LZ78 – Exercise

Multimedia Technology

Tutorial 2, section 3a

LZ78 - exercise

Consider that we apply LZ78 encoding and start with an empty dictionary.

Show how the sequence "aababbbbbbaa" is encoded and how the output produced by the encoder is decoded.

aababbbbbaa

In	Out	Dictionary
a		

The first symbol I see is 'a'.

aababbbbbaa

In	Out	Dictionary
a	(0 , a)	

The output is (0,a)

0 indicates that I have not seen 'a' before.

aababbbbbaa

In	Out	Dictionary
a	(0 , a)	1: a

Add symbol 'a' to the Dictionary at index 1.

aababbbbbaa

In	Out	Dictionary
a	(0 , a)	1: a
ab		

The next string I see is 'ab'.
Symbol 'a' is the only symbol in the string I have seen before.

aababbbbbaa

In	Out	Dictionary
a	(0 , a)	1: a
ab	(1 , b)	

The output is: (1 , b)

1 in the code above is the index 1 of the Dictionary, meaning a.

'b' is the next symbol.

aababbbbbaa

In	Out	Dictionary
a	(0 , a)	1: a
ab	(1 , b)	2: ab

Add 'ab' in index 2 of the Dictionary.

aababbbbbaa

In	Out	Dictionary
a	(0 , a)	1: a
ab	(1 , b)	2: ab
abb		

The next string is abb.

I have seen 'a' and 'ab' before.

aababbbbbaa

In	Out	Dictionary
a	(0 , a)	1: a
ab	(1 , b)	2: ab
abb	(2 , b)	

The output is: (2 , b)

2 is the index of the Dictionary, meaning 'ab'

The next symbol is 'b'

aababbbbbaa

In	Out	Dictionary
a	(0 , a)	1: a
ab	(1 , b)	2: ab
abb	(2 , b)	3: abb

Add 'abb' at index 3 in the Dictionary.

aababbbbbaa

In	Out	Dictionary
a	(0 , a)	1: a
ab	(1 , b)	2: ab
abb	(2 , b)	3: abb
b		

The next string I see is 'b'.

aababbbbbaa

In	Out	Dictionary
a	(0 , a)	1: a
ab	(1 , b)	2: ab
abb	(2 , b)	3: abb
b	(0 , b)	

The output is: (0 , b)

0 indicates that I have not seen 'b' before.

aababbbbbaa

In	Out	Dictionary
a	(0 , a)	1: a
ab	(1 , b)	2: ab
abb	(2 , b)	3: abb
b	(0 , b)	4: b

Add 'b' at index 4 in the Dictionary.

aababbbbbaa

In	Out	Dictionary
a	(0 , a)	1: a
ab	(1 , b)	2: ab
abb	(2 , b)	3: abb
b	(0 , b)	4: b
bb		

The next string is 'bb'.
I have only seen the symbol 'b' before, not the string 'bb'.

aababbbbbaa

In	Out	Dictionary
a	(0 , a)	1: a
ab	(1 , b)	2: ab
abb	(2 , b)	3: abb
b	(0 , b)	4: b
bb	(4 , b)	

The output is: (4 , b)

4 indicates the index of the Dictionary, meaning 'b'.

The next symbol is 'b'.

aababbbbbaa

In	Out	Dictionary
a	(0 , a)	1: a
ab	(1 , b)	2: ab
abb	(2 , b)	3: abb
b	(0 , b)	4: b
bb	(4 , b)	5: bb

Add 'bb' at index 5 in the Dictionary.

aababbbbbbaa

In	Out	Dictionary
a	(0 , a)	1: a
ab	(1 , b)	2: ab
abb	(2 , b)	3: abb
b	(0 , b)	4: b
bb	(4 , b)	5: bb
aa		

The next string is 'aa'.

aababbbbbaa

In	Out	Dictionary
a	(0 , a)	1: a
ab	(1 , b)	2: ab
abb	(2 , b)	3: abb
b	(0 , b)	4: b
bb	(4 , b)	5: bb
aa	(1 , a)	

The output is: (1 , α)

1 indicates the index 1 from the Dictionary.

