

Lesson 4

Shannon-Fano Coding

Oleh :

Ledya Novamizanti

Astri Novianty

Prodi S1 Teknik Telekomunikasi
Fakultas Teknik Elektro
Universitas Telkom



Shannon-Fano Coding

- Shannon-Fano coding, dikembangkan oleh **Claude Shannon** di Bell Labs dan **Robert Fano** di MIT
- Merupakan algoritma pertama untuk membangun satu himpunan variable-length code terbaik.
- Shannon-fano coding ditentukan oleh probabilitas (atau frekuensi) kemunculan dari tiap simbol pada suatu pesan.
- Pembentukan diagram pohon dari Shannon-Fano coding dimulai dari **akar ke daun** (top-bottom)



Shannon-Fano Coding

- Sifat kode yang dihasilkan, yaitu:
 - ❖ Kode yang berbeda memiliki jumlah bit yang berbeda
 - ❖ Kode untuk simbol dengan probabilitas kemunculan rendah memiliki bit lebih banyak, dan kode untuk simbol dengan probabilitas kemunculan tinggi memiliki bit lebih sedikit
 - ❖ Walaupun setiap kode memiliki panjang bit yang berbeda, masing-masing dapat di-decode secara unik

Algoritma Shannon-Fano Coding

1. Baca semua simbol pada pesan (teks), dan tentukan probabilitas (atau frekuensi) kemunculan masing-masing simbol
2. Urutkan simbol berdasarkan probabilitas kemunculan secara descending. Jika probabilitas sama, urutkan indeks simbol ascending
3. Bagi rangkaian simbol menjadi 2 subset sedemikian hingga selisih total probabilitas antara 2 subset seminimal mungkin
4. Semua simbol dalam satu bagian diberi kode 0, dan di bagian lain diberi kode 1
5. Lakukan terus langkah ke-3 dan 4, hingga tidak ada lagi subset yang tersisa

Contoh 1

Diketahui 7 simbol A, B, C, D, E, F, G dengan probabilitas kemunculan masing-masing 0.25, 0.20, 0.15, 0.15, 0.10, 0.10, 0.05.

- a. Tentukan Shannon-Fano Code untuk ketujuh simbol tersebut, menggunakan tabel dan diagram pohon
- b. Hitung redundancy
- c. Hitung efficiency

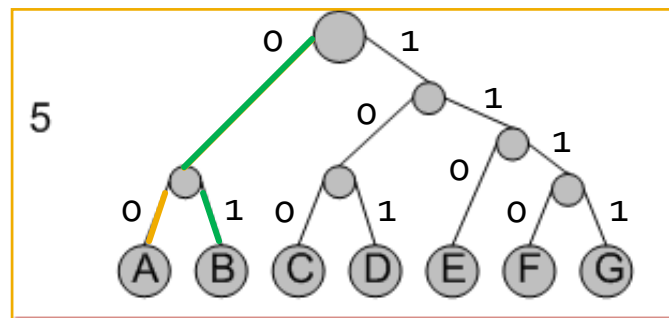
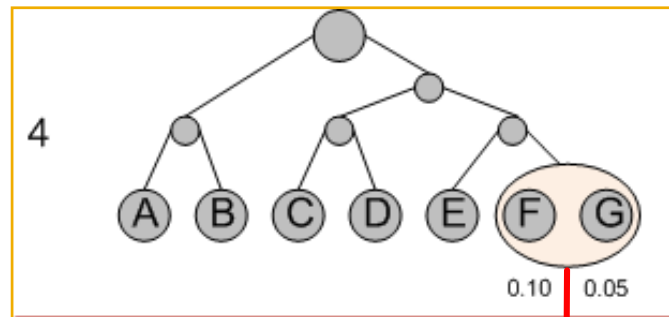
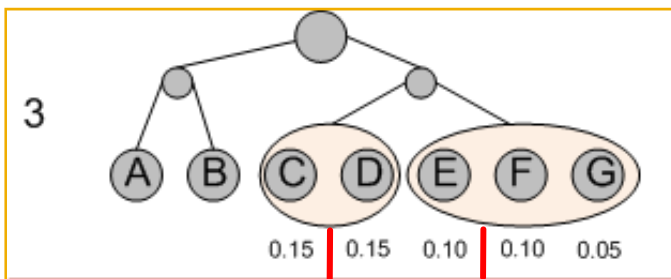
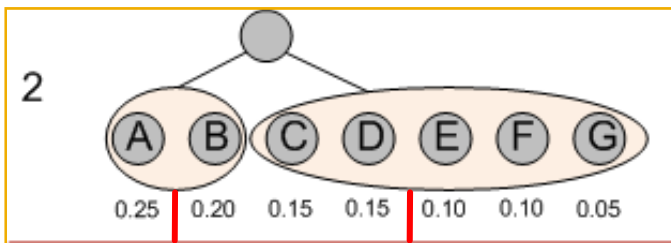
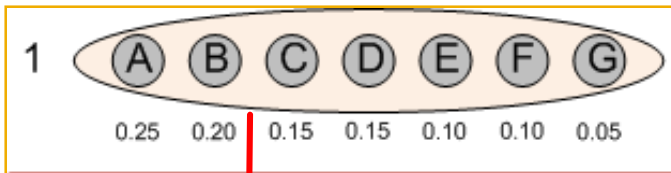
Solusi Contoh 1

