Б-деревья

сильно ветвящиеся, у которых есть весьма практическая обл. применения: формирование и поддержка крупномасштабных деревьев поиска в которых необходимы и включение новых элементов, и удаление старых, но для которых либо не хватает ОЗУ, либо она слишком дорога чтобы использовать ее для долговременного хранения. то есть что имеется ввиду раньше деревья поиска были все в ОЗУ теперь например не хватает памяти для хранения всех элементов. Порция данных, кот. хранится в ОЗУ, называется страница. в ОЗУ не хранится весь набор данных. в ОЗУ +еще указатели во внешнюю память (адреса памяти на жестком диске) . в 1970 году (это очень важно!) эти два товарища – коммуниста или сотрудники сформулировали след. критерий: каждая страница кроме одной должна содержать при заданном N от N до 2N вершин. здесь речь идет о Б-деревьях порядка N. Они обладают след свойствами:

1. Каждая страница содержит не более 2N элементов (ключей)
2. Каждая страница кроме корневой содержит не менее N элементов
3. Каждая страница либо представляет собой лист либо имеет M+1 го потомка (M-кол-во ключей на этой странице)
4. Все страницы-листья находятся на одном уровне

Алгоритм добавления, удаления,

Пусть N=2 тогда любая страница – не более 2N ключей

меньше всех

m=3

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Key |  |  |  |  |
| Count |  |  |  |  |
| P0 |  |  |  |  |

m=4

Добавление данных в Б-дерево.

1 5 13 222

Ключ всегда помещается в лист.

если происходит переполнение, то эл-т вытесняется наверх и лист расщепляется.

Удаление элементов.

Мы должны удалять эл-ты с условием неизменения свойств и чтобы не нарушалась упорядоченность дерева.

Если лист, то просто удаляется. если 1 потомок, то он становится на место удал-го эл-та. иначе ….

У узла не может быть только 1 ссылка. Надо делать слияние страниц .

Лексикографический смежный эл-т либо !самый правый из левого поддерева! или самый левый из правого поддерева.

Если есть правая страница, то берем с правой, иначе с левой. давайте ради прикола удалим листовой элемент.

При слиянии страниц верхний элемент опускается вниз.

Еще бывают Б+ деревья. Истинные значения ключей Б+ дерева содержатся только в листьях. а в нетерминальных узлах находятся ключи-разделители, создающие диапазон изменения ключей для поддеревьев. Еще одна разновидность. (2-3) дерево. это когда N=1. В Б+деревьях слева <= а справа > Средний элемент вверх, и остается в левом. Недостаток в том что дублирование информации. При удалении – если могут два листа объединится, объединяются. центральный – поднимается. Все реально размещенные элементы хранятся в листьях. Отсюда преимущество Б+деревьев – появляется возможность линейного поиска. Каждый лист знает своего соседа справа.