

Структура дерева кодирования

Структура дерева кодирования в алгоритме Хаффмана представляет собой бинарное дерево, где каждый лист дерева соответствует символу, а каждый внутренний узел представляет собой суммарную частоту символов в своих поддеревьях. Вот ключевые аспекты структуры дерева кодирования Хаффмана:

1. Узлы

- Листья:
 - Листья дерева соответствуют уникальным символам из исходного набора данных.
 - Каждый лист содержит символ и его частоту в исходных данных.
 - Не имеют дочерних узлов (left и right равны null).
- Внутренние узлы:
 - Внутренние узлы представляют суммарную частоту символов своих поддеревьев.
 - Не содержат символов (поле ch устанавливается в '\0').
 - Имеют ссылки на левые и правые поддеревья (left и right).

2. Частоты

- Каждый узел имеет ассоциированную частоту.
- Частота листа соответствует частоте символа.
- Частота внутреннего узла — сумме частот его дочерних узлов.

3. Построение дерева

- Дерево строится путем соединения узлов с наименьшей частотой.
- Процесс повторяется, пока не останется единственный узел — корень дерева.

4. Кодирование

- Коды Хаффмана генерируются на основе пути от корня к листьям.
- При переходе влево добавляется "0", при переходе вправо — "1".

5. Уникальность кодов

- Коды Хаффмана являются префиксными кодами, что означает, что ни один код символа не является префиксом другого.
- Это свойство обеспечивает однозначную декодируемость закодированных данных.

6. Эффективность кодирования

- Часто встречающиеся символы получают более короткие коды.
- Это повышает эффективность сжатия данных.