Data Compression
Lecture 7

Dr. Mona M.Soliman

Faculty of Computers and Artificial Intelligence

Cairo University

Fall 2022

Arithmetic Coding
Binary Coding

Binary Representation of Arithmetic Coding

Problems of Floating Point Representation of Arithmetic Coding

- Need high precision
- No output is generated until the entire sequence is encoded.

Binary Representation of Arithmetic Coding

Solving Precision Problem Using Scaling

If the whole **Range** is in the lower half (i.e. **0**<=**Range** <**0.5**)

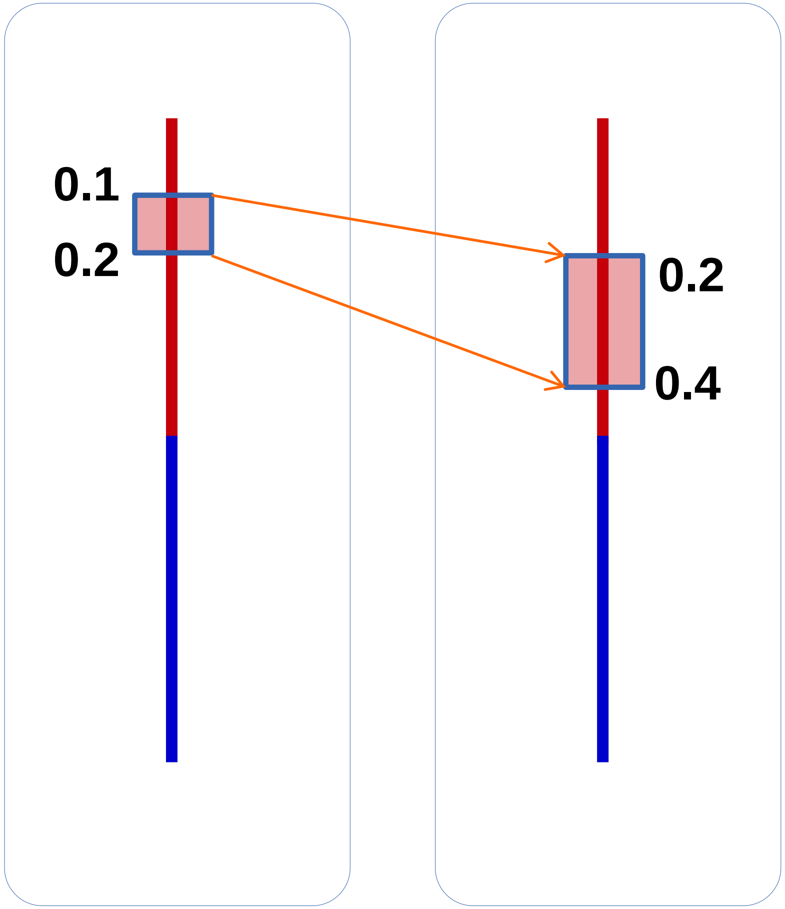
(Note: 0.5 is not included in the range)

Multiply Upper and Lower values of the range by "2"

The new range will be **0**<= **Range** < **1**

Call this scaling "**E1**" , indicate it by binary "**0**"

