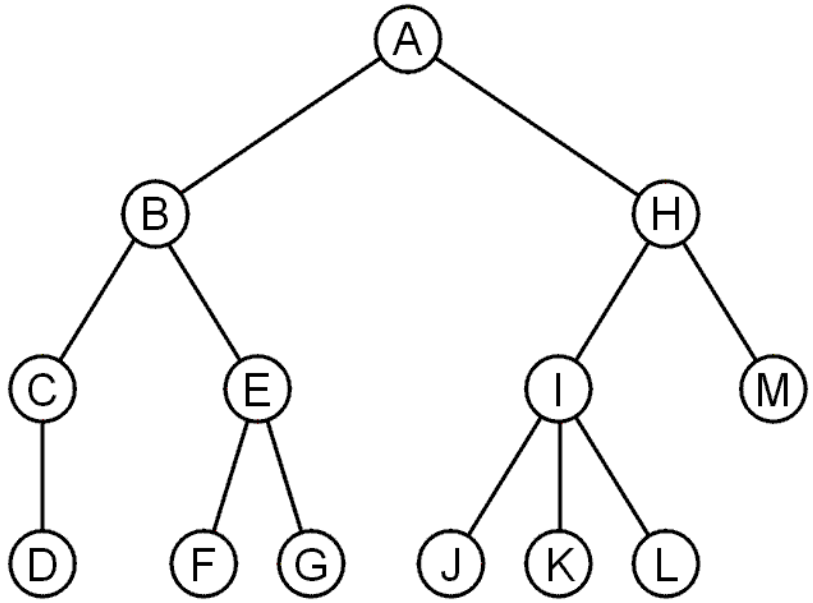
- Trees are the first data structure different from what you've seen in your first-year programming courses



Trees

A rooted tree data structure stores information in *nodes*

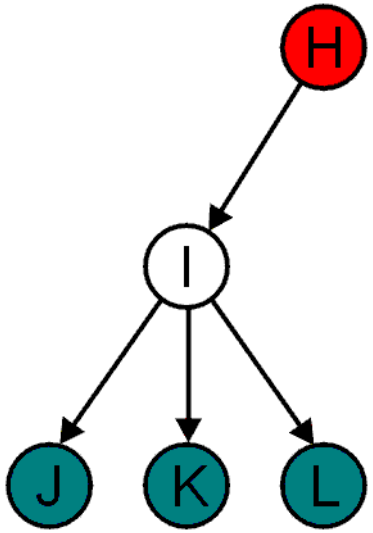
- Similar to linked lists:
  - There is a first node, or *root*
  - Each node has variable number of references to successors
  - Each node, other than the root, has exactly one node pointing to it



Terminology

- All nodes will have zero or more child nodes or children
  - I has three children: J, K and L

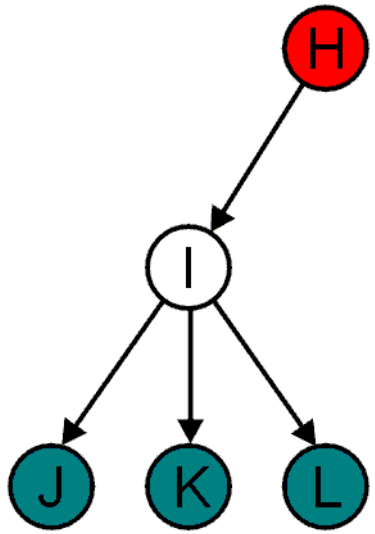
- For all nodes other than the root node, there is one parent node
  - H is the parent I



Terminology

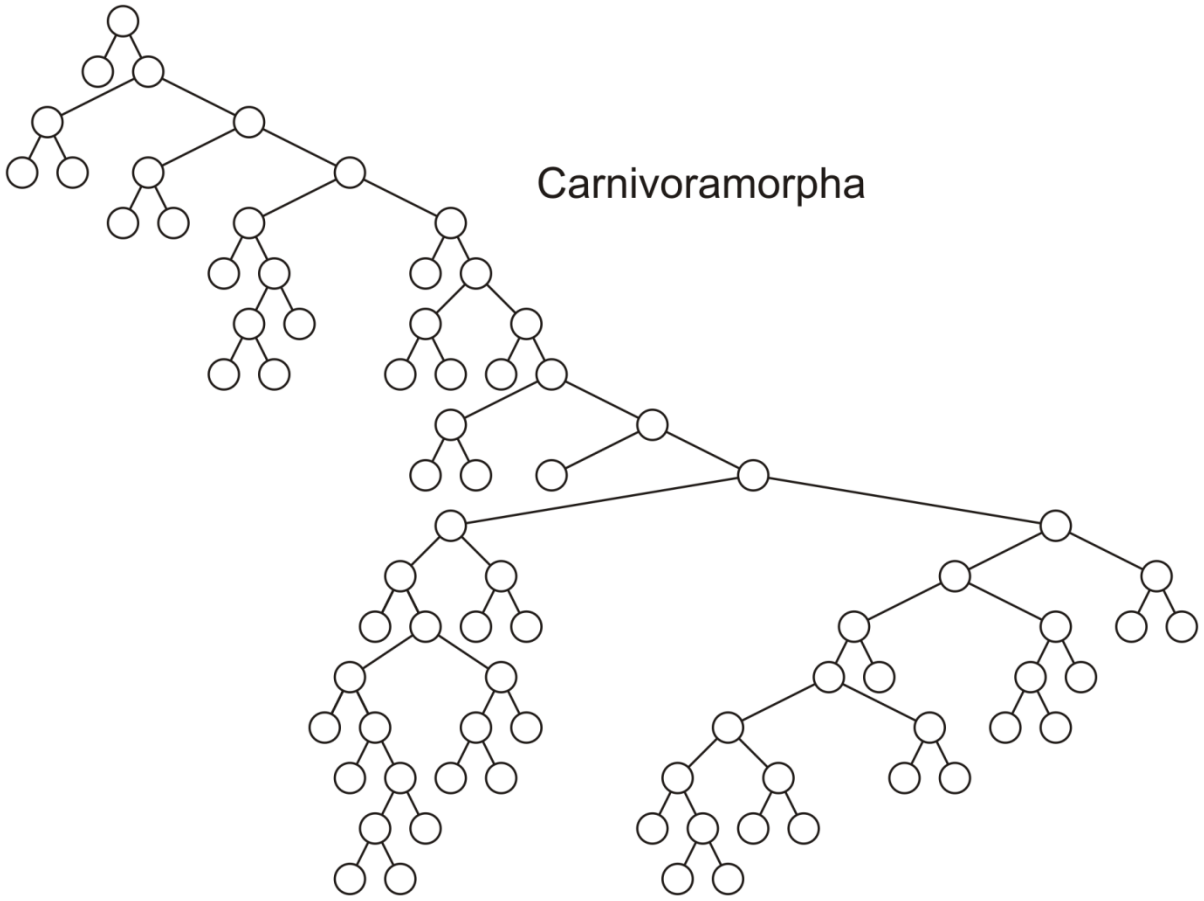
- The degree of a node is defined as the number of its children: deg(I) = 3

