Variable Length Coding: Shannon-Fano Algorithm

Introduction to Lossless Compression

Compression: the process of coding that will effectively reduce the total number of bits needed to represent certain information.
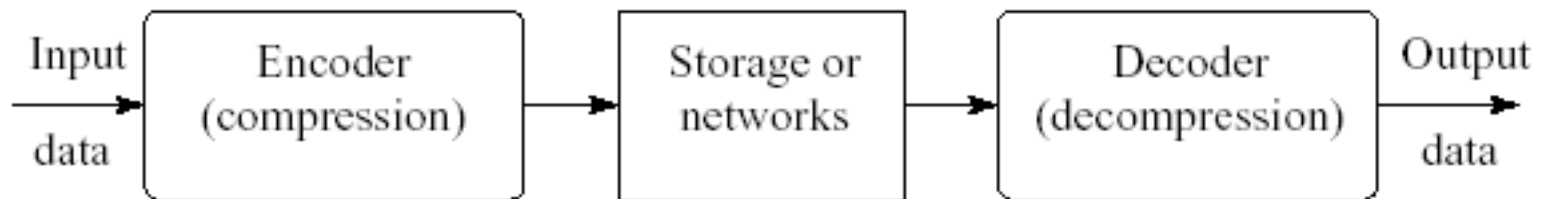

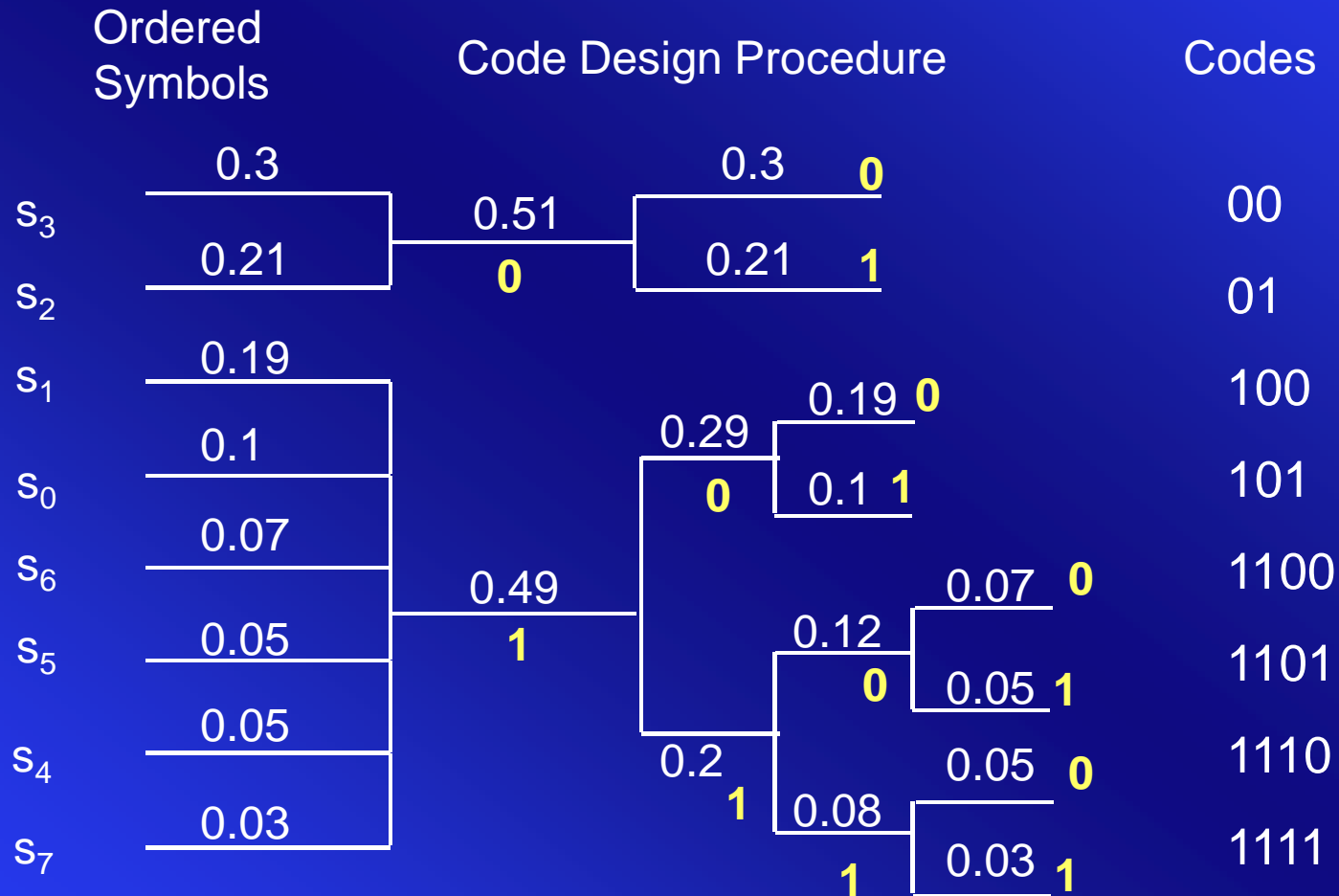Fig. 7.1: A General Data Compression Scheme.

Variable-Length Coding

- **Shannon-Fano Algorithm** a top-down approach
- 1. Sort the symbols according to the frequency count of their occurrences.
- 2. Recursively divide the symbols into two parts, each with approximately the same number of counts, until all parts contain only one symbol.

Shannon-Fano Coding

1. Sort set of symbols S in order of non-increasing probabilities.

2. Divide S into two parts S_1, S_2 such that each part has approximately equal probability

3. Assign '0'('1') to symbols of S_1(S_2)

4. Continue (2) and (3) on each of the parts until each part contains only one symbol

An Example of Shannon-Fano Coding

Ordered Symbols

Code Design Procedure

Codes

s_3 0.3

0.51   **0**

0.3   **0**   00

s_2 0.21

0.21   **1**   01

s_1 0.19

0.19   **0**   100

0.29   **0**

s_0 0.1

0.1   **1**   101

s_6 0.07

0.49   **1**

0.07   **0**   1100

0.12   **0**

s_5 0.05

0.05   **1**   1101

s_4 0.05

0.2   **1**

0.05   **0**   1110

0.08   **1**

s_7 0.03

0.03   **1**   1111

Assignment

Q.1. Explain Shanon-Fano encoding.