

# Тема: Алгоритм Хаффмана

Сергей Витальевич Рыбин  
[svrybin@etu.ru](mailto:svrybin@etu.ru)

СПбГЭТУ «ЛЭТИ», кафедра «Алгоритмической математики»

21 января 2023 г.



СПбГЭТУ «ЛЭТИ»  
ПЕРВЫЙ ЭЛЕКТРОТЕХНИЧЕСКИЙ



- 1 Клод Шеннон и Роберт Фано (Fano) предложили конструкцию **кода переменной длины**. При таком коде у каждого символа своя длина кодовой последовательности, поэтому необходимо побеспокоиться о том, как определять конец кода отдельного символа.

- 1 Клод Шеннон и Роберт Фано (Fano) предложили конструкцию **кода переменной длины**. При таком коде у каждого символа своя длина кодовой последовательности, поэтому необходимо побеспокоиться о том, как определять конец кода отдельного символа.
- 2 Предлагается такое ограничение на код: **никакая кодовая последовательность не является началом другой кодовой последовательности**.

- 1 Клод Шеннон и Роберт Фано (Fano) предложили конструкцию **кода переменной длины**. При таком коде у каждого символа своя длина кодовой последовательности, поэтому необходимо побеспокоиться о том, как определять конец кода отдельного символа.
- 2 Предлагается такое ограничение на код: **никакая кодовая последовательность не является началом другой кодовой последовательности**.
- 3 Это свойство называют **свойством префикса**, а код, обладающий таким свойством, — **беспрефиксным кодом**.

- 1 Клод Шеннон и Роберт Фано (Fano) предложили конструкцию **кода переменной длины**. При таком коде у каждого символа своя длина кодовой последовательности, поэтому необходимо побеспокоиться о том, как определять конец кода отдельного символа.
- 2 Предлагается такое ограничение на код: **никакая кодовая последовательность не является началом другой кодовой последовательности**.
- 3 Это свойство называют **свойством префикса**, а код, обладающий таким свойством, — **беспрефиксным кодом**.
- 4 Шеннон и Фано предложили метод построения кода, близкого к оптимальному: разбить все символы на две группы с приблизительно одинаковой суммарной частотой появления, коды первой группы начать с нуля, а второй — с единицы. Внутри группы делать то же самое, пока в каждой группе не останется только по одному символу.

- 1 Клод Шеннон и Роберт Фано (Fano) предложили конструкцию **кода переменной длины**. При таком коде у каждого символа своя длина кодовой последовательности, поэтому необходимо побеспокоиться о том, как определять конец кода отдельного символа.
- 2 Предлагается такое ограничение на код: **никакая кодовая последовательность не является началом другой кодовой последовательности**.
- 3 Это свойство называют **свойством префикса**, а код, обладающий таким свойством, — **беспрефиксным кодом**.
- 4 Шеннон и Фано предложили метод построения кода, близкого к оптимальному: разбить все символы на две группы с приблизительно одинаковой суммарной частотой появления, коды первой группы начать с нуля, а второй — с единицы. Внутри группы делать то же самое, пока в каждой группе не останется только по одному символу.
- 5 Элегантный алгоритм для точного решения этой задачи, основанный на нескольких очевидных свойствах оптимальной кодовой последовательности, предложил Дэвид Хаффман.

- 1 Клод Шеннон и Роберт Фано (Fano) предложили конструкцию **кода переменной длины**. При таком коде у каждого символа своя длина кодовой последовательности, поэтому необходимо побеспокоиться о том, как определять конец кода отдельного символа.
- 2 Предлагается такое ограничение на код: **никакая кодовая последовательность не является началом другой кодовой последовательности**.
- 3 Это свойство называют **свойством префикса**, а код, обладающий таким свойством, — **беспрефиксным кодом**.
- 4 Шеннон и Фано предложили метод построения кода, близкого к оптимальному: разбить все символы на две группы с приблизительно одинаковой суммарной частотой появления, коды первой группы начать с нуля, а второй — с единицы. Внутри группы делать то же самое, пока в каждой группе не останется только по одному символу.
- 5 Элегантный алгоритм для точного решения этой задачи, основанный на нескольких очевидных свойствах оптимальной кодовой последовательности, предложил Дэвид Хаффман.
- 6 Два символа кодируем 0 и 1, а если символов больше, то соединяем два самых редких символа в один новый символ, решаем получившуюся задачу, а затем вновь разделяем этот новый символ на два, приписав соответственно 0 и 1 к его кодовой последовательности.