- Nodes with the same parent are siblings
  - J, K, and L are siblings



Terminology

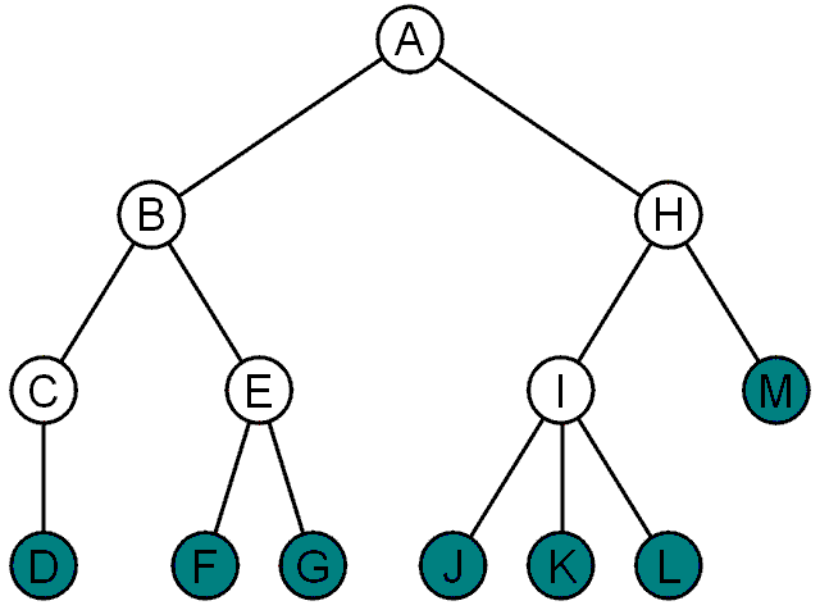
Phylogenetic trees have nodes with degree 2 or 0:



Wesley-Hunt, G. D.; Flynn, J. J. “Phylogeny of the Carnivora: basal relationships among the Carnivoramorphans, and assessment of the position of ‘Miacoidea’