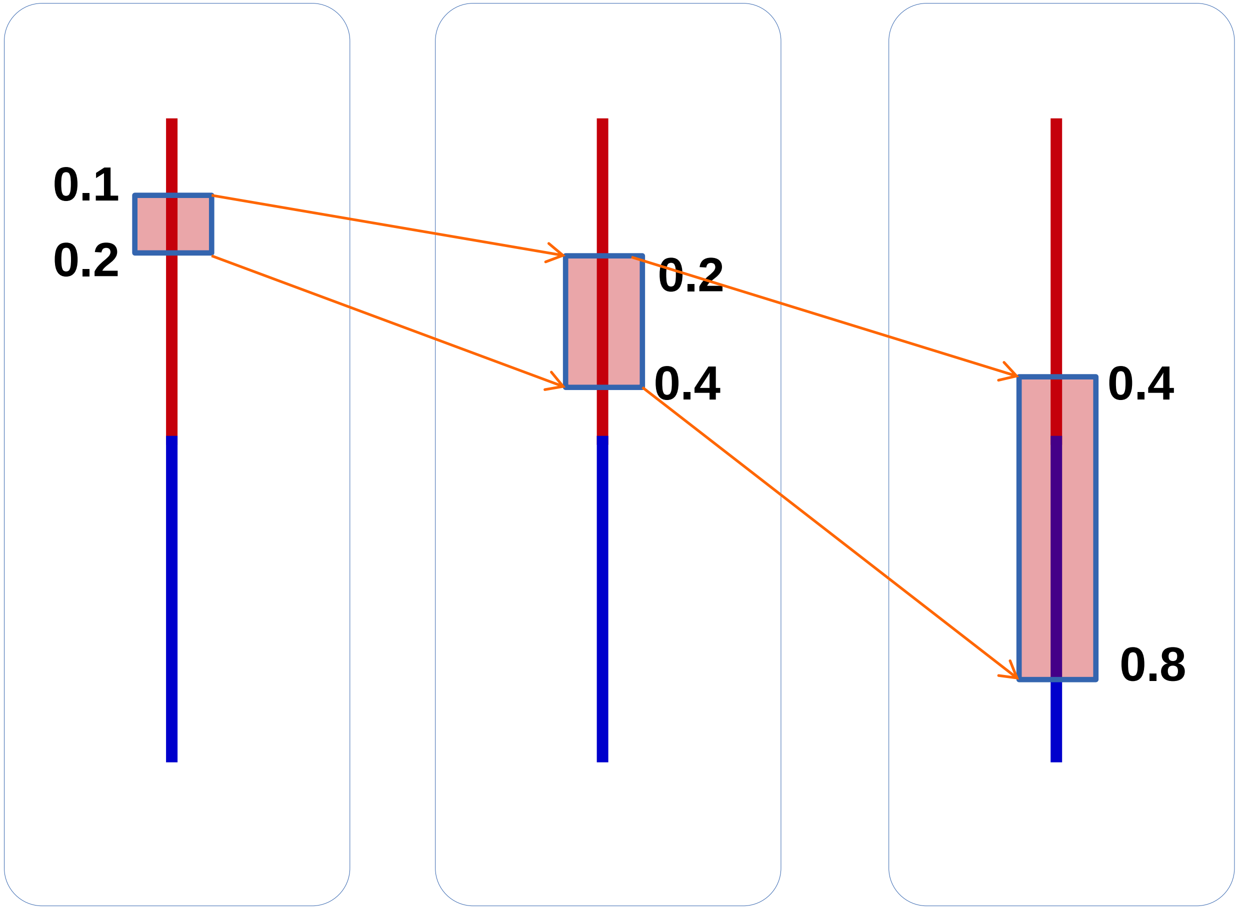
STOP when the mid point (0.5) will lay inside the range



E1 Scaling , Indicated by "0"

Binary Representation of Arithmetic Coding

STOP when The mid point (0.5) will lay inside the range



Scaling Code: E1,E1: 00

Binary Representation of Arithmetic Coding

Solving Precision Problem Using Scaling

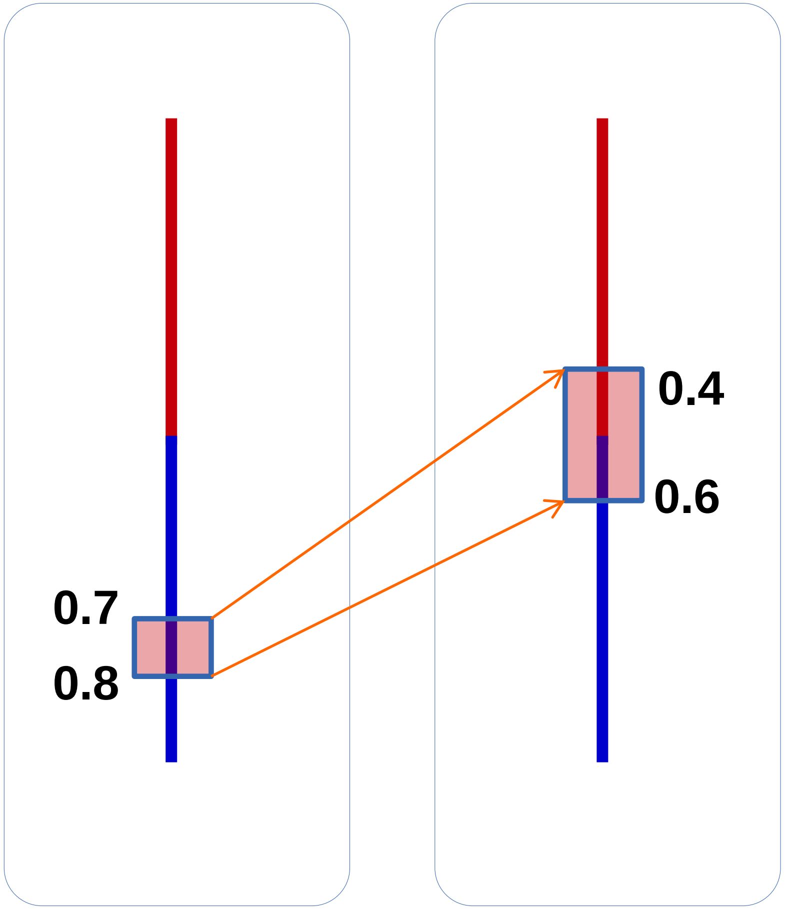
If the whole **Range** is in the upper half