- 1 Клод Шеннон и Роберт Фано (Fano) предложили конструкцию **кода переменной длины**. При таком коде у каждого символа своя длина кодовой последовательности, поэтому необходимо побеспокоиться о том, как определять конец кода отдельного символа.
- 2 Предлагается такое ограничение на код: **никакая кодовая последовательность не является началом другой кодовой последовательности**.
- 3 Это свойство называют **свойством префикса**, а код, обладающий таким свойством, — **беспрефиксным кодом**.
- 4 Шеннон и Фано предложили метод построения кода, близкого к оптимальному: разбить все символы на две группы с приблизительно одинаковой суммарной частотой появления, коды первой группы начать с нуля, а второй — с единицы. Внутри группы делать то же самое, пока в каждой группе не останется только по одному символу.
- 5 Элегантный алгоритм для точного решения этой задачи, основанный на нескольких очевидных свойствах оптимальной кодовой последовательности, предложил Дэвид Хаффман.
- 6 Два символа кодируем 0 и 1, а если символов больше, то соединяем два самых редких символа в один новый символ, решаем получившуюся задачу, а затем вновь разделяем этот новый символ на два, приписав соответственно 0 и 1 к его кодовой последовательности.
- 7 Для определенности, символу с большей частотностью приписываем **справа 1**, с меньшей — 0. Очевидно, что при обратном правиле получится инверсный код.

- 1 Клод Шеннон и Роберт Фано (Fano) предложили конструкцию **кода переменной длины**. При таком коде у каждого символа своя длина кодовой последовательности, поэтому необходимо побеспокоиться о том, как определять конец кода отдельного символа.
- 2 Предлагается такое ограничение на код: **никакая кодовая последовательность не является началом другой кодовой последовательности**.
- 3 Это свойство называют **свойством префикса**, а код, обладающий таким свойством, — **беспрефиксным кодом**.
- 4 Шеннон и Фано предложили метод построения кода, близкого к оптимальному: разбить все символы на две группы с приблизительно одинаковой суммарной частотой появления, коды первой группы начать с нуля, а второй — с единицы. Внутри группы делать то же самое, пока в каждой группе не останется только по одному символу.
- 5 Элегантный алгоритм для точного решения этой задачи, основанный на нескольких очевидных свойствах оптимальной кодовой последовательности, предложил Дэвид Хаффман.
- 6 Два символа кодируем 0 и 1, а если символов больше, то соединяем два самых редких символа в один новый символ, решаем получившуюся задачу, а затем вновь разделяем этот новый символ на два, приписав соответственно 0 и 1 к его кодовой последовательности.
- 7 Для определенности, символу с большей частотностью приписываем **справа 1**, с меньшей — 0. Очевидно, что при обратном правиле получится инверсный код.

**Недостаток:** предполагается, что вероятность появления текущего символа не зависит от предыдущих.



## Задача

Пусть имеется 7 кодовых символов  $a, b, c, d, e, f, g$  с частотами появления соответственно

$0.22(a), 0.20(b), 0.15(c), 0.13(d), 0.12(e), 0.10(f), 0.08(g)$ .

## Задача

Пусть имеется 7 кодовых символов  $a, b, c, d, e, f, g$  с частотами появления соответственно

$0.22(a), 0.20(b), 0.15(c), 0.13(d), 0.12(e), 0.10(f), 0.08(g)$ .

С помощью алгоритма Хаффмена построить код Шеннона–Фэно.

## Задача

Пусть имеется 7 кодовых символов  $a, b, c, d, e, f, g$  с частотами появления соответственно

$0.22(a), 0.20(b), 0.15(c), 0.13(d), 0.12(e), 0.10(f), 0.08(g)$ .

