**Задание помеченных деревьев кодом Прюфера**

Английским математиком А.Кэли (1821-1895) в середине 19 века было доказано, что число помеченных деревьев с n ≥ 2 вершинами равно nn-2. Немецкий математик Х.Прюфер (1896-1934) предложил способ однозначного кодирования *n* - вершинного помеченного дерева упорядоченной последовательностью из *n*-2 номеров его вершин. Так как число таких последовательностей равно *nn-2*, то каждому коду Прюфера можно сопоставить некоторое дерево, и наоборот, то есть между ними существует [взаимно однозначное соответствие](https://ru.wikipedia.org/wiki/%D0%91%D0%B8%D0%B5%D0%BA%D1%86%D0%B8%D1%8F).

**Построение кода Прюфера по дереву**

Достаточно *n-2* раза выполнить следующую процедуру:

1. Выберем в текущем (исходном) дереве висячую вершину с наименьшим номером.
2. Запишем в слово (слева-направо) номер смежной с ней вершины.
3. Удалим висячую вершину вместе с инцидентным ребром.

**Пример.** Помеченное 7- вершинное дерево и его код Прюфера

1

2

3

6

5

7

4

→ 57517

**Построение списка ребер по коду Прюфера**

Запишем вспомогательный массив, содержащий числа от 1 до *n* и выполним  *n-2* раза следующую процедуру:

1. Удалим из вспомогательного массива минимальное число, которого нет в слове (исходном, а затем в текущем).
2. Удалим из текущего (исходного) слова первый элемент.
3. Составим ребро, концевыми вершинами которого являются вершины с номерами, соответствующими удаленным числам.

Последнее (*n-1)*-ребро составляется из двух вершин с номерами, оставшимися во вспомогательном массиве после выполнения *n-2* раза предыдущей процедуры.

Из всех ребер, отождествляя вершины с одинаковыми номерами, конструируется итоговый граф. Покажем, что этот граф не содержит циклов.

Введем обозначения: *b1b2…bn-2*  **-** код Прюфера;

*a1* **–** минимальное число, не входящее в *b1b2…bn-2* , генерируемое ребро *(a1,b1)*

*a2* **–** минимальное число, не входящее в *b2 b3…bn-2* , генерируемое ребро *(a2,b2)*

**. . .**

*an-2* **–** минимальное число, не совпадающее с *bn-2* , генерируемое ребро *(an-2,bn-2)*

*(an-1,an)* **–** последнее (*n-1)*- ребро.

Ацикличность конструируемого графа покажем индукцией по числу ребер, перечисляя ребра в порядке обратном их генерации:

1. Граф с одним ребром *(an-1,an),* очевидно, не содержит циклов.
2. Предположим, что добавление к графу ребер *(an-2,bn-2),*…, *(an-i,bn-i)* не образует циклов.
3. Покажем, что добавление к графу ребра *(an-i-1,bn-i-1)* не сможет образовать цикл в текущем графе. Действительно, учитывая способ генерации ребер, номер *an-i-1* не совпадает ни с одним из номеров *bn-i-1* , *bn-i*,*…*, *bn-2*. Кроме того, числа *ai≠ aj* при *i≠j*. Таким образом, хотя бы одна из концевых вершин добавляемого ребра *(an-i-1,bn-i-1)* не принадлежит уже построенному графу и цикл при этом образоваться не может.

Связность графа покажем методом от противного. Пусть граф содержит *k* компонент связности. Каждая из них является деревом, поскольку граф не содержит циклов. При этом число вершин *ni* и ребер *mi* каждой из *k* компонент удовлетворяют соотношениям *mi= ni-1*. Просуммировав их, имеем следующие равенство , где  и . Таким образом, получаем, что . Но в построенном графе . Следовательно,  и граф содержит лишь одну компоненту связности.

**Пример.** Код Прюфера и восстановленное по нему дерево.

73442 6

*n=7* (1,7)(3,5)(3,4)(4,6)(2,4)(2,7) 5 3 2 7 1

1 2 3 4 5 6 7 4

**Степень оптимальности кода Прюфера**

Пусть *ℑ* **-** некоторый класс графов, *ℑ****n,m***– подмножество графов из *ℑ*, содержащих *n* вершин и *m* ребер; *B={0,1}*, *B\*-*множество слов в алфавите *B*. ***Кодированием*** графов из класса *ℑ* называется семейство взаимно однозначных отображений *Φ = {ϕ n,m : n=1,2,…; m=1,2,… ;}*, где *ϕ n,m*: *ℑ****n,m→*** *B\*.* Слово из *B\**, сопоставляемое графу *G∈ℑ****n,m***, называется ***кодом*** графа. Величина *|ϕ n,m(G)|* - ***длина*** кода графа *G∈ℑ****n,m***. При любом способе кодирования имеет место соотношение

*log2|ℑ****n,m*** *|* ***≤*** *max(|ϕ n,m(G)|, G∈ℑ****n,m****)*.(1)

Для обоснования приведенного соотношения достаточно заметить, что если перечислить в двоичной системе счисления графы из множества *ℑ* ***n,m***, то наибольшая длина кодов будет не менее *log2|ℑ****n,m***|бит. Поскольку коды графов должны обеспечивать еще и возможность их однозначного восстановления, то это в общем случае потребует дополнительного увеличения длины кода.

Если в соотношении (1) для любых *n* и *m* выполняется равенство, то соответствующее кодирование называется ***оптимальным***. Если имеет место

 , (2)

то кодирование называется ***асимптотически оптимальным****.*

Пусть *Tn* класс всех *n*- вершинных деревьев*.* Для графов, являющихся деревьями, будем пользоваться обозначением *t,* учитывая (1), получаем для деревьев соотношение

*log2|T****n*** *|*≤ *max(|ϕ n(t)|, t∈T****n****)**.*

По теореме Кэли имеем *|Tn|=nn-2* и следовательно *log2|T****n****|≥ (n-2)*⎣*log2 n*⎦ . Так как код Прюфера состоит из *(n-2)* чисел, каждое из которых принадлежит диапазону от 1 до *n*, то его длина |*ϕ n(t)*| ≤ *(n-2)*⎡*log2 n*⎤. Объединяя оценки, имеем *(n-2)*⎣*log2 n*⎦ ≤ *log2|T****n*** *|*≤ *max(|ϕ n(t)|, t∈T****n****)*≤ *(n-2)*⎡*log2 n*⎤ .

При этом предел (2) равен единице

.

Следовательно, рассмотренное кодирование помеченных деревьев является асимптотически оптимальным. На подпоследовательности *n=2i, i=1,2,…* кодирование является оптимальным.