The next symbol is 'a'.

aababbbbbbaa

In	Out	Dictionary
a	(0 , a)	1: a
ab	(1 , b)	2: ab
abb	(2 , b)	3: abb
b	(0 , b)	4: b
bb	(4 , b)	5: bb
aa	(1 , a)	6: aa

Add 'aa' at index 6 in the Dictionary.

Decode the output from the encoder.

In	Out	Dictionary
		1: a
		2: ab
		3: abb
		4: b
		5: bb
		6: aa

When decoding I know the dictionary values.
The input is the code.
I have to caclulate the output.

In	Out	Dictionary
(0 , a)		1: a
		2: ab
		3: abb
		4: b
		5: bb
		6: aa

Input: (0 , a)

a

In	Out	Dictionary
(0 , a)	a	1: a
		2: ab
		3: abb
		4: b
		5: bb
		6: aa

The output is a.

a

In	Out	Dictionary
(0 , a)	a	1: a
(1 , b)		2: ab
		3: abb
		4: b
		5: bb
		6: aa

Input: (1 , b)

aab

In	Out	Dictionary
(0 , a)	a	1: a
(**1** , **b**)	ab	2: ab
		3: abb
		4: b
		5: bb
		6: aa

If we check index 1 in the Dictionary, we have 'a'
We add b from the input code
The output is ab

aab

In	Out	Dictionary
(0 , a)	a	1: a
(1 , b)	ab	2: ab
(2 , b)		3: abb
		4: b
		5: bb
		6: aa

Input: (2 , b)

aababb

In	Out	Dictionary
(0 , a)	a	1: a
(1 , b)	ab	2: ab
(2 , b)	abb	3: abb
		4: b
		5: bb
		6: aa

When we check the Dictionary at index 2, we have 'ab'.

Add b from the input code.

The output is abb

aababb

In	Out	Dictionary
(0 , a)	a	1: a
(1 , b)	ab	2: ab
(2 , b)	abb	3: abb
(0 , b)		4: b
		5: bb
		6: aa

Input: (0 , b)

aababb

In	Out	Dictionary
(0 , a)	a	1: a
(1 , b)	ab	2: ab
(2 , b)	abb	3: abb
(0 , b)		4: b
		5: bb
		6: aa

0 indicates that I had not seen 'b' before, when I was assigning that code.

aababbb

In	Out	Dictionary
(0 , a)	a	1: a
(1 , b)	ab	2: ab
(2 , b)	abb	3: abb
(0 , b)	b	4: b
		5: bb
		6: aa

The output is 'b'.

aababbb

In	Out	Dictionary
(0 , a)	a	1: a
(1 , b)	ab	2: ab
(2 , b)	abb	3: abb
(0 , b)	b	4: b
(4 , b)		5: bb
		6: aa

Input: (4 , b)

aababbbbb

In	Out	Dictionary
(0 , a)	a	1: a
(1 , b)	ab	2: ab
(2 , b)	abb	3: abb
(0 , b)	b	4: b
(4 , b)	bb	5: bb
		6: aa

If I check index 4 in the Dictionary, I find 'b'.
Add b from the input code.
The output is bb

aababbbbb

In	Out	Dictionary
(0 , a)	a	1: a
(1 , b)	ab	2: ab
(2 , b)	abb	3: abb
(0 , b)	b	4: b
(4 , b)	bb	5: bb
(1 , a)		6: aa

Input: (1 , a)

aababbbbba

In	Out	Dictionary
(0 , a)	a	1: a
(1 , b)	ab	2: ab
(2 , b)	abb	3: abb
(0 , b)	b	4: b
(4 , b)	bb	5: bb
(1 , a)	aa	6: aa

When I check the Dictionary at index 1, I have 'a'.
Add a from the input code.
Output: aa