С помощью алгоритма Хаффмена построить код Шеннона–Фэно.

## Решение: объединение символов

## Задача

Пусть имеется 7 кодовых символов  $a, b, c, d, e, f, g$  с частотами появления соответственно

$0.22(a), 0.20(b), 0.15(c), 0.13(d), 0.12(e), 0.10(f), 0.08(g)$ .

С помощью алгоритма Хаффмена построить код Шеннона–Фэно.

## Решение: объединение символов

1  $0.22(a), 0.20(b), 0.18(fg), 0.15(c), 0.13(d), 0.12(e)$ ;

## Задача

Пусть имеется 7 кодовых символов  $a, b, c, d, e, f, g$  с частотами появления соответственно

$0.22(a), 0.20(b), 0.15(c), 0.13(d), 0.12(e), 0.10(f), 0.08(g)$ .

С помощью алгоритма Хаффмена построить код Шеннона–Фэно.

## Решение: объединение символов

- 1  $0.22(a), 0.20(b), 0.18(fg), 0.15(c), 0.13(d), 0.12(e)$ ;
- 2  $0.25(de), 0.22(a), 0.20(b), 0.18(fg), 0.15(c)$ ;



## Задача

Пусть имеется 7 кодовых символов  $a, b, c, d, e, f, g$  с частотами появления соответственно

$$0.22(a), 0.20(b), 0.15(c), 0.13(d), 0.12(e), 0.10(f), 0.08(g).$$

С помощью алгоритма Хаффмена построить код Шеннона–Фэно.

## Решение: объединение символов

- 1  $0.22(a), 0.20(b), 0.18(fg), 0.15(c), 0.13(d), 0.12(e);$
- 2  $0.25(de), 0.22(a), 0.20(b), 0.18(fg), 0.15(c);$
- 3  $0.33(cfg), 0.25(de), 0.22(a), 0.20(b);$

## Задача

Пусть имеется 7 кодовых символов  $a, b, c, d, e, f, g$  с частотами появления соответственно

$$0.22(a), 0.20(b), 0.15(c), 0.13(d), 0.12(e), 0.10(f), 0.08(g).$$

С помощью алгоритма Хаффмена построить код Шеннона–Фэно.

## Решение: объединение символов

- 1  $0.22(a), 0.20(b), 0.18(fg), 0.15(c), 0.13(d), 0.12(e);$
- 2  $0.25(de), 0.22(a), 0.20(b), 0.18(fg), 0.15(c);$
- 3  $0.33(cfg), 0.25(de), 0.22(a), 0.20(b);$
- 4  $0.42(ab), 0.33(cfg), 0.25(de);$

## Задача

Пусть имеется 7 кодовых символов  $a, b, c, d, e, f, g$  с частотами появления соответственно

$0.22(a), 0.20(b), 0.15(c), 0.13(d), 0.12(e), 0.10(f), 0.08(g)$ .

С помощью алгоритма Хаффмена построить код Шеннона–Фэно.

## Решение: объединение символов

- 1  $0.22(a), 0.20(b), 0.18(fg), 0.15(c), 0.13(d), 0.12(e)$ ;
- 2  $0.25(de), 0.22(a), 0.20(b), 0.18(fg), 0.15(c)$ ;
- 3  $0.33(cfg), 0.25(de), 0.22(a), 0.20(b)$ ;
- 4  $0.42(ab), 0.33(cfg), 0.25(de)$ ;
- 5  $0.58(cfgde), 0.42(ab)$ .

## Задача

Пусть имеется 7 кодовых символов  $a, b, c, d, e, f, g$  с частотами появления соответственно

$0.22(a), 0.20(b), 0.15(c), 0.13(d), 0.12(e), 0.10(f), 0.08(g)$ .

С помощью алгоритма Хаффмена построить код Шеннона–Фэно.

