Problem 2. Given the sequence *cat*Δ*ate*Δ*hat*, encode the sequence using *ppma* algorithm with $N = 1$ and an adaptive arithmetic coder. Assuming a six-letter alphabet {*h*, *e*, *t*, *a*, *c*, Δ}.

Assume the word length of the AC is 6. Thus, initially $l = 000000$ and $u = 111111$.

The Final context table after encoding all letters is:

order	context	symbol occurrence counts							Total
		h	*e*	*t*	*a*	*c*	Δ	<*ESC*>	
1	*c*	0	0	0	1	0	0	1	2
	a	0	0	3	0	0	0	1	4
	t	0	1	0	0	0	1	1	3
	Δ	1	0	0	1	0	0	1	3
	e	0	0	0	0	0	1	1	2
	h	0	0	0	1	0	0	1	2
0	/	1	1	3	3	1	2	1	12
–1	/	1	1	1	1	1	1	0	6

Symbol 1: *c*

No 1st-order or zero order contexts yet, go directly to –1 order context.

Use –1 order context to encode *c*, $l = 101010$, and $u = 110100 \rightarrow E_2$ scale.

Transmitted sequence: 1.

Updated bounds: $l = 010100$, $u = 101001 \rightarrow E_3$ scale.

Updated bounds: $l = 001000$, $u = 110011$, E_3 count = 1.

Symbol 2: *a*

No 1st-order context of *c* yet, goes directly to 0 order context.

Encode <*ESC*>, $l = 011110$, and $u = 110011$.

Use –1 order context to encode *a*, $l = 101001$, and $u = 101011 \rightarrow E_2$ scale.

Transmitted sequence: 110.

Updated bounds: $l = 010010$, $u = 010111$, E_3 count = 0 $\rightarrow E_1$ scale.

Transmitted sequence: 1100.

Updated bounds: $l = 100100$, $u = 101111 \rightarrow E_2$ scale.

Transmitted sequence: 11001.

Updated bounds: $l = 001000$, $u = 011111 \rightarrow E_1$ scale.

Transmitted sequence: 110010.

Updated bounds: $l = 010000$, $u = 111111$.

Symbol 3: *t*

No 1st-order context of *a* yet, goes directly to 0 order context.

Encode <*ESC*>, $l = 110000$, and $u = 111111 \rightarrow E_2$ scale twice.

Transmitted sequence: 11001011.

Updated bounds: $l = 000000$, $u = 111111$.

Use –1 order context to encode t, l = 010101, and u = 011111 → $E_1 + E_2$ scales.
Transmitted sequence: 1100101101.
Updated bounds: l = 010100, u = 111111.

Symbol 4: Δ
No 1st-order context of t yet, goes directly to 0 order context.
Encode <*ESC*>, l = 110101, and u = 111111 → E_2 scale twice.
Transmitted sequence: 110000111.
Updated bounds: l = 010100, u = 111111.
Use –1 order context to encode Δ, l = 111000, u = 111111 → E_2 scale three times.
Transmitted sequence: 110000111111.
Updated bounds: l = 000000, u = 111111.

Symbol 5: a
No 1st-order context of Δ yet, goes directly to 0 order context.
Use 0 order context to encode a, l = 001100, and u = 011000 → E_1 scale.
Transmitted sequence: 1100001111110.
Updated bounds: l = 011000, u = 110001.

Symbol 6: t
Use 1st order context of a to encode t, l = 011000, and u = 100100 → E_3 scale.
Updated bounds: l = 010000, u = 101001, E_3 count = 1 → E_3 scale.
Updated bounds: l = 000000, u = 110011, E_3 count = 2.

Symbol 7: e
No e in 1st order context of t, encode <*ESC*>, l = 011010, and u = 110011.
Encode <*ESC*> for 0 order context, l = 110000, and u = 110011 → E_2 scale twice.
Transmitted sequence: 11000011111101100.
Updated bounds: l = 000000, u = 001111, E_3 count = 0 → E_1 scale twice.
Transmitted sequence: 1100001111110110000.
Updated bounds: l = 000000, u = 111111.
Use –1 order context to encode e, l = 001010, and u = 010100 → E_1 scale.
Transmitted sequence: 11000011111101100000.
Updated bounds: l = 010100, u = 101001 → E_3 scale.
Updated bounds: l = 000000, u = 110011, E_3 count = 1.

Symbol 8: Δ
No 1st-order context of e yet, goes directly to 0 order context.
Use 0 order context to encode Δ, l = 100111, u = 101100 → E_2 scale.
Transmitted sequence: 1100001111110110000010.
Updated bounds: l = 001110, u = 011001, E_3 count = 0 → E_1 scale.

Transmitted sequence: 110000111111011000000100.

Updated bounds: $l = 011100$, $u = 110011$.

Symbol 9: *h*

No *h* in the 1st order context of Δ, encode <*ESC*> using the conditional probabilities of the context Δ, $l = 101000$, and $u = 110011 \rightarrow E_2$ scale.

Transmitted sequence: 1100001111110110000001001.

Updated bounds: $l = 010000$, $u = 100111 \rightarrow E_3$ scale.

Updated bounds: $l = 000000$, $u = 101111$, E_3 count = 1.

No 0 order context for *h*, encode <*ESC*>, $l = 101010$, $u = 101111 \rightarrow E_2$ scale.

Transmitted sequence: 110000111111011000000100110.

Updated bounds: $l = 010100$, $u = 011111$, E_3 count $= 0 \rightarrow E_1$ scale.

Transmitted sequence: 1100001111110110000001001100.

Updated bounds: $l = 101000$, $u = 111111 \rightarrow E_2$ scale.

Transmitted sequence: 11000011111101100000010011001.

Updated bounds: $l = 010000$, $u = 111111$.

Use –1 order context to encode *h*, $l = 010000$, $u = 010111 \rightarrow E_1 + E_2 + E_1$ scales.

Transmitted sequence: 11000011111101100000010011001010.

Updated bounds: $l = 000000$, $u = 111111$.

Symbol 10: *a*

No 1st-order context of *h* yet, create an entry for context '*h*', then go directly to 0 order context.

Use 0 order context to encode *a*, $l = 011001$, and $u = 011111 \rightarrow E_1 + E_2 + E_2$ scales.

Transmitted sequence: 11000011111101100000010011001010011.

Updated bounds: $l = 001000$, $u = 111111$.

Symbol 11: *t*

Use 1st-order context of *a* to encode *t*, $l = 001000$, $u = 101100$.

We transmit the lower bound to signal the final letter *t*.

Transmitted sequence: 11000011111101100000010011001010011001000.

Problem 4. A sequence is encoded using the Burrows-Wheeler transform. Given *L* = *elbkkee*, and index = 5 (we start counting from 1, not 0), find the original sequence.

The decoded sequence is *kelebek*.