Тема 2. СТРУКТУРЫ ДАННЫХ

2.5. Деревья

2.5.1. Основные определения

Конечное *корневое дерево T* формально определяется как связный ориентированный ациклический граф, удовлетворяющий следующим условиям:

1) имеется точно одна вершина, называемая *корнем*, в которую не входит ни одно ребро;

2) в каждую вершину (за исключением корня) входит ровно одно ребро;

3) из корня к каждой вершине существует путь (последовательность ребер), причем единственный.

Ориентированный граф, состоящий из нескольких деревьев, называется *лесом*.

Обычно считается, что в дереве все ребра имеют направление от корня, поэтому при графическом изображении ориентацию ребер можно не указывать. В описании соотношений между вершинами дерева используется терминология, принятая в генеалогических деревьях. Пусть v – произвольная вершина дерева с корнем r. Все вершины, входящие в единственный путь из r в v, называются *предками* вершины v. Если вершина w является предком вершины v, то вершина v называется *потомком* вершины w. Для каждой вершины w все ее потомки образуют *поддерево*, корнем которого является вершина w. Если (v, w) – ребро дерева, то вершину v называют *отцом* вершины w, а вершину w – *сыном* вершины v. Число сыновей вершины называется ее *степенью*. Дерево, степень каждой вершины которого не превышает некоторого целого m, называется m-*арным*. Вершины, не имеющие сыновей, называются *листьями*. Вершины, имеющие сыновей (которых будем называть *братьями*), называются *внутренними*. Таким образом, в корневом дереве все вершины являются потомками его корня, и наоборот, корень есть предок всех своих потомков.

Длина пути от корня до произвольной вершины v называется *уровнем* (*глубиной*) вершины v и определяется числом ребер в этом пути. Следовательно, уровень корня дерева равен нулю, а уровень любой другой вершины имеет уровень на единицу больше уровня своего отца. *Высота* вершины есть длина самого длинного пути от этой вершины до какого-нибудь листа. *Высотой дерева* называется высота его корня. Другими словами, высота дерева равна максимальному из уровней его листьев.

Введенные понятия проиллюстрированы на рис. 2.12. Лес состоит из двух деревьев с корнями a и n. Листьями являются вершины c, d, g, h, i, l, m, p, q, r. Вершина a является отцом вершин b, e и j, которые являются сыновьями вершины a и братьями по отношению друг к другу. Высоты деревьев равны соответственно 3 и 2. Уровень вершины j равен 1, а высота равна 2.

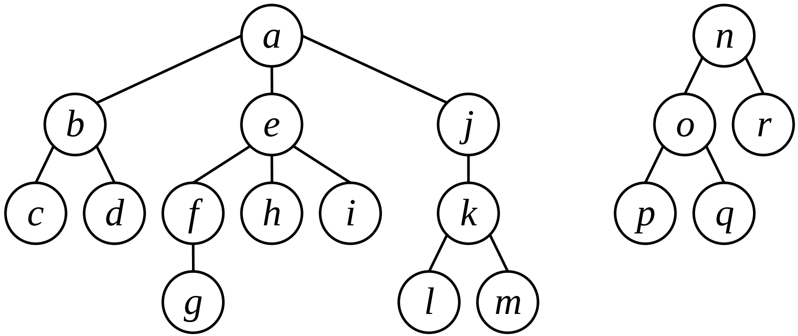
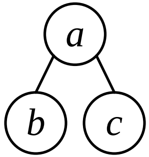
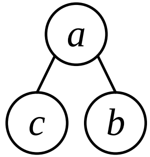


Рис. 2.12. Лес, состоящий из двух деревьев

Все рассматриваемые деревья будем считать упорядоченными, т. е. для них важен относительный порядок поддеревьев каждой вершины. Обычно предполагается, что в упорядоченном дереве множество сыновей каждой вершины упорядочено слева направо. Поэтому деревья

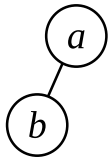


и

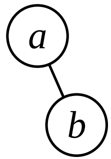


считаются различными.

