

Arithmetic Coding

—

Exercise - decoding

Multimedia Technology

Tutorial 2, section 1b

Arithmetic Coding

An information source is encoded using arithmetic coding based on the interval table A: $[0,0.4)$, B: $[0.4,0.6)$, Г: $[0.6,0.8)$, \$ (terminal symbol): $[0.8,1)$.

If the output of the encoder is the decimal number 0.089, calculate what the input string of the encoder was.

Based on the information I have can I make any assumptions about the string?

No, because the exercise does not tell me that the intervals reflect the actual probability of the symbols appearing in the input.

If the exercise stated that these were the appearance probabilities in the input, then I could guess that the string has the same count of B, C, \$, each with probability 0.2, and double that of A with probability 0.4.

Since \$ is the terminal symbol and therefore appears once, I would have two A's, one B, one C (in unknown order), and the \$ (terminal) at the end.

The intervals give us the encoding ranges for each symbol.

Decoding Algorithm

```
input n;  
repeat {  
    find s so that n is in  
        [l[s], h[s]);  
    output s;  
    r = h[s] - l[s];  
    n = (n - l[s]) / r;  
} until s = $;
```

$L[s]$ and $h[s]$

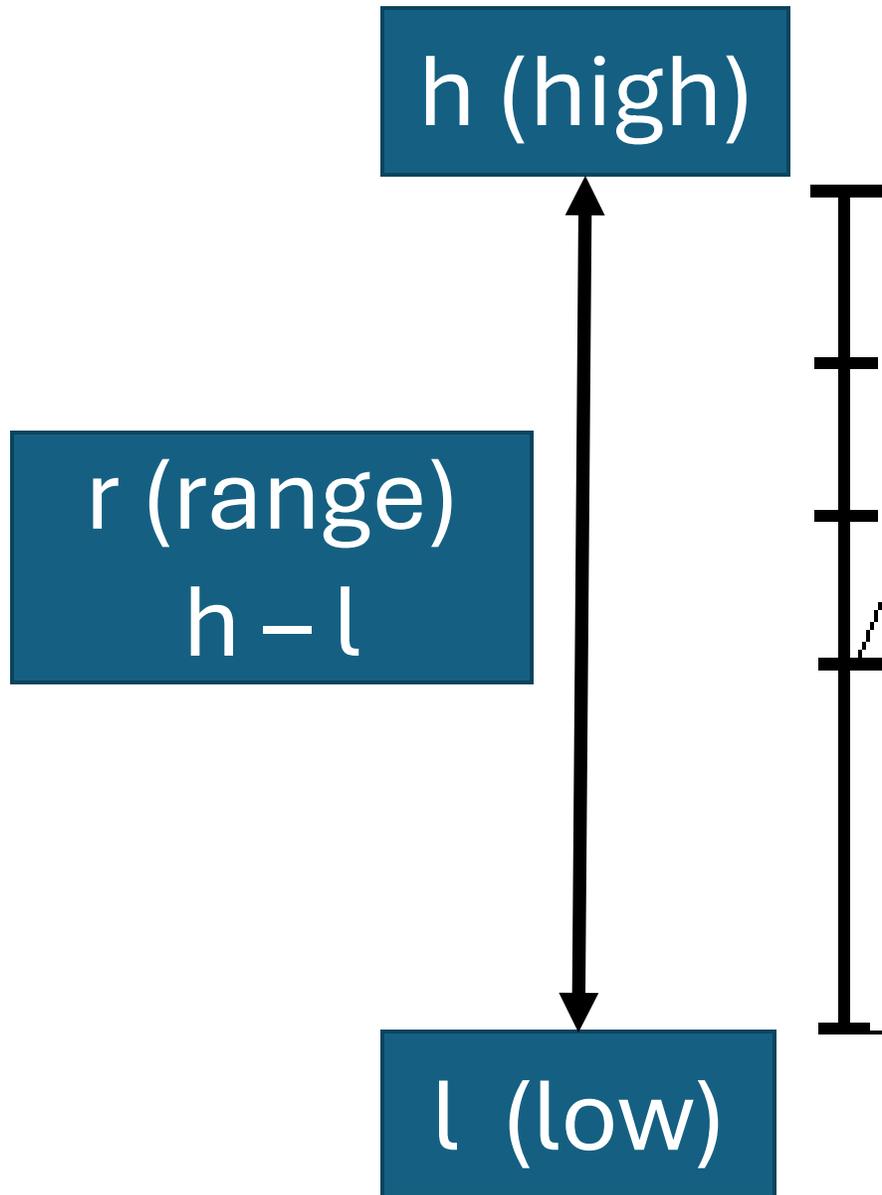
Lower and upper bounds of the interval for each symbol s .

$A: [0, 0.4)$

$l[A]$

$h[A]$

high, low and range



Current values for the upper bound, range, and lower bound of the interval as we decode each symbol in the encoded sequence

Step 1

- $n = 0.089$
- Belongs in the interval for A: $[0, 0.4)$
- Thus, output: A
- Find new value for n
- $r = 0.4 - 0 = 0.4$
- $n = (0.089 - 0) / 0.4 = 0.2225$

Step 2

- $n = 0.2225$
- Belongs in the interval for A: $[0, 0.4)$
- Thus, output: A
- Find new value for n
- $r = 0.4 - 0 = 0.4$
- $n = (0.2225 - 0) / 0.4 = 0.55625$

Step 3

- $n = 0.55625$
- Belongs in the interval for B: $[0.4, 0.6)$
- Thus, output: B
- Find new value for n
- $r = 0.6 - 0.4 = 0.2$
- $n = (0.55625 - 0.4) / 0.2 = 0.78125$

Step 4

- $n = 0.78125$
- Belongs in the interval for C: $[0.6, 0.8)$
- Thus, output: C
- Find new value for n
- $r = 0.8 - 0.6 = 0.2$
- $n = (0.78125 - 0.6) / 0.2 = 0.90625$

Step 5

- $n = 0.90625$
- Belongs in the interval for $\$$: $[0.8, 0.1)$
- We found $\$$ - this symbol signifies the end of the algorithm
- We don't need to continue the calculations
- The initial string was: AABC
- The $\$$ is a terminal symbol, it signifies the end of the sequence, it is not part of the string