

# Описание структуры дерева кодирования

29 декабря 2023 г.

## 1 Введение

Дерево кодирования представляет собой бинарное дерево Хаффмана, подчиняющееся префиксному правилу (никакое кодовое слово не должно быть началом другого). Это позволяет декодировать однозначно и за один проход по коду. Каждый лист дерева - некоторый уникальный символ (представленный в виде байта), соответствующий некоторой бинарной последовательности, составленной из пути от корня к листу. Все левые ветви помечаются нулем, правые - единицами. Чем чаще попадается символ, тем короче его бинарное представление.

## 2 Структура класса Vertex

Класс Vertex хранится в файле <src/Vertex.java>, объект класса представляет собой узлы дерева, не являющиеся листьями. Класс реализует интерфейс Comparable, имеет ссылки на детей (при этом у изначально созданных листьев детей нет).

## 3 Построение дерева

Дерево строится в методе buildTree() класса Tree, используя очередь с приоритетом, на этапе инициализации заполнившейся листьями.

## 4 Структура закодированного файла

Строка 1: "P O L Y N U F F M A N". Это аналог ватермарки, наличие данной строки говорит о том, что ее можно расшифровывать.

Строка 2: число n - количество символов, присутствующих в исходном файле.

Строка 3: n строк хэш-таблицы кодов, идут сплошным текстом (символ и код без разделителей).

Строка 4: Закодированный файл.