Важной разновидностью корневых деревьев является класс *бинарных* (*двоичных*) *деревьев*. Бинарное дерево T либо пустое, либо состоит из корня и двух бинарных поддеревьев: левого T_l и правого T_r. Все вершины в поддереве T_l расположены левее всех вершин в поддереве T_r. Таким образом, в бинарных деревьях каждая вершина имеет не более двух сыновей, причем каждый сын интерпретируется либо как *левый* сын, либо как *правый* сын. Поэтому два следующих дерева



и



представляют собой два различных бинарных дерева. Это является существенным отличием от других деревьев. В бинарном дереве важно, каким является единственный сын вершины (левым или правым), для других деревьев такого различия нет.

2.5.2. Представления деревьев

Большинство представлений множеств с помощью деревьев как структур данных основано на связных распределениях. Каждая вершина представляет собой узел, состоящий из поля *info* и нескольких полей для хранения указателей. Такое представление называется *узловым*.

Проще всего задача представления решается для бинарных деревьев. Для этого используются узлы фиксированного размера, состоящие из полей *left* (для хранения указателя на левого сына), *info* и *right* (для хранения указателя на правого сына). Если у узла, на который ссылается указатель *p*, нет левого или правого сына, то *p*.*left* = Λ или *p*.*right* = Λ. С бинарным деревом связан специальный внешний указатель *root* (*указатель дерева*), который указывает на его корень. Если *root* = Λ, то дерево пустое. Пример узлового представления бинарного дерева с корнем *n* из рис. 2.12 показан на рис. 2.13. При необходимости в ряде прикладных алгоритмов можно добавить указатель отца *father* для облегчения движения от потомков к предкам.

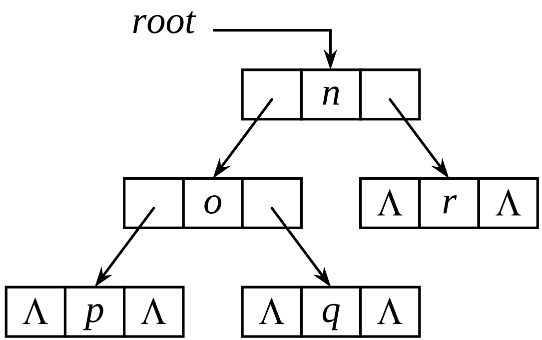


Рис. 2.13. Узловое представление бинарного дерева

Что касается деревьев с произвольным ветвлением, то задача их представления решается сложнее. Один из способов представления заключается в следующем: все узлы имеют фиксированный размер и состоят из поля *info* и поля связи *father*, указывающего на отца узла. При этом для доступа к узлам дерева необходим набор внешних указателей для каждого узла или, по крайней мере, для каждого листа. Пример такого представления дерева с корнем *a* из рис. 2.12 показан на рис. 2.14 (внешние указатели не показаны). Рассмотренный способ представления полезен, если необходимо подниматься по дереву от потомков к предкам. Такая операция встречается редко; чаще требуется движение от предков к потомкам.