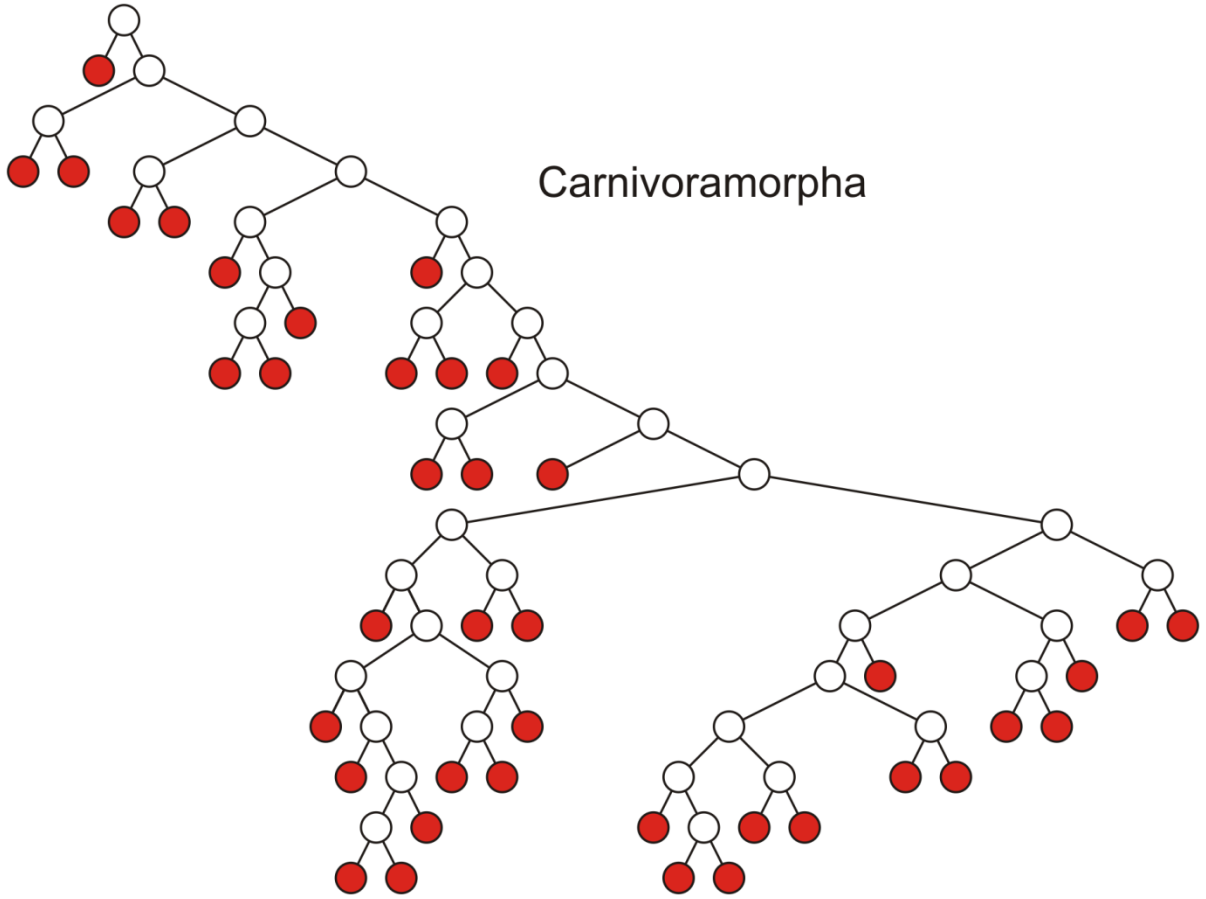
Terminology

- Nodes with degree zero are also called leaf nodes
-
- All other nodes are said to be internal nodes, that is, they are internal to the tree



Terminology

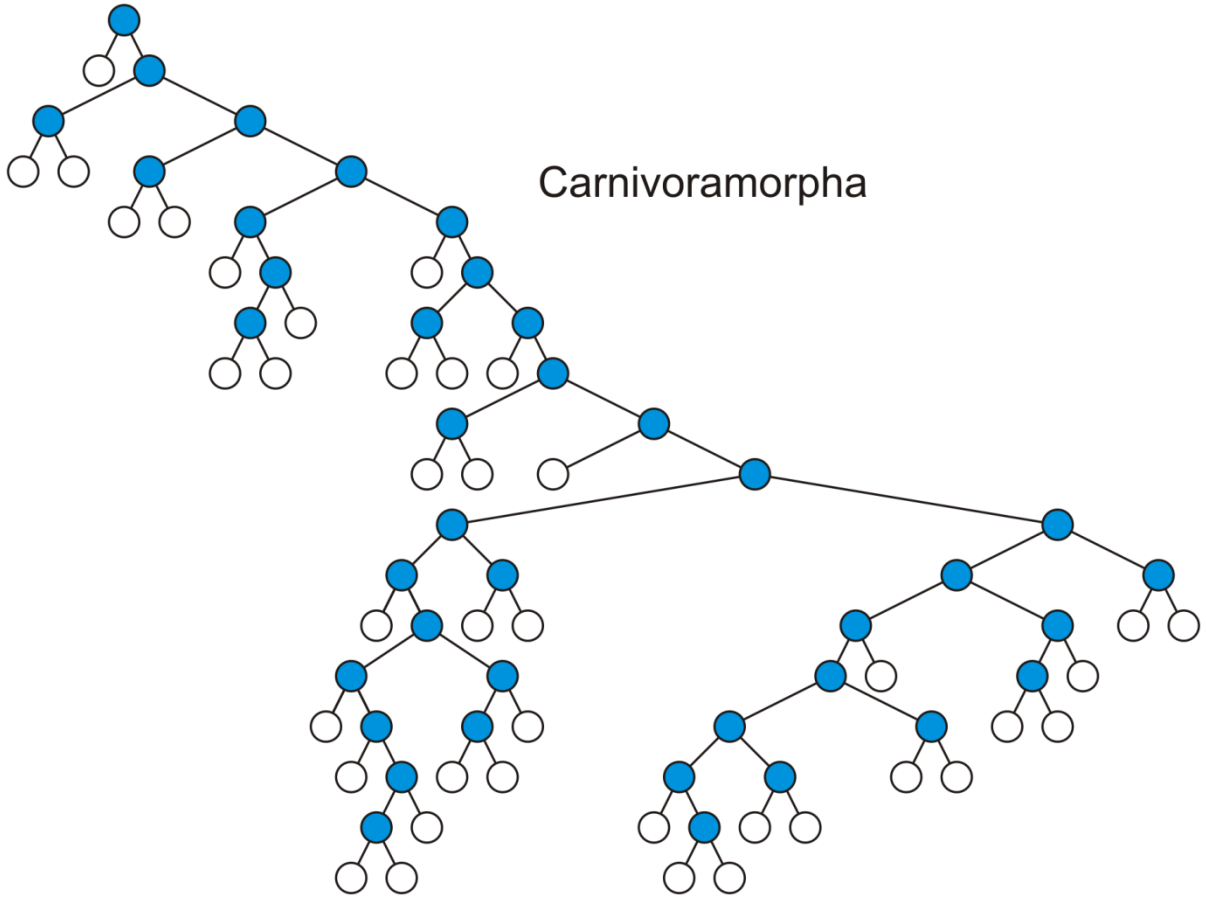
Leaf nodes:



Wesley-Hunt, G. D.; Flynn, J. J. “Phylogeny of the Carnivora: basal relationships among the Carnivoramorphans, and assessment of the position of ‘Miacoidea’

Terminology

Internal nodes:

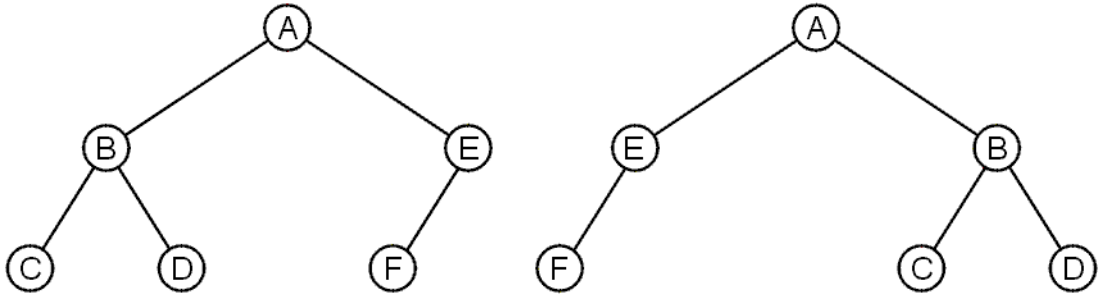


Wesley-Hunt, G. D.; Flynn, J. J. “Phylogeny of the Carnivora: basal relationships among the Carnivoramorphans, and assessment of the position of ‘Miacoidea’

Terminology

These trees are equal if the order of the children is ignored

- *unordered trees*

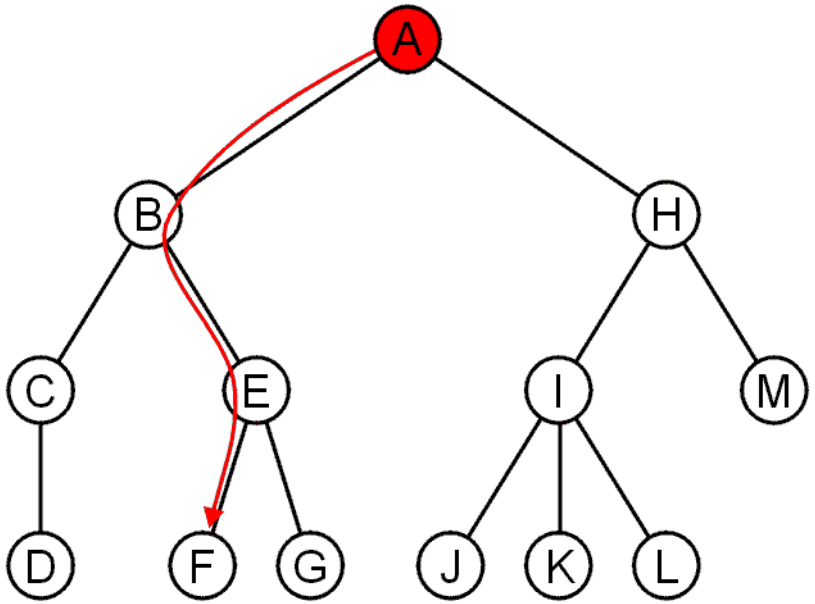


They are different if order is relevant (*ordered trees*)

- We will usually examine ordered trees (linear orders)
- In a hierarchical ordering, order is not relevant

Terminology

The shape of a rooted tree gives a natural flow from the *root node*, or just *root*