(i.e. **0.5**<= **Range** <**1**)

(Note: 0.5 is not included in the range)

- Subtract "0.5" from Upper and Lower values of the range, then multiply Upper and Lower values by "2"

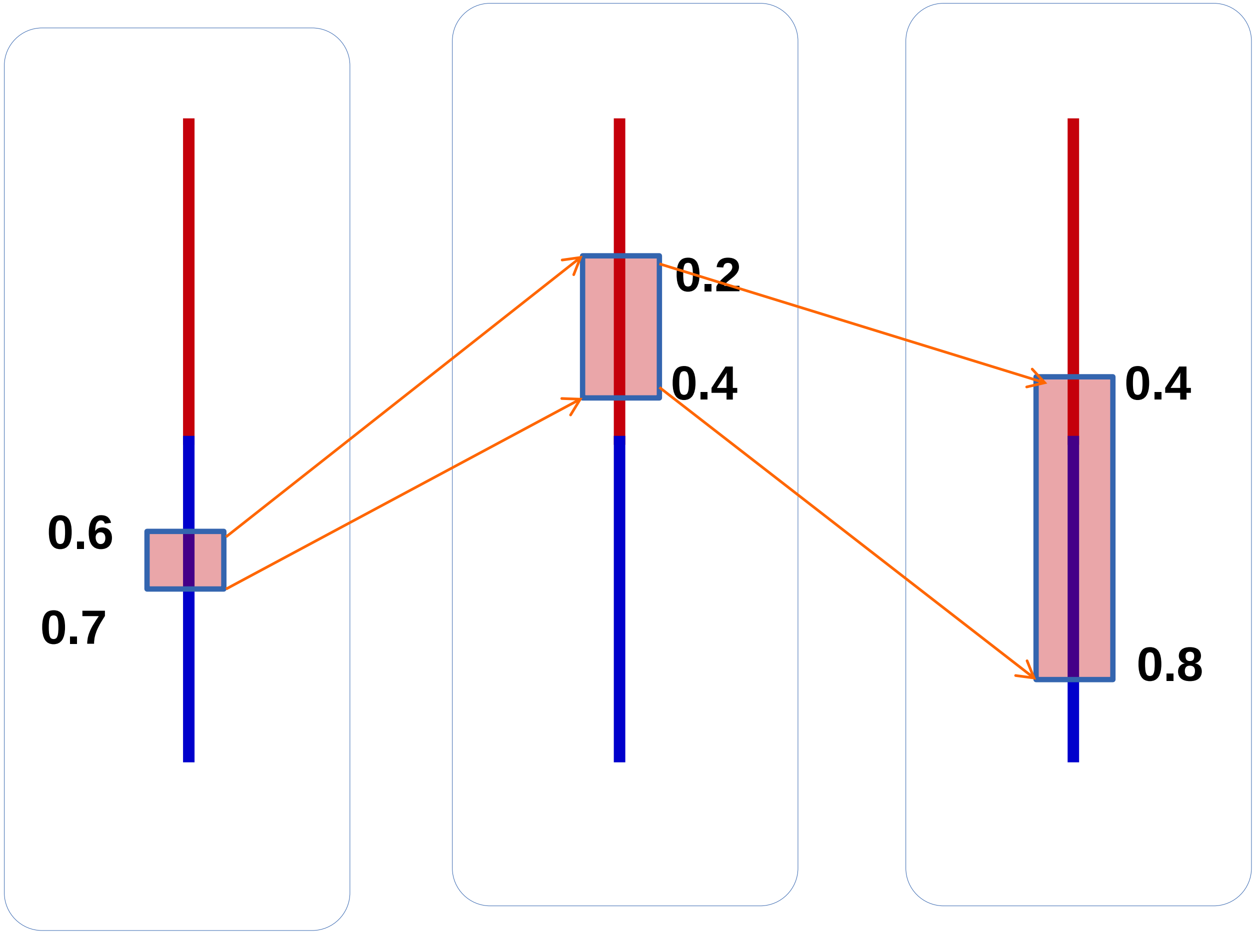
- The new range will be **0**<= **Range** <**1**
- Call this scaling "**E2**" , indicate it by binary "**1**"



