

Предварительная структуризация и В-Деревья

Алгоритмы 7 - Айрат Хасьянов

Головоломка

- $2\ 2\ 2\ 2 = 9$
- $(,), *, +, -, /$
- Как?

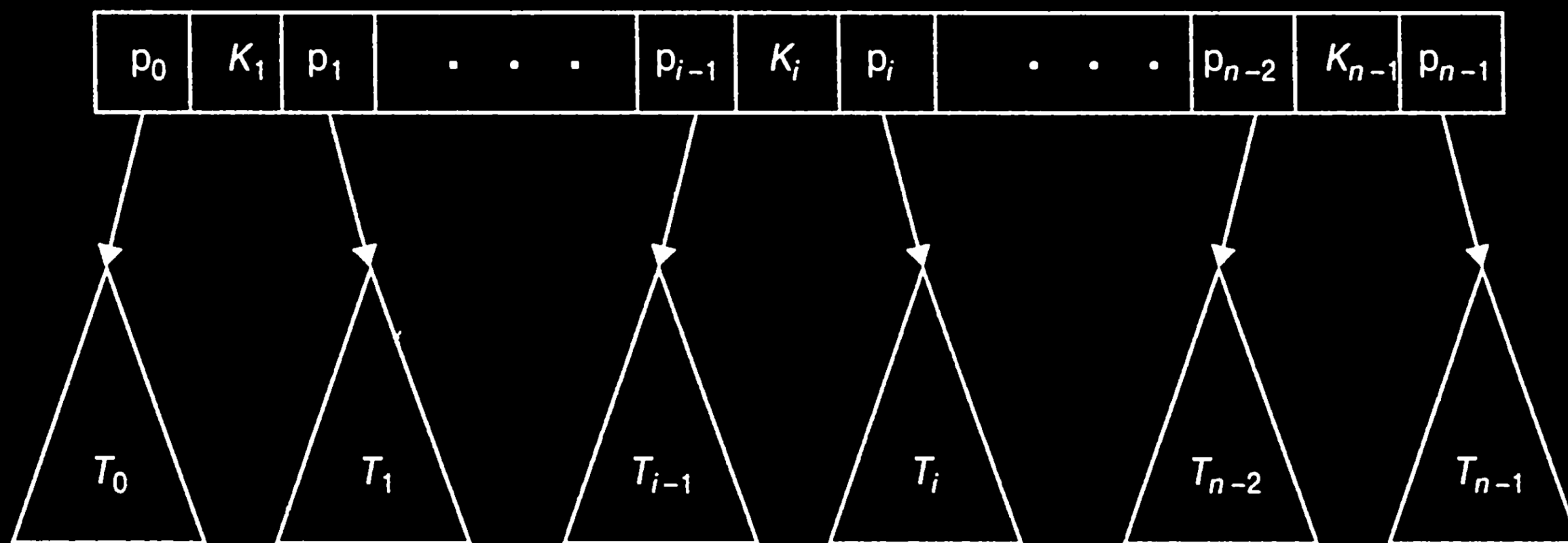
О В-Деревьях

- **Индекс** - предоставляет информацию о размещении записи по значению ключа.
- **В-Дерево** - способ организации индексов.

Определение

- **В-Деревья** предложены Р. Бэйером и Э. МакКрейтом;
- **В-Дерево** является расширением 2-3 дерева, в котором узлы могут иметь много ключей.
- В **В-Дереве** все записи хранятся в листьях в порядке возрастания значений ключей.
- Каждый родительский узел содержит $n-1$ упорядоченных ключей $K_1 < \dots < K_{n-1}$.
- Все ключи в поддереве T_0 меньше K_0 , в поддереве T_1 - меньше T_1 и т.д., в поддереве $T_{n-1} < K_{n-1}$.