Terminology

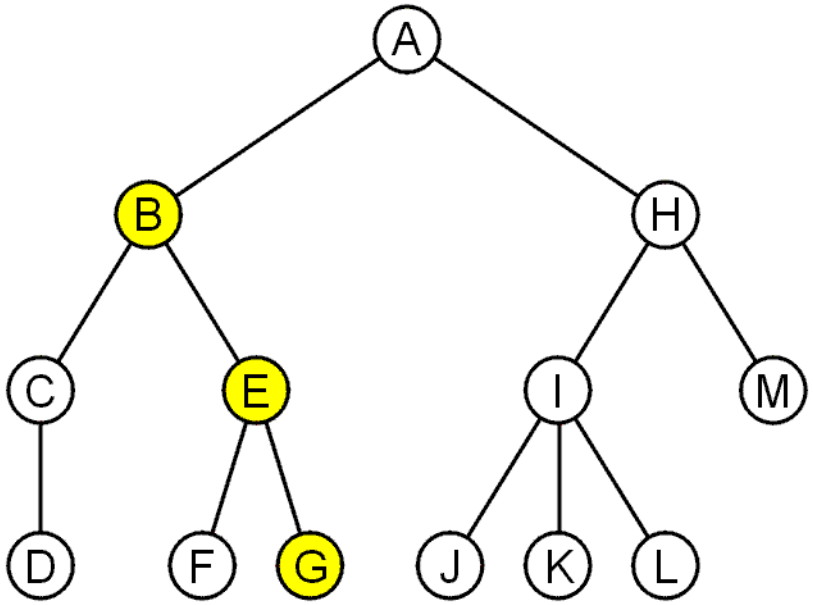
A path is a sequence of nodes

$$(a_0, a_1, ..., a_n)$$

where a_{k+1} is a child of a_k is

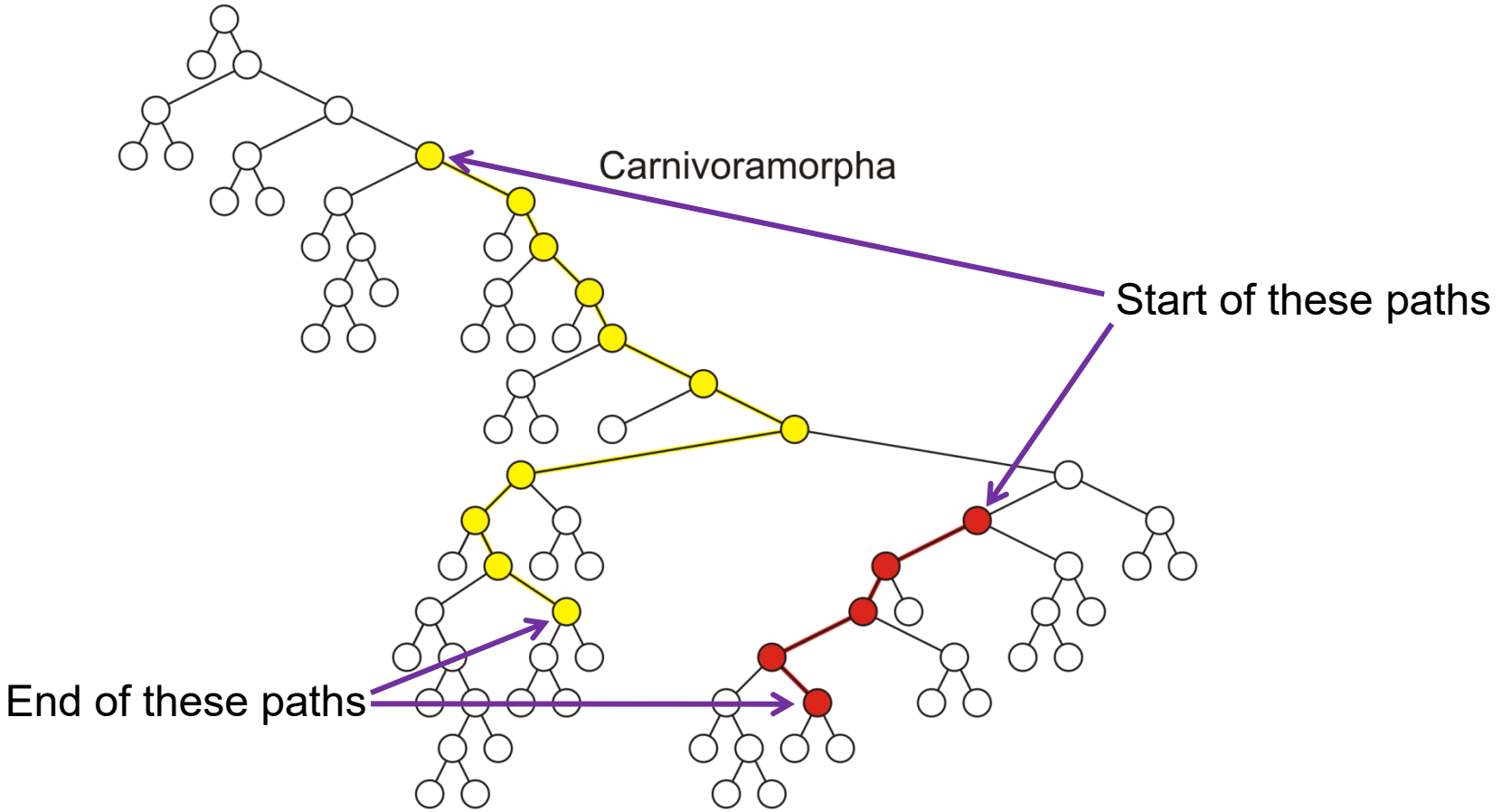
The length of this path is n

E.g., the path (B, E, G)
has length 2



Terminology

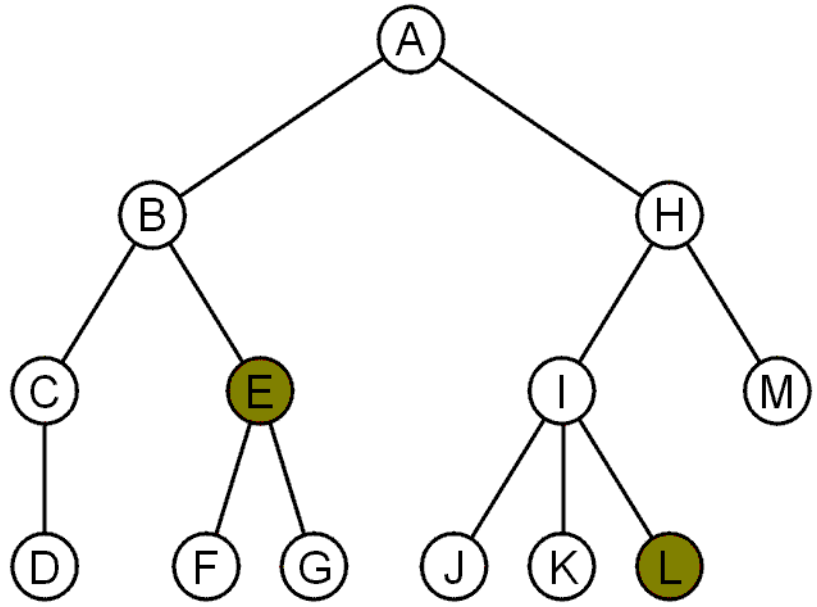
Paths of length 10 (11 nodes) and 4 (5 nodes)



Wesley-Hunt, G. D.; Flynn, J. J. “Phylogeny of the Carnivora: basal relationships among the Carnivoramorphans, and assessment of the position of ‘Miacoidea’

