Деревья

Алексей Владыкин

СПбГУ ИТМО

15 марта 2010

Определения

- Корневое дерево — это:
  - пустое дерево;
  - отдельный узел (node);
    он же является корнем (root) этого дерева;
  - узел n с поддеревьями $T_1, \ldots, T_k$; при этом
    $n_1, \ldots, n_k$ (корни $T_1, \ldots, T_k$) — дети (children) узла n;
    n — корень дерева и родитель (parent) узлов $n_1, \ldots, n_k$.
- Внешняя вершина, лист (leaf) — вершина без детей.
- Внутренняя вершина — вершина с детьми.
- Ребро (edge) — связь между родителем и ребенком.
- Лес — множество деревьев.

Разновидности деревьев

- Упорядоченные деревья
- Позиционные деревья
- k-ичные деревья
- Помеченные деревья
  (пометки на вершинах или на рёбрах)

Примеры

- Структура книги: главы, разделы, подразделы
- Диаграмма классов
  (в языках без множественного наследования)
- Генеалогическое дерево
- Файловая система
- Дерево решений (decision tree)
- Суффиксное дерево (suffix tree)
- Префиксное дерево (prefix tree)
- Абстрактное синтаксического дерево (abstract syntax tree)

Запись деревьев

- Графическое изображение
- Структура вложенных скобок
- Список с отступами
- Круги Эйлера / диаграммы Венна

Свойства

- Свойства отношения «родитель — ребёнок»:
  - антирефлексивность;
  - антирефлексивность транзитивного замыкания (ацикличность);
  - антисимметричность;
  - если $(i, j) \in R$, то $\forall k \neq i : (k, j) \notin R$ (требование единственности родителя).
- Если n и e — количество узлов и рёбер, то $e = n - 1$.
- Количество деревьев из n помеченных узлов
  $\text{trees}(n) = n^{n-2}$ (теорема Кэли).
- Количество двоичных деревьев из n помеченных узлов
  $\text{trees}(n) = \frac{1}{n+1} C_{2n}^{n}$ (числа Каталана).
- Количество узлов в полном k-ичном дереве высоты h
  $\text{nodes}(k, h) = 1 + k + k^2 + \ldots + k^h = \frac{k^h - 1}{k - 1}$.

Обход дерева

- Прямой порядок (префиксный):
  корень — поддерево 1 — ... — поддерево k;
- Обратный порядок (постфиксный):
  поддерево 1 — ... — поддерево k — корень;
- Симметричный порядок (инфиксный):
  поддерево 1 — корень — поддерево 2 — ... — поддерево k.

Деревья поиска

В каждой вершине дерева поиска:

- поддеревья $t_1, t_2, \ldots, t_r$;
- набор ключей $k_1 < k_2 < \ldots < k_{r-1}$;
- все ключи поддерева t_1 меньше k_1;
- все ключи поддерева t_i $(1 < i < r)$ между k_{i-1} и k_i;
- все ключи поддерева t_r больше k_{r-1}.

В сбалансированном дереве поиска операции добавления, удаления и поиска ключа выполняются за $O(\log_r n)$.

Представление деревьев в программах

- Структура на указателях, в каждом узле — указатели на детей

```
struct node_t {
    //node_t* parent;
    node_t* children[];
    data_t data;
};
```

- Структура на указателях, в каждом узле указатель на первого ребёнка и следующего брата

```
struct node_t {
    //node_t* parent;
    node_t* first_child;
    node_t* next_sibling;
    data_t data;
};
```

- Для k-ичного дерева ограниченного размера — массив