

Data Compression



`data.zip`

ASCII

A	01000001
B	01000010
C	01000011
D	01000100
E	01000101

Huffman Example

A	01
B	0000
C	0001
D	001
E	1

Prefix Property



A	0
B	1
C	01
D	11



A	0
B	10
C	110
D	111

Step 1: Frequency Analysis

String:

**“ECEABEADCAEDEEEECEADEEEEEEDB
AAEABDBBAAEAAACDDCCEABEEDCB
EEDEAEFFFFFFAEEDBCBEEADEAEEDA
EBCDEDEAEEDCEEAE”**

character	A	B	C	D	E
frequency	0.2	0.1	0.1	0.15	0.45

Step 2: Forest of Single Trees

B	0.1

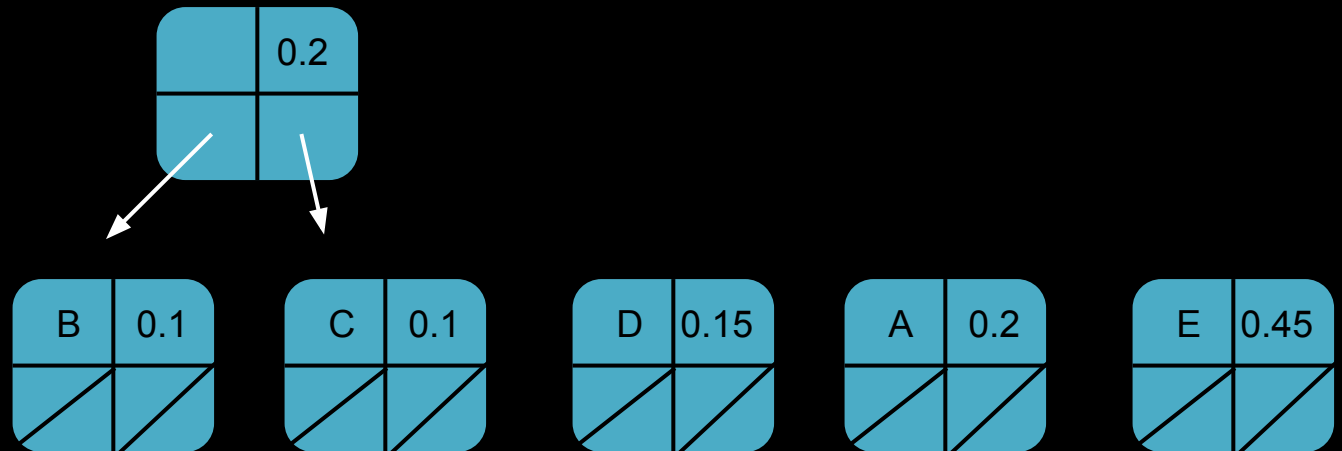
C	0.1

D	0.15