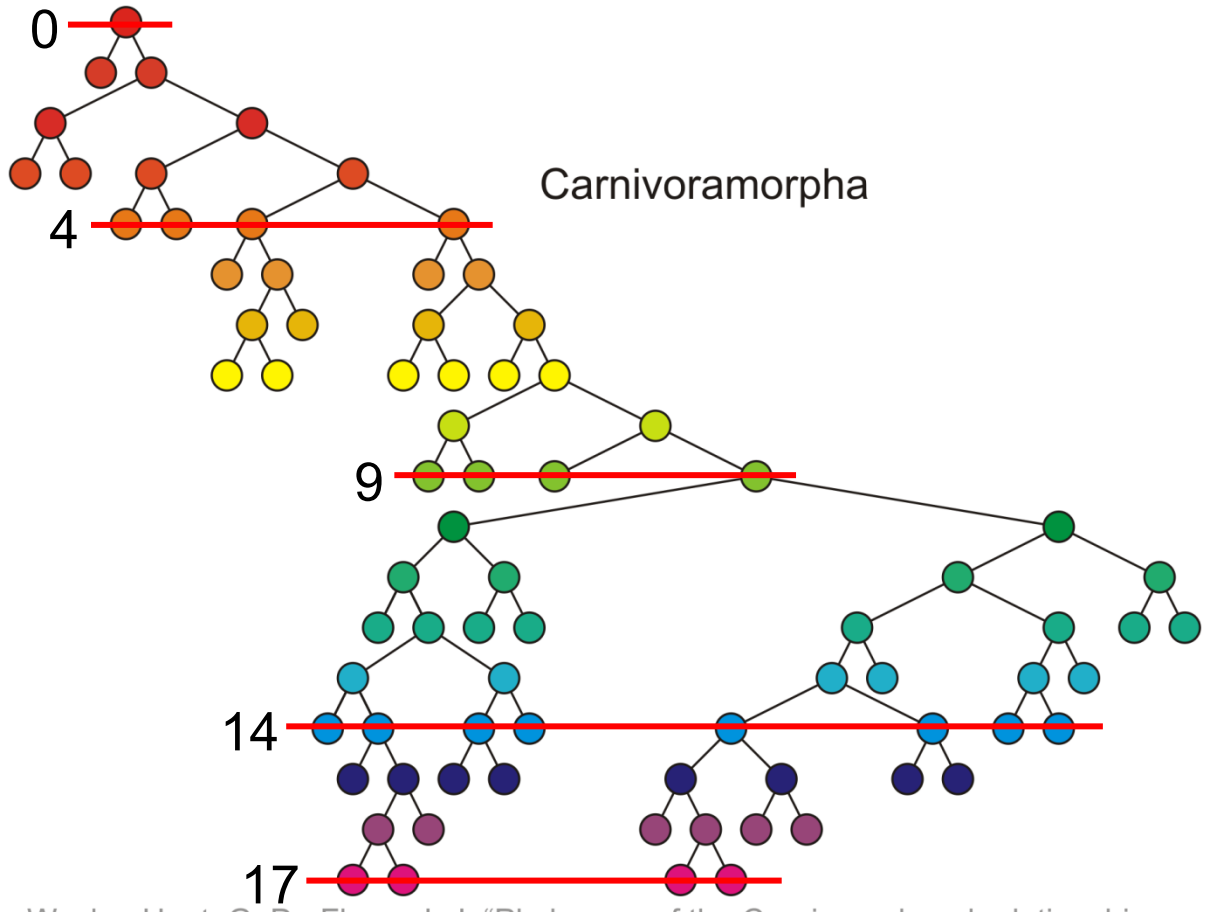
Terminology

- For each node in a tree, there exists a unique path from the root node to that node
-
- The length of this path is the depth of the node, e.g.,
  - E has depth 2
  - L has depth 3



Terminology

Nodes of depth up to 17



Wesley-Hunt, G. D.; Flynn, J. J. “Phylogeny of the Carnivora: basal relationships among the Carnivoramorphans, and assessment of the position of ‘Miacoidea’

Terminology

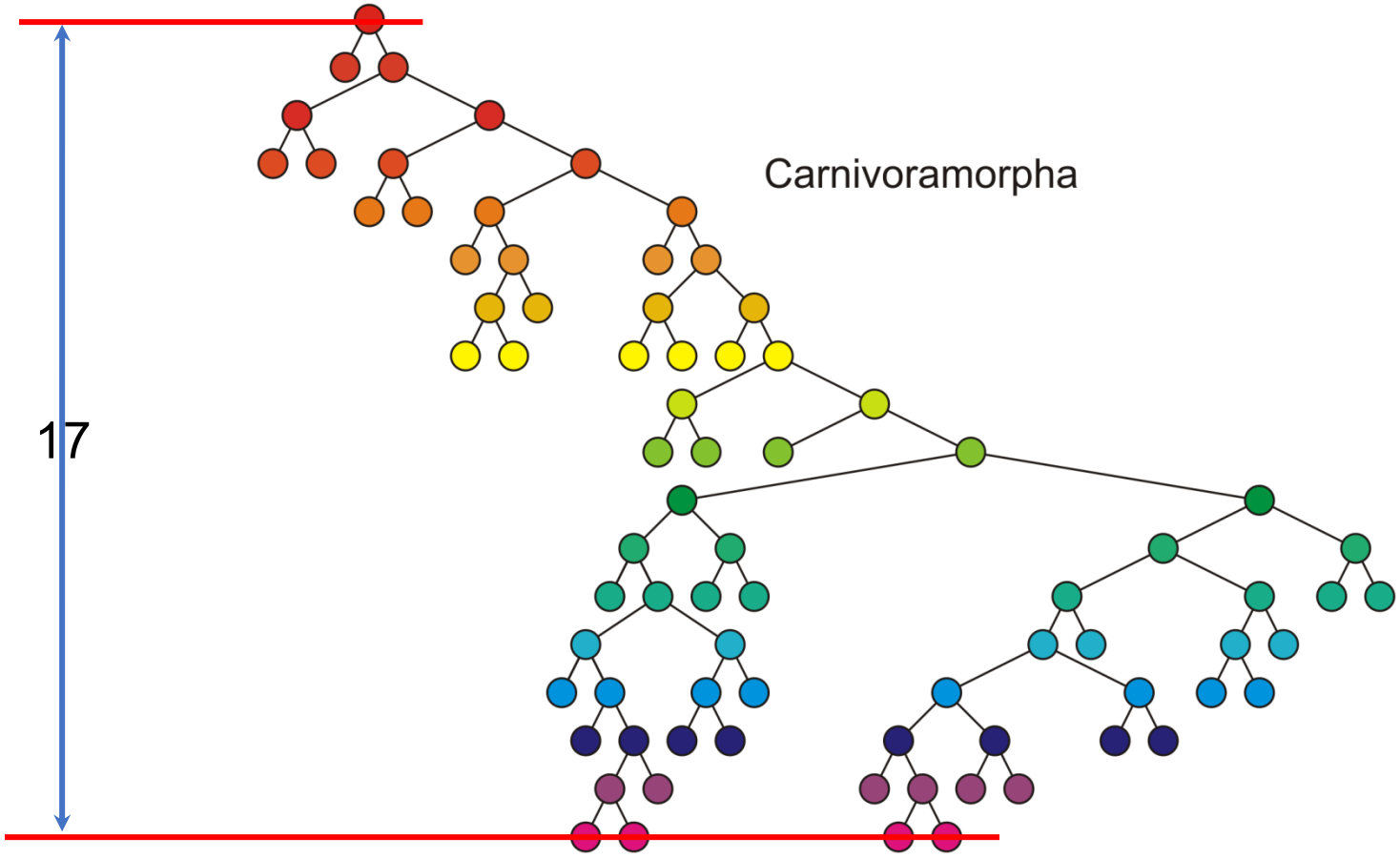
- The height of a tree is defined as the maximum depth of any node within the tree