## Решение: объединение символов

- 1  $0.22(a), 0.20(b), 0.18(fg), 0.15(c), 0.13(d), 0.12(e)$ ;
- 2  $0.25(de), 0.22(a), 0.20(b), 0.18(fg), 0.15(c)$ ;
- 3  $0.33(cfg), 0.25(de), 0.22(a), 0.20(b)$ ;
- 4  $0.42(ab), 0.33(cfg), 0.25(de)$ ;
- 5  $0.58(cfgde), 0.42(ab)$ .

## Решение: расщепление символов

## Задача

Пусть имеется 7 кодовых символов  $a, b, c, d, e, f, g$  с частотами появления соответственно

$$0.22(a), 0.20(b), 0.15(c), 0.13(d), 0.12(e), 0.10(f), 0.08(g).$$

С помощью алгоритма Хаффмена построить код Шеннона–Фэно.

## Решение: объединение символов

- 1  $0.22(a), 0.20(b), 0.18(fg), 0.15(c), 0.13(d), 0.12(e);$
- 2  $0.25(de), 0.22(a), 0.20(b), 0.18(fg), 0.15(c);$
- 3  $0.33(cfg), 0.25(de), 0.22(a), 0.20(b);$
- 4  $0.42(ab), 0.33(cfg), 0.25(de);$
- 5  $0.58(cfgde), 0.42(ab).$

## Решение: расщепление символов

- 1  $0(ab), 1(cfgde);$

# Пример

## Задача

Пусть имеется 7 кодовых символов  $a, b, c, d, e, f, g$  с частотами появления соответственно

$$0.22(a), 0.20(b), 0.15(c), 0.13(d), 0.12(e), 0.10(f), 0.08(g).$$

С помощью алгоритма Хаффмена построить код Шеннона–Фэно.

## Решение: объединение символов

- 1  $0.22(a), 0.20(b), 0.18(fg), 0.15(c), 0.13(d), 0.12(e);$
- 2  $0.25(de), 0.22(a), 0.20(b), 0.18(fg), 0.15(c);$
- 3  $0.33(cfg), 0.25(de), 0.22(a), 0.20(b);$
- 4  $0.42(ab), 0.33(cfg), 0.25(de);$
- 5  $0.58(cfgde), 0.42(ab).$

## Решение: расщепление символов

- 1  $0(ab), 1(cfgde);$
- 2  $0(ab), 11(cfg), 10(de);$

## Задача

Пусть имеется 7 кодовых символов  $a, b, c, d, e, f, g$  с частотами появления соответственно

$$0.22(a), 0.20(b), 0.15(c), 0.13(d), 0.12(e), 0.10(f), 0.08(g).$$

С помощью алгоритма Хаффмена построить код Шеннона–Фэно.

## Решение: объединение символов

- 1  $0.22(a), 0.20(b), 0.18(fg), 0.15(c), 0.13(d), 0.12(e);$
- 2  $0.25(de), 0.22(a), 0.20(b), 0.18(fg), 0.15(c);$
- 3  $0.33(cfg), 0.25(de), 0.22(a), 0.20(b);$
- 4  $0.42(ab), 0.33(cfg), 0.25(de);$
- 5  $0.58(cfgde), 0.42(ab).$

## Решение: расщепление символов

- 1  $0(ab), 1(cfgde);$
- 2  $0(ab), 11(cfg), 10(de);$
- 3  $01(a), 00(b), 11(cfg), 10(de);$

## Задача

Пусть имеется 7 кодовых символов  $a, b, c, d, e, f, g$  с частотами появления соответственно

$$0.22(a), 0.20(b), 0.15(c), 0.13(d), 0.12(e), 0.10(f), 0.08(g).$$

С помощью алгоритма Хаффмена построить код Шеннона–Фэно.

## Решение: объединение символов

- 1  $0.22(a), 0.20(b), 0.18(fg), 0.15(c), 0.13(d), 0.12(e);$
- 2  $0.25(de), 0.22(a), 0.20(b), 0.18(fg), 0.15(c);$
- 3  $0.33(cfg), 0.25(de), 0.22(a), 0.20(b);$
- 4  $0.42(ab), 0.33(cfg), 0.25(de);$
- 5  $0.58(cfgde), 0.42(ab).$

## Решение: расщепление символов

- 1  $0(ab), 1(cfgde);$
- 2  $0(ab), 11(cfg), 10(de);$
- 3  $01(a), 00(b), 11(cfg), 10(de);$
- 4  $01(a), 00(b), 110(c), 111(fg), 10(de);$



## Задача

Пусть имеется 7 кодовых символов  $a, b, c, d, e, f, g$  с частотами появления соответственно

$$0.22(a), 0.20(b), 0.15(c), 0.13(d), 0.12(e), 0.10(f), 0.08(g).$$

С помощью алгоритма Хаффмена построить код Шеннона–Фэно.