a. Pencarian Shannon-Fano code menggunakan tabel

Symbol	Probability	Steps				Code
A	0.25	0	0	⊗	⊗	00
B	0.20	0	1	⊗	⊗	01
C	0.15	1	0	0	⊗	100
D	0.15	1	0	1	⊗	101
E	0.10	1	1	0	⊗	110
F	0.10	1	1	1	0	1110
G	0.05	1	1	1	1	1111

Solusi Contoh 1

Pencarian Shannon-Fano code menggunakan diagram pohon



Symbol	Code
A	00
B	01
C	100
D	101
E	110
F	1110
G	1111

Solusi Contoh 1

b. Redundancy, $R = L - H$

- Average length

$$L = \sum P(S_i) \times n(S_i)$$

$$L = 0.25 \times 2 + 0.20 \times 2 + 0.15 \times 3 + 0.15 \times 3 + 0.10 \times 3 + 0.10 \times 4 + 0.05 \times 4 = 2.7 \text{ bits/symbol}$$

- Entropy

$$H = - \sum P(S_i) \times \log_b P(S_i)$$

$$H = - (0.25 \log_2 0.25 + 0.20 \log_2 0.20 + 2 \times 0.15 \log_2 0.15 + 2 \times 0.10 \log_2 0.10 + 0.05 \log_2 0.05) \approx 2.67 \text{ bits/symbol}$$

Redundancy, $R = L - H = 2.7 - 2.67 = 0.03 \text{ bits/symbol}$ (1% lebih banyak dari entropy)

Symbol	Probability	Code
A	0.25	00
B	0.20	01
C	0.15	100
D	0.15	101
E	0.10	110
F	0.10	1110
G	0.05	1111

Solusi Contoh 1

Average Length (L) 2.7 bits/symbol
dan Entropy (H) 2.67 bits/symbol

c. Efficiency

$$\text{efficiency} = \frac{H}{L} \times 100\%$$

$$\text{efficiency} = \frac{2.67}{2.7} \times 100\% \approx 98.89\%$$

Contoh 2

Diketahui 6 simbol A, B, C, D, E, F dengan probabilitas kemunculan masing-masing 0.25, 0.25, 0.125, 0.125, 0.125, 0.125.

