Tutorial 8 Exercise Solutions

COMP526: Efficient Algorithms

25–26 November, 2024

Exercise 1. What is the result of applying the Lempel-Ziv-Welch (LZW) compression scheme to the text $S = \texttt{ABABABACABABA}$ with alphabet $\Sigma = \{\texttt{A}, \texttt{B}, \texttt{C}\}$ using codewords 4 bits? Write both the encoded text and the dictionary when the procedure terminates.

Solution. Using the LZW encoding algorithm, the binary text encodes to
`000000010011010100100101010 0`
In a more human-readable format, this is the list $0, 1, 3, 5, 2, 5, 4$. When the execution terminates, the dictionary contains the following values:

codeword	phrase
0000	A
0001	B
0010	C
0011	AB
0100	BA
0101	ABA
0110	ABAC
0111	CA
1000	ABAB

□

Exercise 2. Use the LZW decoding algorithm to decode the encoded text
`000101000100001000110000 1001`
(or as a decimal list $1, 4, 4, 2, 3, 0, 9$) where the alphabet is $\Sigma = \{\texttt{!}, \texttt{A}, \texttt{G}, \texttt{H}\}$ and the codeword length is 4 bits. Also record the dictionary after decoding the text.

Solution. The coded text decodes to `AAAAAGH!!!`. The dictionary's contents is

codeword	phrase
0000	!
0001	A
0010	G
0011	H
0100	AA
0101	AAA
0110	AAG
0111	GH
1000	H!
1001	!!

□

Exercise 3. Use the inverse move-to-front transform to decode the encoded text $1,2,3,1,4,4,2,2,2$ using the alphabet $\Sigma = \{\texttt{A},\texttt{C},\texttt{H},\texttt{I},\texttt{U}\}$. Write the state of the alphabet after each decoded letter

Solution. The decoded text is `CHIHUAHUA`. The full execution is depicted below.

index	decoded character	alphabet
		ACHIU
1	C	CAHIU
2	H	HCAIU
3	I	IHCAU
1	H	HICAU
4	U	UHICA
4	A	AUHIC
2	H	HAUIC
2	U	UHAIC
2	A	AUHIC

□

Exercise 4. Use the inverse Burrows-Wheeler transform to decode the text `dnenb$aaraab`

Solution. Recall that to apply the inverse Burrows-Wheeler transform, we first form character-index pairs:

$$(d,0)$$
$$(n,1)$$
$$(e,2)$$
$$(n,3)$$
$$(b,4)$$
$$(\$,5)$$
$$(a,6)$$
$$(a,7)$$
$$(r,8)$$
$$(a,9)$$
$$(a,10)$$
$$(b,11)$$

Sorting this by first element gives

0	$(\$, 5)$
1	$(a, 6)$
2	$(a, 7)$
3	$(a, 9)$
4	$(a, 10)$
5	$(b, 4)$
6	$(b, 11)$
7	$(d, 0)$
8	$(e, 2)$
9	$(n, 1)$
10	$(n, 3)$
11	$(r, 8)$

Using the second entries of the pairs as "links" and following the linked list starting from \$ gives, `bananabread$`. □