- The height of a tree with one node is 0
  - Just the root node

- For convenience, we define the height of the empty tree to be –1

Terminology

- The height of this tree is 17



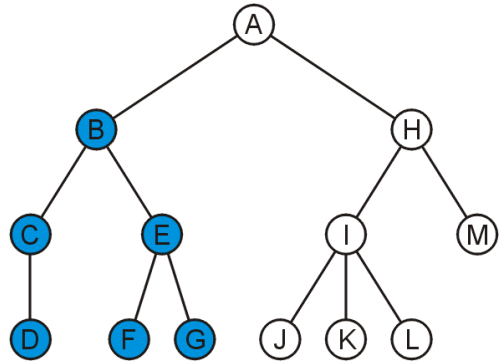
Wesley-Hunt, G. D.; Flynn, J. J. “Phylogeny of the Carnivora: basal relationships among the Carnivoramorphans, and assessment of the position of ‘Miacoidea’

Terminology

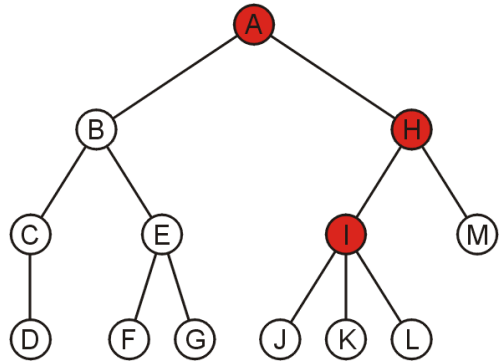
- If a path exists from node a to node b:
  - a is an ancestor of b
  - b is a descendent of a
-
- Thus, a node is both an ancestor and a descendant of itself
  - We can add the adjective strict to exclude equality: a is a strict descendent of b if a is a descendant of b but $a \neq b$
-
- The root node is an ancestor of all nodes

Terminology

- The descendants of node B are B, C, D, E, F, and G:

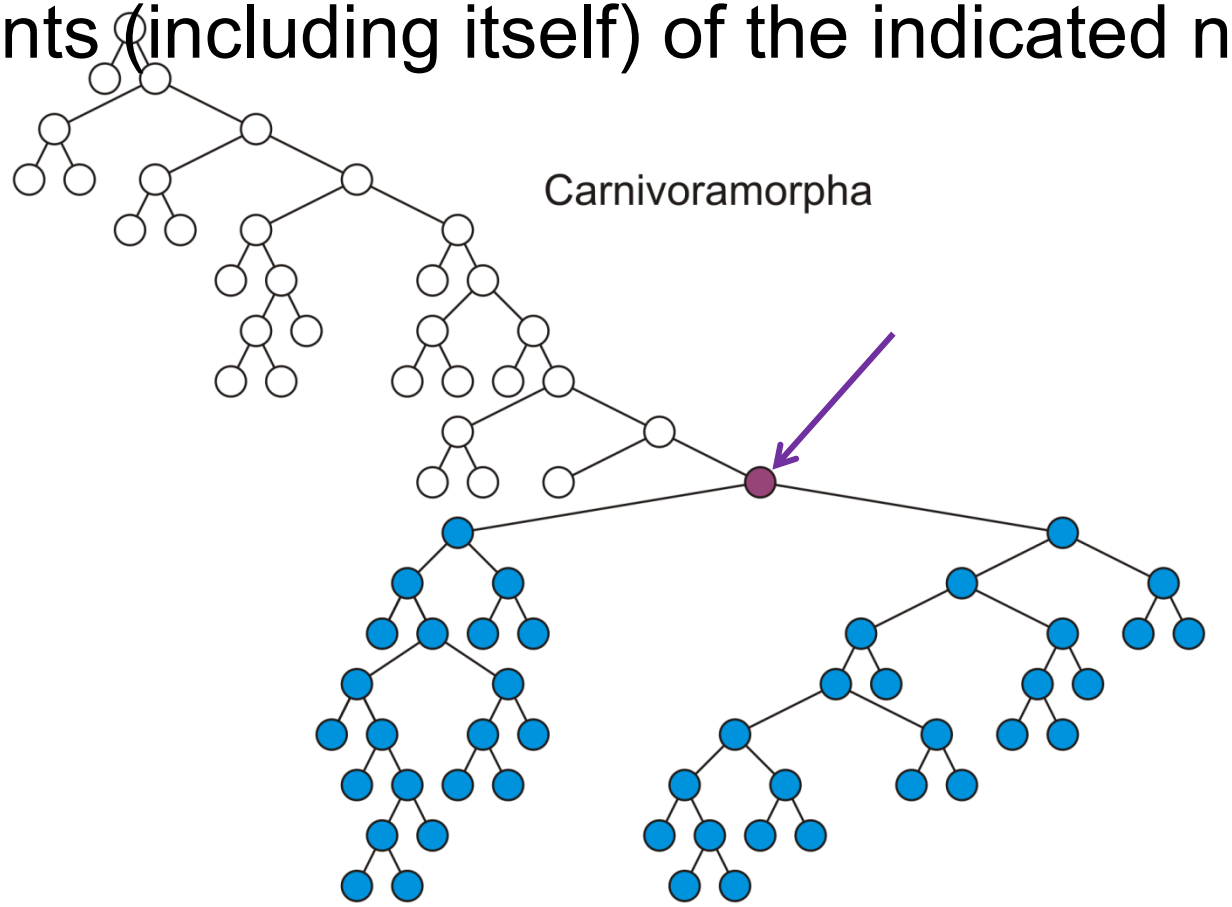


- The ancestors of node I are I, H, and A:



Terminology

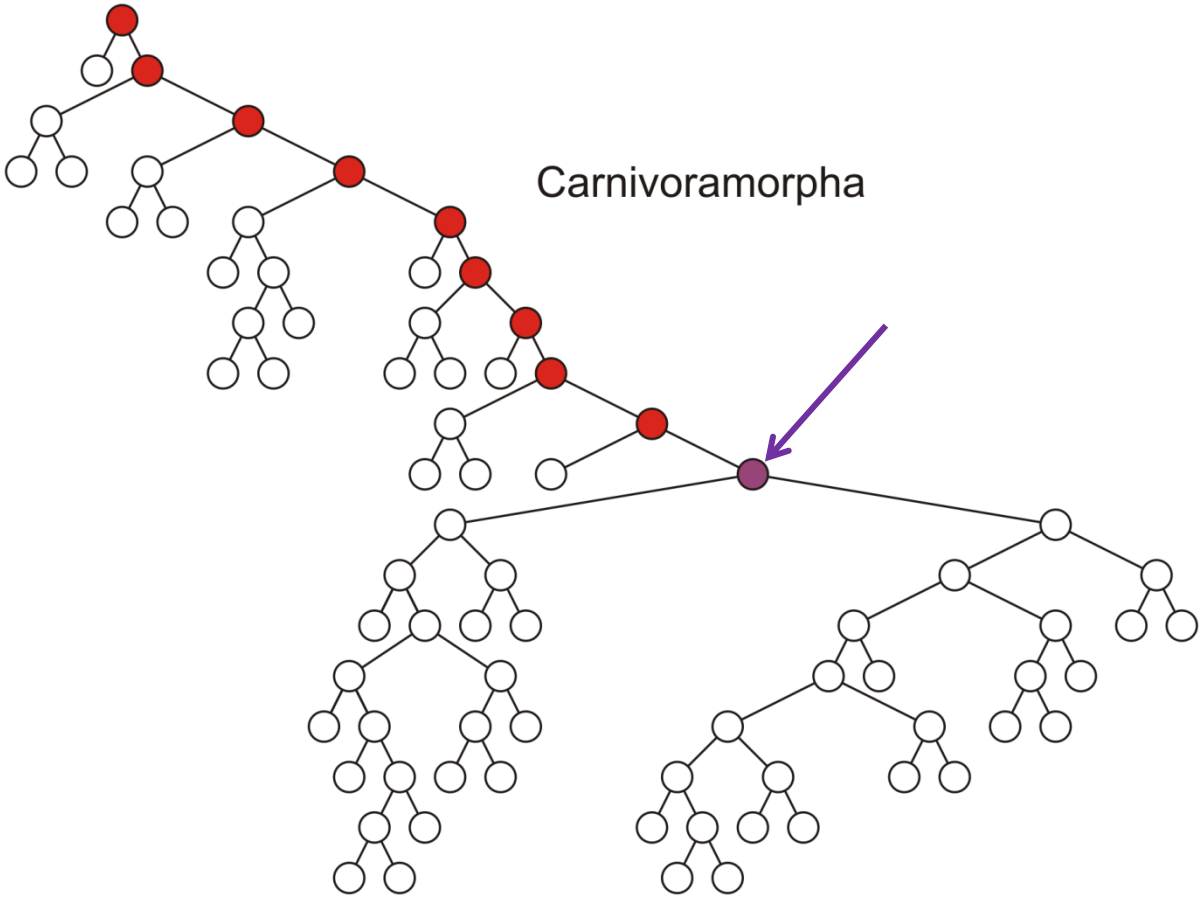
All descendants (including itself) of the indicated node



Wesley-Hunt, G. D.; Flynn, J. J. “Phylogeny of the Carnivora: basal relationships among the Carnivoramorphans, and assessment of the position of ‘Miacoidea’

Terminology

- All ancestors (including itself) of the indicated node



Wesley-Hunt, G. D.; Flynn, J. J. “Phylogeny of the Carnivora: basal relationships among the Carnivoramorphans, and assessment of the position of ‘Miacoidea’

Terminology

- Another approach to a tree is to define the tree recursively:
  - A degree-0 node is a tree
  - A node with degree n is a tree if it has n children and all of its children are disjoint trees (i.e., with no intersecting nodes)

- Given any node a within a tree with root r, the collection of a and all of its descendants is said to be a subtree of the tree with root a