E2 Scaling , Indicated by "1"

Binary Representation of Arithmetic Coding

<u>STOP when The mid point (0.5)</u> will lay inside the range



Scaling Code: E2,E1: 10

Binary Representation of Arithmetic Coding

Define Smallest number of Bis required for the Code

1. Find The Smallest Range in symbols is

2- Find the Min number of Bits required to store a Code <u>LESS than the Smallest Range</u>

Example: Given a long Sequence of characters A, B, and C
the probabilities of Characters are

- P(A)=0.8
- P(B)=0.02
- P(C)= 0.18
- Compress the following part of the sequence "**ACBA**" using Arithmetic Coding

Range: 0-0.8
,0.8-0.82,
0.82-1

Smallest Range in symbols is : 0.82-0.8=0.02

Min number of Bits required to store a Code LESS than the Smallest Range : 6 Bits

For Example :

(Smallest number represented in 6 bits is 0.000001 = 0.015625)

K=5

(note: for K=5; $1/2^5$ = 1/32 = 0.03125 > 0.02)

K=6

($1/2^6$ = 1/64 = 0.015625<0.02)

Example(1) Arithmetic Coding

Lower (Symbol) = Lower+ Range * Low_Range(Symbol)
Upper (Symbol)= Lower+ Range * High_Range(Symbol)

Low_range(A)=0
High_Range(A)=0.8

Low_range(B)=0.8
High_Range(B)=0.82

Low_range(C)=0.82
High_Range(C)=1.0

First Symbol is "A"

Lower(A)=0 ✓

Upper(A)=0.8

Second Symbol is "C"

Lower(C)=0+ (0.8-0)*0.82=0.656 ✓

Upper(C)=0+ (0.8- 0)*1= 0.8

Need E2 Scaling

1

Lower(C)= (0.656-0.5)*2= 0.312

Upper(C) = (0.8-0.5)*2 = 0.6

Third Symbol is "B"

Lower (B)= 0.312+ (0.6-0.312)*0.8=0.5424

Upper(B)=0.312+(0.6-0.312)*0.82=0.54816

Need E2 Scaling 1

Lower (B)= (0.5424-0.5)*2=0.0848

Upper(B)= (0.54816-0.5)*2=0.09632

Need E1 Scaling 0

Lower (B)= 0.0848 *2=0.1696

Upper(B)= 0.09632 *2= 0.19264

Need E1 Scaling 0

Lower (B) = 0.1696 *2= 0.3392

Upper(B)=0.19264*2= 0.33528

Need E1 Scaling 0

Lower(B)= 0.3392 *2= 0.6784

Upper(B)=0.33528 *2=0.77056

Need E2 Scaling 1

Lower(B)= (0.6784 -0.5) *2= 0.3568

Upper(B) = (0.77056-0.5) *2= 0.54112

✓

Fourth Symbol is "A"

Lower (A)= 0.3568+ (0.54112-0.3568)*0=0.3568

Upper(A)=0.3568+(0.54112-0.3568)*0.8=0.504256 ✓

0.5

Pick any Value in "A" Range

*0.5 in "**K**" Bits, K=6*

1 1 0 0 0 1 100000

Compressed Code is

1 1 0 0 0 1 1 0 0 0 0 0

Remember:
0.1 =
0.10 =
0.100 =
0.1000 =
0.10000 = …..

