Оптимальный префиксный код с длиной кодового слова не более L бит — это код, в котором длина каждого кодового слова не должна превышать заданной константы. Здесь будет приведен алгоритм, решающий эту задачу за время O(nL), где L — максимальная длина кодового слова, n — размер алфавита, с помощью сведения задачи к задаче о рюкзаке.

Пусть L — ограничение на длину кодового слова, а P={p1,p2,...,pn} — частоты символов алфавита. Алгоритм генерации кода будет следующим:

1. ***Отсортируем символы алфавита в порядке возрастания их частот.***
2. ***Для каждого символа создадим L предметов ценностью 2^{-1} … 2^{-L}, каждый из которых имеет вес pi.***
3. ***С помощью задачи о рюкзаке выберем набор предметов суммарной ценностью n−1 (nn — размер алфавита) с минимальным суммарным весом.***
4. ***Посчитаем массив H={h1,h2,...,hn}, где hi — количество предметов ценностью pi, которые попали в наш набор.***

При этом hi — это длина кодового слова для i-го символа.

Пример работы алгоритма

Пусть A = {A, B, C} – алфавит из трех различных символов, P = {1, 2, 3} – соответствующий ему набор. Пусть L = 2 – ограничение на длину кодового слова. Сначала создадим набор предметов.

Символ	Частота	Предметы
A	1	$(2^{-1};1),(2^{-2};1)$
B	2	$(2^{-1};2),(2^{-2};2)$
C	3	$(2^{-1};3),(2^{-2};3)$

Решим задачу о рюкзаке для заданного набора и выберем предметы суммарной ценностью n – 1 = 2 с минимальным суммарным весом. В нашем случаи в оптимальный набор попадут следующие предметы.

$$(2^{-1};1),(2^{-1};2),(2^{-1};3),(2^{-2};1),(2^{-2};2)$$

Посчитаем массив H

H = {2, 2, 1}

Итак мы получили длины кодовых слов для символов.