

מבני נתונים ואלגוריתמים

אלגוריתם למפל זיו

המילון שלנו בהתחלה הוא:

$$\begin{aligned} a &= 00 \\ b &= 01 \\ c &= 10 \\ d &= 11 \end{aligned}$$

באופן כללי:

1. כל מילה חדשה מקודדים כזוג סדור (j, k) כאשר j מס' המילה שקיימת k האות הנוספת החדשה.

2. הופכים את הזוגות הסדורים למחרוזת בינארית - את המספר j נכתוב כמספרים בינאריים עם $\log_2 i$ אותיות כאשר i הוא המספר הסידורי של המילה ואת האות k לפי המילון.

נקודת את המילה:

aabccdaabcdba

המילים שלנו יהיו:

i	$\lceil \log_2 i \rceil$	מילה	חלק במחרוזת	קידוד בבינארית i בבינארית ב $\lceil \log_2 i \rceil$ ספרות ואז הקוד של האות)
1	0	$(0, a)$	<i>a</i>	00
2	1	$(1, b)$	<i>ab</i>	101
3	2	$(0, c)$	<i>c</i>	0010
4	2	$(3, d)$	<i>cd</i>	1111
5	3	$(1, a)$	<i>aa</i>	00100
6	3	$(0, b)$	<i>b</i>	00001
7	3	$(4, b)$	<i>cdb</i>	10001
8	3	(1)	<i>a</i>	001

הקידוד של המחרוזת יהיה:

0010100101111001000000110001001

כעת לשחזר. נניח שיש לנו את המחרוזת המקודדת הנ"ל.

ניקח $i = 1$ עד שנגמרת המחרוזת,

$\lceil \log_2 i \rceil$ - המספר הסידורי,

מס' ספרות k ע"פ המחרוזת.

במקרה שלנו $k = 2$.

i	$\lceil \log_2 i \rceil$	מילה במחרוזת המקודדת $(k + \log_2 i)$ ספרות)	מילה משוחזרת
1	0	$, 00 = (0, a)$	<i>a</i>
2	1	$1, 01 = (1, b)$	<i>ab</i>
3	2	$00, 10 = (0, c)$	<i>c</i>
4	2	$11, 11 = (3, d)$	<i>cd</i>
5	3	$001, 00 = (1, a)$	<i>aa</i>
6	3	$000, 01 = (0, b)$	<i>b</i>
7	3	$100, 01 = (4, b)$	<i>cdb</i>
8	3	$001, = (1)$	<i>a</i>

המילה המשוחזרת היא:

aabccdaabcdba

שאלה בנושא דחיסה:

מצאו מחרוזת (באורך 1000 תווים לפחות) שקידוד הומפן מצליח לדחוס, ולעומת זאת למפל-זיו כמעט ולא דחס אותה.