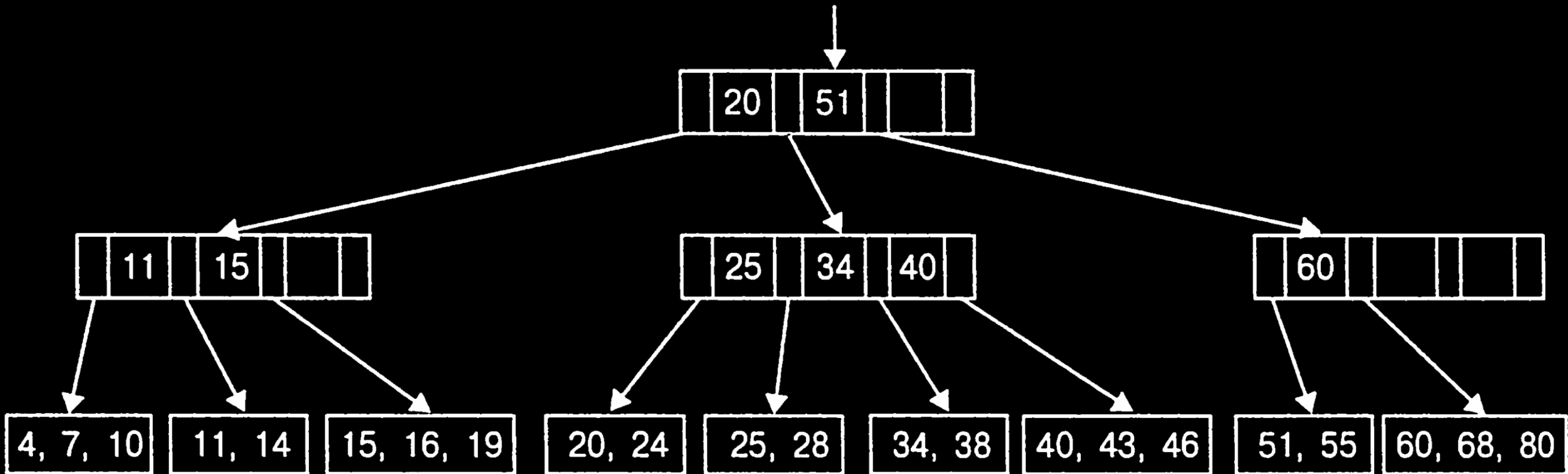
Родительский узел



У дерева порядка m :

- Корень либо является листом, либо имеет от 2 до m потомков;
- Каждый узел, кроме корня, имеет от $m/2$ до m потомков;
- Дерево идеально сбалансировано.

Пример В-Дерева



Основной мерой сложности является количество обращений к узлам. Эта величина всегда равна высоте дерева $h + 1$.

Высота В-дерева с n узлами

- Корень содержит как минимум 1 ключ.
- На уровне 1 как минимум 2 узла, минимальное количество ключей в каждом узле: $\lceil m/2 \rceil - 1$
- Количество узлов на втором уровне: $2\lceil m/2 \rceil$
- Количество ключей в каждом узле: $\lceil m/2 \rceil - 1$
- Всего ключей на 2-м уровне: $2\lceil m/2 \rceil (\lceil m/2 \rceil - 1)$
- Количество ключей на i -м уровне: $2\lceil m/2 \rceil^{i-1} (\lceil m/2 \rceil - 1)$
- Количество листьев: $2\lceil m/2 \rceil^{h-1}$

Как оценить общее количество узлов в дереве?

Высота В-дерева с n узлами (2)

$$n \geq 1 + \sum_{i=1}^{h-1} 2 \lceil m/2 \rceil^{i-1} (\lceil m/2 \rceil - 1) + 2 \lceil m/2 \rceil^{h-1}$$

$$n \geq 4 \lceil m/2 \rceil^{h-1} - 1$$

$$h \leq \left\lceil \log_{\lceil m/2 \rceil} \frac{n+1}{4} \right\rceil + 1$$

- Таким образом, сложность поиска в дереве, равная h лежит в $O(\log n)$.

Задачи

1. Напишите алгоритм добавления элемента в B-дерево (он будет рекурсивным);
2. Напишите алгоритм нахождения максимального ключа в B-дереве.

Домашнее задание

- Реализуйте хранение словаря в файле (файлах). Используйте индекс на базе B-дерева для быстрого доступа к записям в словаре. Конечно же, приложение должно иметь полное покрытие модульными тестами.