## Решение: объединение символов

- 1  $0.22(a), 0.20(b), 0.18(fg), 0.15(c), 0.13(d), 0.12(e);$
- 2  $0.25(de), 0.22(a), 0.20(b), 0.18(fg), 0.15(c);$
- 3  $0.33(cfg), 0.25(de), 0.22(a), 0.20(b);$
- 4  $0.42(ab), 0.33(cfg), 0.25(de);$
- 5  $0.58(cfgde), 0.42(ab).$

## Решение: расщепление символов

- 1  $0(ab), 1(cfgde);$
- 2  $0(ab), 11(cfg), 10(de);$
- 3  $01(a), 00(b), 11(cfg), 10(de);$
- 4  $01(a), 00(b), 110(c), 111(fg), 10(de);$
- 5  $01(a), 00(b), 110(c), 111(fg), 101(d), 100(e);$

# Пример

## Задача

Пусть имеется 7 кодовых символов  $a, b, c, d, e, f, g$  с частотами появления соответственно

$$0.22(a), 0.20(b), 0.15(c), 0.13(d), 0.12(e), 0.10(f), 0.08(g).$$

С помощью алгоритма Хаффмена построить код Шеннона–Фэно.

## Решение: объединение символов

- 1  $0.22(a), 0.20(b), 0.18(fg), 0.15(c), 0.13(d), 0.12(e);$
- 2  $0.25(de), 0.22(a), 0.20(b), 0.18(fg), 0.15(c);$
- 3  $0.33(cfg), 0.25(de), 0.22(a), 0.20(b);$
- 4  $0.42(ab), 0.33(cfg), 0.25(de);$
- 5  $0.58(cfgde), 0.42(ab).$

## Решение: расщепление символов

- 1  $0(ab), 1(cfgde);$
- 2  $0(ab), 11(cfg), 10(de);$
- 3  $01(a), 00(b), 11(cfg), 10(de);$
- 4  $01(a), 00(b), 110(c), 111(fg), 10(de);$
- 5  $01(a), 00(b), 110(c), 111(fg), 101(d), 100(e);$
- 6  $01(a), 00(b), 110(c), 1111(f), 1110(g), 101(d), 100(e).$

# Пример

## Задача

Пусть имеется 7 кодовых символов  $a, b, c, d, e, f, g$  с частотами появления соответственно

$$0.22(a), 0.20(b), 0.15(c), 0.13(d), 0.12(e), 0.10(f), 0.08(g).$$

С помощью алгоритма Хаффмена построить код Шеннона–Фэно.

## Решение: объединение символов

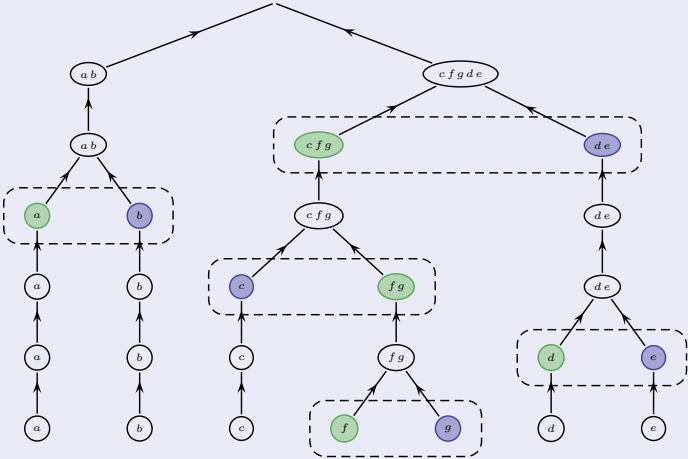
- 1  $0.22(a), 0.20(b), 0.18(fg), 0.15(c), 0.13(d), 0.12(e);$
- 2  $0.25(de), 0.22(a), 0.20(b), 0.18(fg), 0.15(c);$
- 3  $0.33(cfg), 0.25(de), 0.22(a), 0.20(b);$
- 4  $0.42(ab), 0.33(cfg), 0.25(de);$
- 5  $0.58(cfgde), 0.42(ab).$