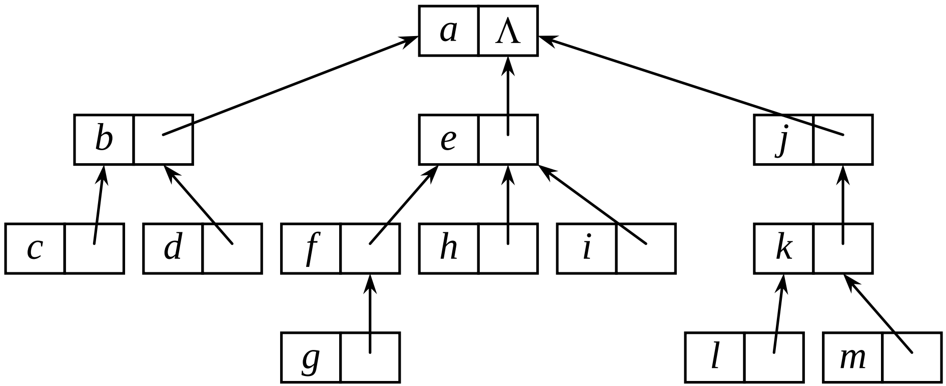


Рис. 2.14. Представление дерева с помощью узлов с полем *father*

Представление дерева с использованием указателей, ведущих от предков к потомкам, создает определенные трудности. Если известно, что число сыновей каждой вершины ограничено сверху константой k, то дерево можно реализовать аналогично бинарному дереву, помещая в узел k указателей на сыновей. Проблема заключается в том, что если у большинства вершин число сыновей существенно меньше k, то бесполезно тратится большой объем памяти. Другая проблема заключается в том, что такая реализация невозможна, если число сыновей может быть любым и неизвестна верхняя граница k (заранее неизвестно, сколько полей для указателей необходимо выделять).

Решением указанных проблем является метод преобразования произвольного дерева (леса) в бинарное. Такое преобразование называется *естественным соответствием* между лесами и бинарными деревьями и заключается в следующем. Поле *left* бинарного дерева предназначается для указания самого левого сына данной вершины, а поле *right* – для указания следующего брата данной вершины. Другими словами, поле *left* каждого узла указывает на связный список сыновей этого узла; список связывается с помощью полей *right*. Пример такого представления леса (см. рис. 2.12) показан на рис. 2.15 (для упрощения не показаны поля узлов дерева).

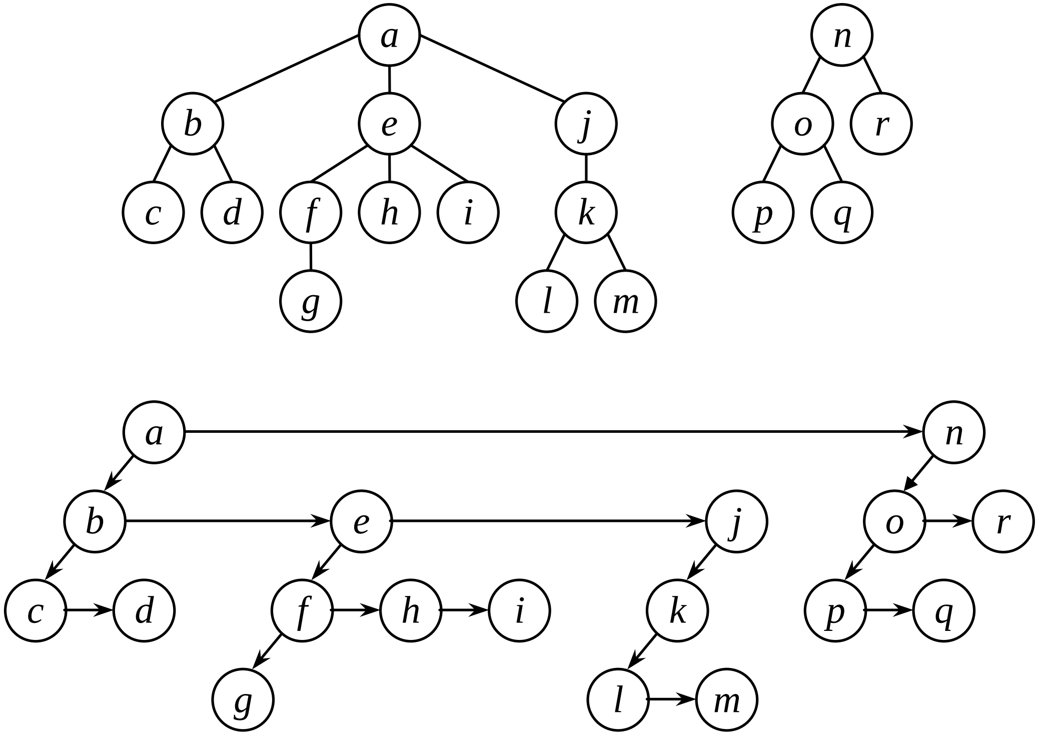


Рис. 2.15. Бинарное дерево, соответствующее лесу на рис. 2.12

Возможны и другие представления деревьев, конкретный выбор представления определяется спецификой задачи.