Which equal "0.5" in decimal

Compressed Code is equivalent to 0.7734375

Example(1) Arithmetic Coding (Decoding)

Compressed Code is $$0.\underbrace{1\ 1\ 0\ 0\ 0\ 1}_{\textit{Use first ``K'' bits}}\ 1\ 0\ 0\ 0\ 0\ 0$$

110001 (Binary) = (32+16+1) / 2^6 = 49/64 = 0.765625

$0 < 0.765625 < 0.8$ **<u>First Symbol is “A”</u>**

$$\left.\begin{array}{l}\text{Lower (A)} = 0 + (1-0) * 0 = 0 \\ \text{Upper (A)} = 0 + (1-0) * 0.8 = 0.8\end{array}\right\} ✓$$

Code= (0.765625 -0)/ (0.8 -0) = 0.957

$0.82 < 0.957 < 1.0$ **Second <u>Symbol is “C”</u>**

Example(1) Arithmetic Coding (Decoding)

$$\left.\begin{array}{l}\text{Lower (C)} = 0 + (0.8 - 0) * 0.82 = 0.656 \\ \text{Upper (C)} = 0 + (0.8 - 0) * 1.0 = 0.8\end{array}\right\} \text{Need E2 Scaling} \quad 1$$

$$\left.\begin{array}{l}\text{Lower (C)} = (0.656 - 0.5) *2 = 0.312 \\ \text{Upper (C)} = (0.8 - 0.5) *2 = 0.6\end{array}\right\} \quad ✓$$

Use Code“100011” $0.\,1\,\underbrace{1\,0\,0\,0\,1\,1}_{\textit{Shift ONE Bit, Use “}\boldsymbol{K}\textit{” bits}}\,0\,0\,0\,0\,0$

Example(1) Arithmetic Coding (Decoding)

Use Code“100011”

$$0.\boxed{1}\,\underbrace{1\ 0\ 0\ 0\ 1\ 1}_{\textit{Shift ONE Bit, Use “}\mathbf{K}\textit{” bits}}\ 0\ 0\ 0\ 0\ 0$$

100011 (Binary) = (32+2+1) / 2^6 = 35/64 = 0.546875

Code = (0.546875 – 0.312)/(0.6 – 0.312) = 0.815538

0.8 < 0.815538 < 0.82 Third Symbol is “B”

Lower (B) = 0.312 + (0.6 - 0.312)*0.8= 0.5424

Upper (B) = 0.312 + (0.6 - 0.312)*0.82= 0.54816

Need E2 Scaling

1

Lower (B) = (0.5424 -0.5)*2= 0.0848

Upper (B) = (0.54816 -0.5)*2=0.09632

Need E1 Scaling

0

Example(1) Arithmetic Coding (Decoding)

Use Code“100011” $0.\ 1\ \underbrace{1\ 0\ 0\ 0\ 1\ 1}_{\textit{Shift ONE Bit, Use “}\mathbf{K}\textit{” bits}}\ 0\ 0\ 0\ 0\ 0$

$$\left.\begin{array}{l}\text{Lower (B) = 0.0848*2=0.1696}\\ \text{Upper (B) = 0.09632*2=0.19264}\end{array}\right\}\ \text{Need E1 Scaling}\quad 0$$

$$\left.\begin{array}{l}\text{Lower(B)= 0.1696*2=0.3392}\\ \text{Upper(B)= 0.19264*2=0.38528}\end{array}\right\}\ \text{Need E1 Scaling}\quad 0$$

$$\left.\begin{array}{l}\text{Lower (B) = 0.3392*2= 0.6784}\\ \text{Upper(B) = 0.38528 *2= 0.77056}\end{array}\right\}\ \text{Need E2 Scaling}\quad 1$$

Example(1) Arithmetic Coding (Decoding)

$$\left.\begin{array}{l}\textbf{Lower (B) = (0.6784-0.5)*2 = 0.3568}\\ \textbf{Upper (B) = (0.77056 -0.5) *2= 0.54112}\end{array}\right\}\ \checkmark$$

Use Code"100011"

$$\mathbf{0.\ 1\ \boxed{1\ 0\ 0\ 0\ 1}\ \underbrace{1\ 0\ 0\ 0\ 0\ 0}_{\textit{Shift Five Bit, Use ``}\textbf{\textit{K}}\textit{'' bits}}}$$

100000 (Binary) = 32/64 = 0.5

Code = (0.5 - 0.3568) / (0.54112 - 0.3568) = 0.7769

0.0 < 0.7769 < 0.8 Fourth <u>Symbol is "A"</u>