## Решение: расщепление символов

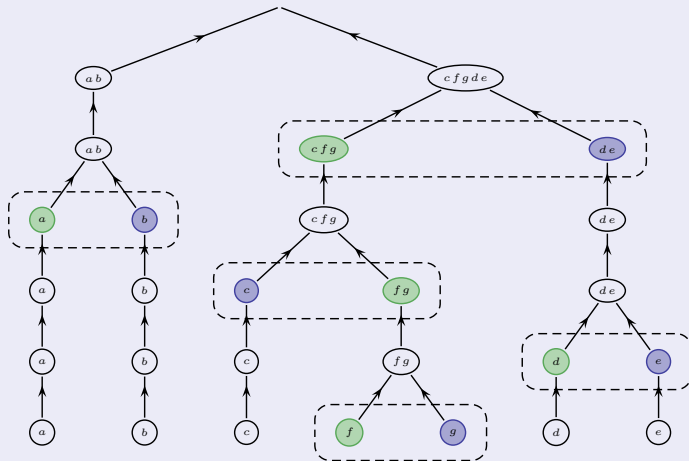
- 1  $0(ab), 1(cfgde);$
- 2  $0(ab), 11(cfg), 10(de);$
- 3  $01(a), 00(b), 11(cfg), 10(de);$
- 4  $01(a), 00(b), 110(c), 111(fg), 10(de);$
- 5  $01(a), 00(b), 110(c), 111(fg), 101(d), 100(e);$
- 6  $01(a), 00(b), 110(c), 1111(f), 1110(g), 101(d), 100(e).$

Полученный код:  $01 - a, 00 - b, 110 - c, 101 - d, 100 - e, 1111 - f, 1110 - g$



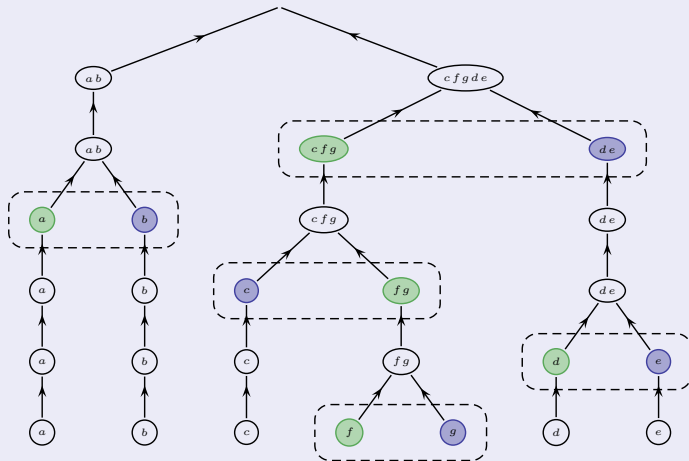


# Пример. Схема решения



**i** Алгоритм работает «снизу вверх» (движение по стрелкам). Объединяемые наборы символов отмечены серым цветом и пунктирными линиями.

# Пример. Схема решения



- i Алгоритм работает «снизу вверх» (движение по стрелкам). Объединяемые наборы символов отмечены серым цветом и пунктирными линиями.
- i При объединении символы с большей частотностью отмечены зеленым, а с меньшей — синим цветом.





- 1 С. В. Рыбин. Дискретная математика и информатика. — Лань, 2022.

- 1 С. В. Рыбин. Дискретная математика и информатика. — Лань, 2022.
- 2 С. Н. Поздняков, С. В. Рыбин. Дискретная математика. — Издательский центр «Академия», 2008.