3 The Tree Abstraction

Trees, and especially binary trees, are perhaps the most important of non-linear data structures in computer science. Therefore, considerable attention (in the form of five chapters) is devoted to this topic. Terms applicable to trees and their attributes are defined in §3.1. A sampling of applications and uses for tree data structures is provided in §3.2. The advanced reader may wish to briefly review the first two sections before moving to the operation specifications in §3.3 through §3.5. Readers to whom the tree concept is new may wish to consult any one of the following references for more expansive coverage of the applicable terminology: [2, 7, 16, 24, 26, 28, 30].

3.1 Concepts and Definitions

3.1.1 Basic Tree Definitions

Elements forming a tree are termed nodes, each containing some piece of information and belonging to a given data type, T. Imposed on this collection of nodes is a hierarchical structure enforcing a particular relation between the nodes. This relationship is termed parenthood [2]. One node, the root, serves as the entry node to the tree and its component nodes. As with the list abstraction, (see Volume 2 Chapter 3), the formal definition of a tree is recursive. The definition below is derived from Knuth [16] where a tree is formed of a finite set of one or more nodes such that

1. there is one specially designated node, the root, giving entry to the structure;
2. \( n \) subtrees of the root are formed by partitioning the remaining nodes (without the root) into disjoint sets;

where \( n \geq 0 \), and each subtree forms a disjoint set of nodes, each itself a tree. Obviously, every node is the root node of a subtree. A leaf (or terminal) node contains zero subtrees. Whereas, an interior node has one or more subtrees. Figures 3.1 and 3.2, on the following page, provide a graphic picture of binary and multiway tree data structures, respectively. In the diagram, circles represent nodes while lines connecting the nodes (branches) show the hierarchical relationship between nodes. Leaf nodes are shown with bold circles while the root is shaded. The other terms shown in the diagram and not explained above are defined in the sections that follow.
3.1.2 Attributes of Nodes

*Degree* is the term for the number of subtrees rooted at a given node. The degree of a tree is the largest degree over all its nodes. The binary tree in Figure 3.1 above has a degree of two while the tree in figure 3.2 has a degree of three. The *leaves* of a tree