- a. Tentukan Shannon-Fano Code untuk ke-enam simbol tersebut
- b. Hitung redundancy
- c. Hitung efficiency

Solusi Contoh 2

a. Pencarian Shannon-Fano code menggunakan tabel

Symbol	Probability	Steps			Code
A	0.25	0	0	⊗	00
B	0.25	0	1	⊗	01
C	0.125	1	0	0	100
D	0.125	1	0	1	101
E	0.125	1	1	0	110
F	0.125	1	1	1	111

Solusi Contoh 2

b. Redundancy, $R = L - H$

■ Average length, $L = \sum P(S_i) \times n(S_i)$

$$L = 0.5 \times 2 + 0.5 \times 3 = 2.5 \text{ bit/simbol}$$

■ Entropy, $H = - \sum P(S_i) \times \log_b P(S_i)$

$$H = - (2 \times 0.25 \log_2 0.25 + 4 \times 0.125 \log_2 0.125) = 2.5 \text{ bit/simbol}$$

Redundancy, $R = L - H = 0 \text{ bit/simbol}$

c. Efficiency, $efficiency = \frac{H}{L} \times 100\% = 100\%$

Symbol	Probability	Code
A	0.25	00
B	0.25	01
C	0.125	100
D	0.125	101
E	0.125	110
F	0.125	111

Metode ini menghasilkan hasil terbaik ketika simbol memiliki probabilitas kejadian pangkat negatif dari 2.

Contoh 3

Gunakan tabel kode pada contoh 2, untuk encode pesan FABEDCAB

Solusi :

Symbol	Code
A	00
B	01
C	100
D	101
E	110
F	111

Pesan :	F	A	B	E	D	C	A	B
Encode :	111	00	01	110	101	100	00	01

Contoh 4

Dari hasil encode pada contoh 3, hitung rasio kompresi untuk pesan FABEDCAB

Solusi :

Pesan :	F	A	B	E	D	C	A	B
Encode :	111	00	01	110	101	100	00	01

→ 20 bit

Panjang pesan hasil kompresi = 20 bit

Panjang pesan asli (menggunakan kode ASCII) dari 8 huruf FABEDCAB yaitu 8 byte = 8×8 bit = 64 bit

Solusi Contoh 4

Original size = 64 bit

Compressed size = 20 bit

1.
$$\text{rasio kompresi} = \text{original size} : \text{compressed size}$$

$$\text{rasio kompresi} = 64 : 20 = \underline{3.2 : 1}$$

2.
$$\text{rasio kompresi (\%)} = \left(1 - \frac{\text{compressed size}}{\text{original size}} \right) \times 100$$

$$\text{rasio kompresi (\%)} = \left(1 - \frac{20}{64} \right) \times 100 = \underline{68.75}$$

Contoh 5

Gunakan tabel kode pada contoh 2, untuk decode rangkaian bit 11100011101011000001

Solusi :

Symbol	Code
A	00
B	01
C	100
D	101
E	110
F	111

Decode : 11100011101011000001
F A B E D C A B

Tugas 1

Terdapat source $S=\{a,e,i,o,u,! \}$ dengan probabilitas setiap simbol adalah $P = \{0.2, 0.3, 0.1, 0.2, 0.1, 0.1\}$.

- a. Tentukan kode Shannon-Fano dari source tersebut
- b. Hitung average length dan entropy
- c. Hitung redundancy kode
- d. Hitung variansi kode
- e. Hitung efficiency
- f. Hitung rasio kompresi

Tugas 2

Diketahui pesan string ada ada saja

- a. Tentukan kode Shannon-Fano dari tiap simbol pada string tersebut
- b. Hitung variansi kode
- c. Hitung average length dan entropy
- d. Hitung redundancy kode
- e. Hitung efficiency
- f. Hitung rasio kompresi

Referensi

1. Adam Drozdek, *Elements of Data Compression*, Thomson Brooks/Cole, 2002
2. Khalid Sayood, *Introduction to Data Compression*, Academic Press, 2000.
3. T.M. Cover, J.A. Thomas, *Elements of Information Theory*, John Wiley&Sons.
4. M. Nelson and J.-L. Gailly. *The Data Compression Book*. M&T Books, CA, 1996.
5. D. Salomon. *Data Compression: The Complete Reference*. Springer, 1998.

Thank you 😊