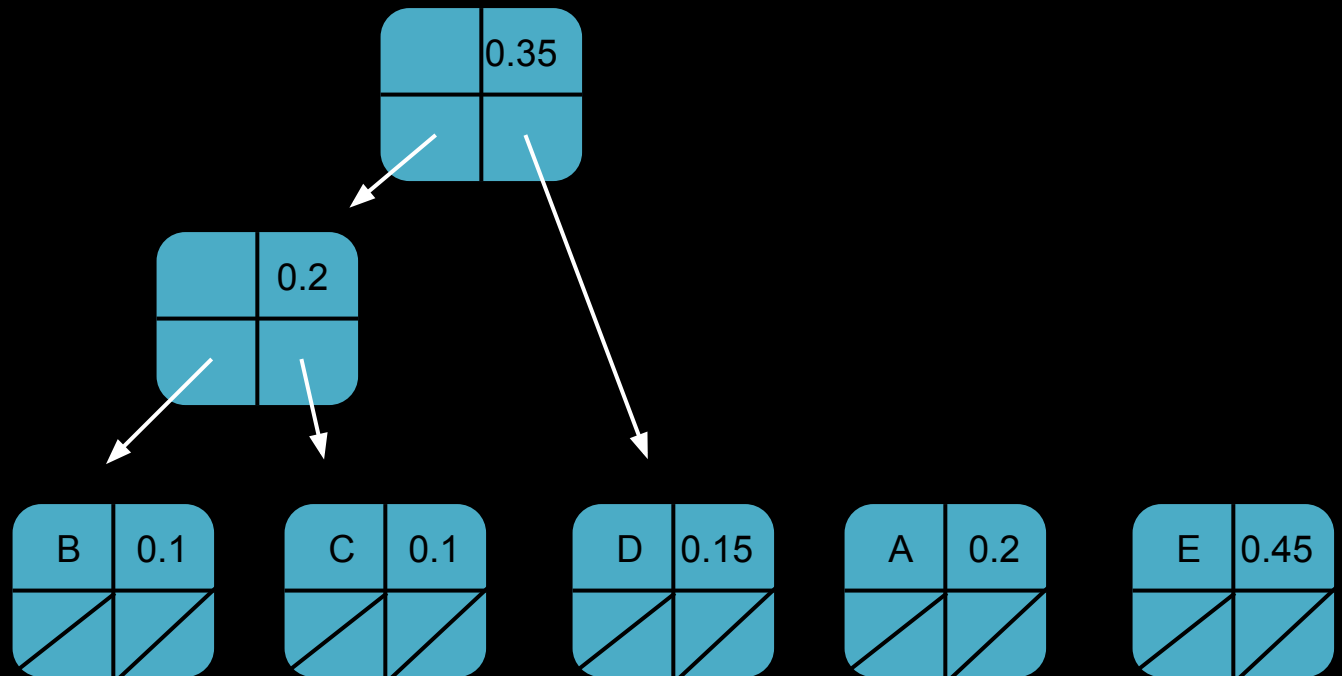
A	0.2

E	0.45

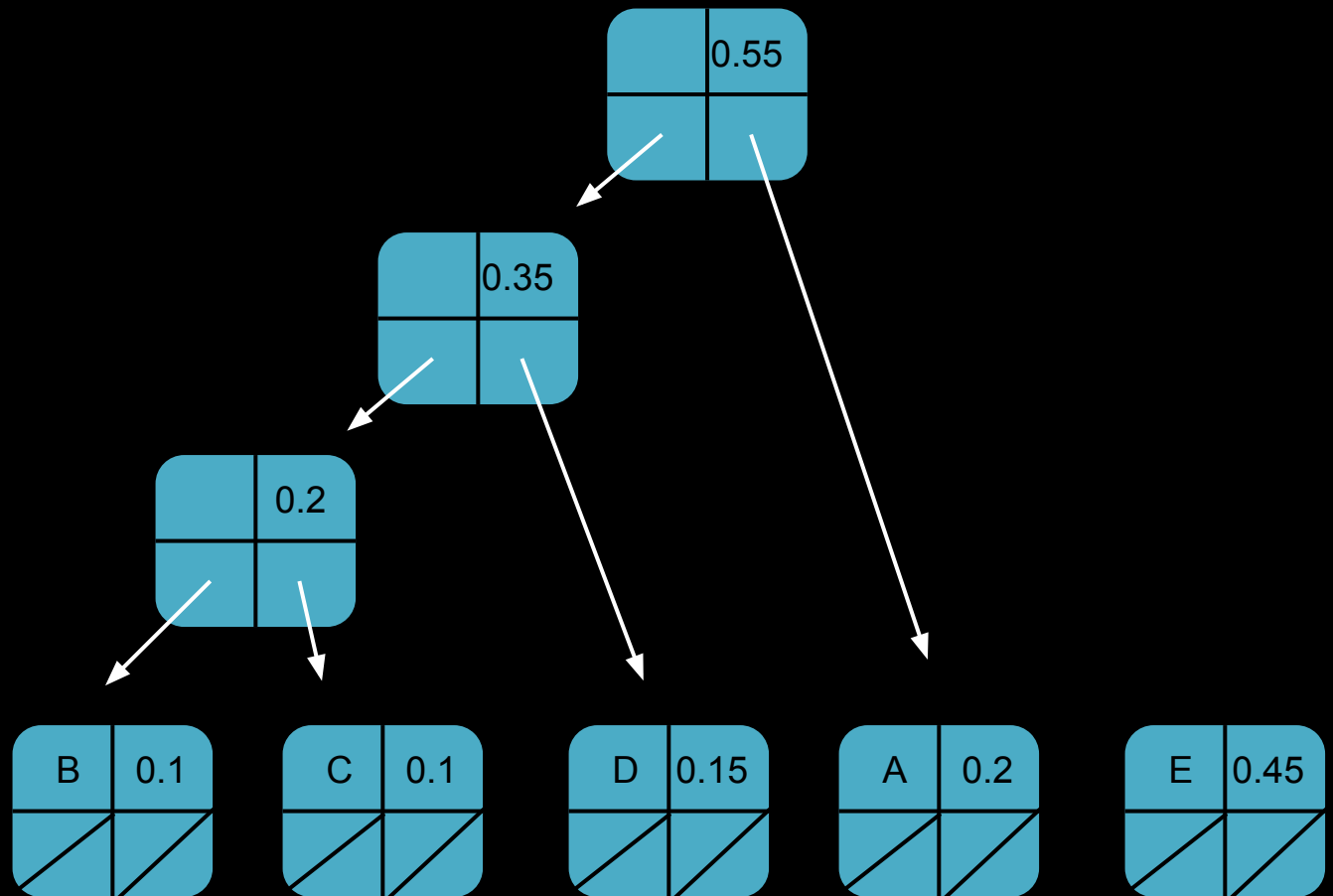
Step 3: Combine Trees



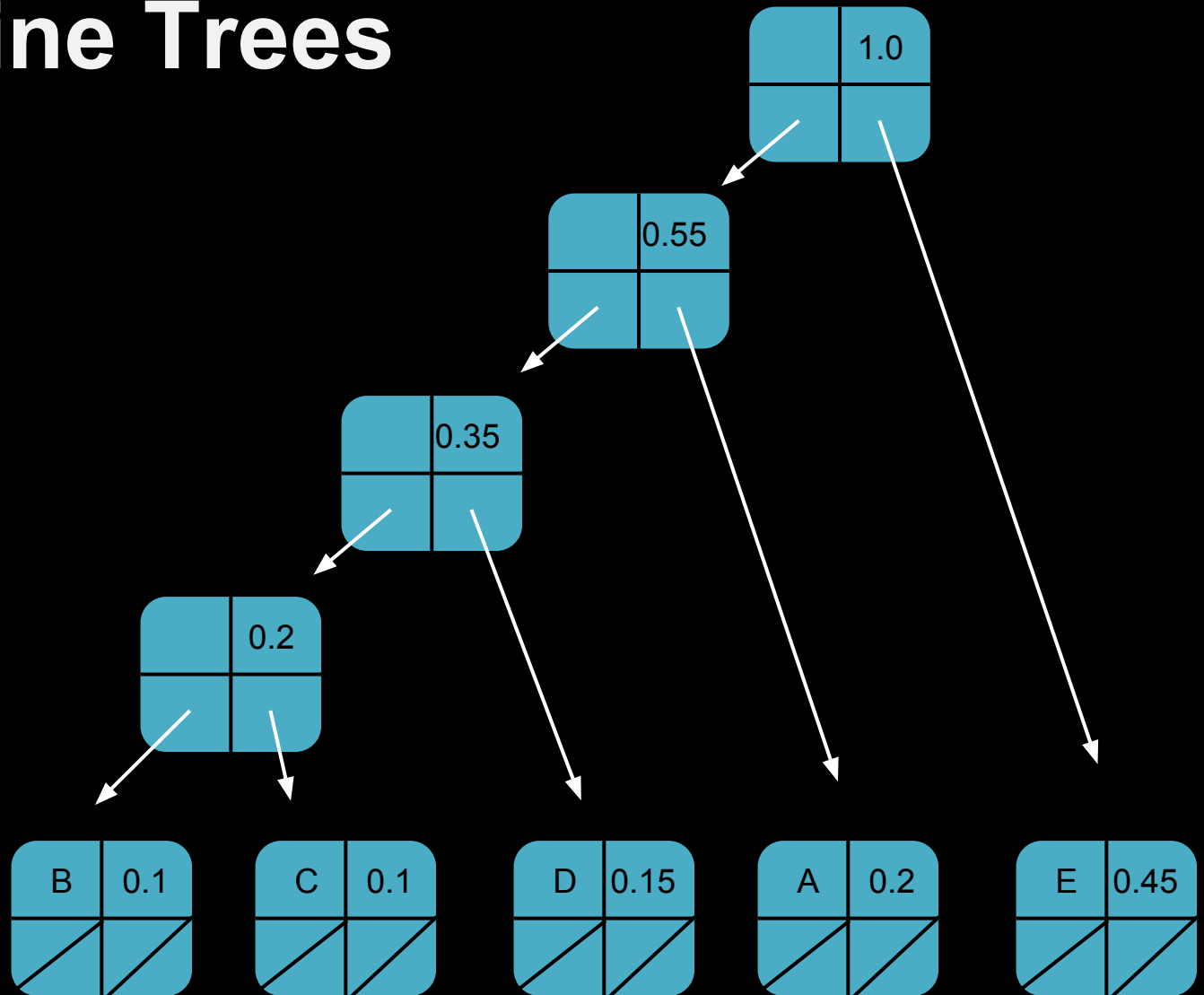
Step 3: Combine Trees



Step 3: Combine Trees



Step 3: Combine Trees



Step 4: Traverse Tree

